# IFMBE Proceedings

Nicos Maglaveras · Ioanna Chouvarda
Paulo de Carvalho (Eds.)

Volume 66

# Precision Medicine Powered by pHealth and Connected Health

ICBHI 2017, Thessaloniki, Greece, 18–21 November 2017


Springer

# IFMBE Proceedings

Volume 66

The International Federation for Medical and Biological Engineering, IFMBE, is a federation of national and transnational organizations representing internationally the interests of medical and biological engineering and sciences. The IFMBE is a non-profit organization fostering the creation, dissemination and application of medical and biological engineering knowledge and the management of technology for improved health and quality of life. Its activities include participation in the formulation of public policy and the dissemination of information through publications and forums. Within the field of medical, clinical, and biological engineering, IFMBE's aims are to encourage research and the application of knowledge, and to disseminate information and promote collaboration. The objectives of the IFMBE are scientific, technological, literary, and educational.

The IFMBE is a WHO accredited NGO covering the full range of biomedical and clinical engineering, healthcare, healthcare technology and management. It is representing through its 60 member societies some 120.000 professionals involved in the various issues of improved health and health care delivery.



More information about this series at http://www.springer.com/series/7403

Nicos Maglaveras · Ioanna Chouvarda
Paulo de Carvalho
Editors

# Precision Medicine Powered by pHealth and Connected Health

ICBHI 2017, Thessaloniki, Greece, 18–21 November 2017

Springer

*Editors*
Nicos Maglaveras
Department of Medicine
Aristotle University of Thessaloniki
Thessaloniki
Greece

and

Department of Industrial Engineering and Management Sciences, Department of Electrical Engineering and Computer Science, McCormick School of Engineering and Applied Science, Center for Engineering and Health
Northwestern University
Evanston, IL
USA

Ioanna Chouvarda
Department of Medicine
Aristotle University of Thessaloniki
Thessaloniki
Greece

Paulo de Carvalho
Department of Informatics Engineering
University of Coimbra
Coimbra
Portugal

ISSN 1680-0737 ISSN 1433-9277 (electronic)
IFMBE Proceedings
ISBN 978-981-10-7418-9 ISBN 978-981-10-7419-6 (eBook)
https://doi.org/10.1007/978-981-10-7419-6

Library of Congress Control Number: 2017959161



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Preface

The IFMBE WG on Health Informatics and eHealth and Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University of Thessaloniki, have organized the ICBHI 2017 in Thessaloniki on November 18–21, 2017.

The ICBHI 2017 conference is the third in a series of scientific events initiated from the IFMBE Working Group on Health Informatics and eHealth bringing together expertise from medical, technological, biodata sciences, regulatory, and public health domains.

The area of biomedical and health informatics is exploding at all scales. The developments in the areas of medical devices (MDD), eHealth, and personalized health (pHealth) as enabler factors for the evolution of precision medicine are quickly developing and demand the development of new scaling tools and integration frameworks and methodologies in the following areas

- Medical devices
- Biomedical environmental and behavioral data acquisition
- Biodata management and analytics
- Advanced mHealth-inspired applications
- User interaction systems and data/information exchange
- Applied connected health and integrated care systems
- Personalized health systems
- Preventive and multimorbid disease management tools
- Patient safety and efficient public health systems
- In silico modeling of pathophysiological mechanisms
- Translational research
- Biomedical informatics
- Biomedical informatics and technology education

It is expected that ICBHI 2017 will catalyze the personal health systems development methodology thus achieving convergence and integration mechanisms of the above-mentioned main scientific areas for the enabling of precision medicine. The need for managing the new findings and capabilities as well as integrating competencies and thematic areas are absolutely necessary for the area of biomedical and health informatics.

Taking into account the current needs in the biomedical and health informatics area, the conference streamlines in two main axes. The first axis includes biosensors, mHealth/uHealth components and systems, biorobotics, electroceuticals and micro-bio-nano-recording, and bioparametric estimation and filtering. The second axis includes machine learning, multiparametric analytics, bioinformatics, big biodata/biomedical computing, medical decision support, patient safety informatics, interoperability, and security. Approaches that can be used for integrating these two streams linked to precision medicine applications were addressed.

The conference featured a scientific challenge which revolved around the recent high interest in the coordinated care theme. The interest is highly related to the increase in aged population and the high numbers of multimorbid patients. It focused on cardiorespiratory

problems related with multimorbid COPD patients. The goals of the challenge focused on detection of wheezes and crackles in respiratory sounds. The scientific challenge rules followed the Physionet challenge rules (http://bhichallenge.med.auth.gr/).

This proceedings volume covers all presentations at the ICBHI 2017. All submissions have been carefully and critically reviewed by at least two independent experts from other than the authors' home countries and additionally by at least one member of the Scientific Program Committee. The editors are indebted to the acknowledged and highly experienced reviewers for having essentially contributed to the quality of the conference and the book at hand. Both the ICBHI 2017 Conference and the publication of the proceedings by Springer would not have been possible without the support and sponsor of the IFMBE and the WG on Health Informatics and eHealth. The editors are also grateful to the dedicated efforts of the Local Organizing Committee members and their supporters for carefully and smoothly preparing and operating the conference. They especially thank all the team members from AUTH Lab of Computing Medical Informatics and Biomedical-Imaging Technologies for their dedication to the event.

Thessaloniki, Greece — Nicos Maglaveras
Thessaloniki, Greece — Ioanna Chouvarda
Coimbra, Portugal — Paulo de Carvalho

# Committees

## Organization

*Conference Chairperson*
Prof. Nicos Maglaveras, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
*Conference Secretary*
Ms. Maria Giaga, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Ms. Lina Nikolopoulou, Mindwork Business Solutions, Greece

## IFMBE Administrative Council Officers

*President*
Prof. James Goh, National University of Singapore, Singapore
*Vice President*
Prof. Shankar Krishnan, Boston, USA
*Past President*
Prof. Ratko Magjarević, University of Zagreb, Croatia
*Secretary-General*
Prof. Kang Ping Lin, Chun-Yuan Christian University, Taiwan
*Treasurer*
Prof. Marc Nyssen, Vrije Universiteit Brussel, Belgium

## IFMBE WG on Health Informatics and eHealth

*Chair*
Paulo de Carvalho
*Co-Chair*
Ratko Magjarevic
*Advisor*
Yuan-Ting Zhang
*Members*
Shankar Krishnan
Zhi-Pei Liang
Nicos Maglaveras
Marc Nyssen
Esteban Pino
Nitish Thakor
May Wang
Bruce Wheeler

## International Scientific Program Committee

Ioanna Chouvarda, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Anthony H. Aletras, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Paolo Bonato, Harvard Medical School, USA
Electra Gizeli, IMBB-FORTH/Department of Biology, University of Crete, Greece
Mark van Gils, VTT, Technical Research Centre, Finland
Anna Bianchi, Polytecnico di Milano, Italy
Elske Ammenwerth, Institute for Biomedical Informatics, UMIT, Austria
Spyros Kitsiou, University of Illinois Chicago, USA
Catherine Chronaki, HL7 International Foundation, Greece
Vassilis Koutkias, Center for Research and Technology—Institute of Applied Biosciences, Greece
Dimitris Kougioumtzis, Department of Electrical and Computer Engineering, Aristotle University, Greece
John Mantas, EFMI, University of Athens, Greece
Rita Paradiso, SMARTEX, Italy
Paulo Carvalho, Universidad de Coimbra, Portugal
Jorge Henriquez, Universidad de Coimbra, Portugal
Alan V. Sahakian, Northwestern University, USA
May Wang, Georgia Institute of Technology, USA
Y. T. Zhang, Chinese University of Hong Kong, China
Konstantina Nikita, NTUA, Greece
Dimitris Fotiadis, University of Ioannina, Greece
Maria Teresa Arredondo, Universidad Politécnica de Madrid, Spain
Pantelis Angelidis, University of Western Macedonia, President of the Thessaloniki Innovation Zone

## Local Organizing Committee

Ioanna Chouvarda, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Anthony H. Aletras, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Anastasios Delopoulos, Department of Electrical and Computer Engineering, Aristotle University, Greece
Eleutherios Angelis, Department of Computer Science, Aristotle University, Greece
Theodoros Agorastos, Department of Medicine, Aristotle University, Greece
Vassilis Koutkias, Center for Research and Technology—Institute of Applied Biosciences, Greece
Andreas Triantafyllidis, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Konstantinos Votis, Center for Research and Technology—Information Technologies Institute (CERTH/ITI)
Christos Maramis, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Vassilis Kilintzis, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece
Katerina Lourida, Lab of Computing Medical Informatics and Biomedical-Imaging Technologies, Aristotle University, Greece

# Contents

**Part III eHealth Systems, Services and Cloud Computing**



**Part IV Biosignals and Biomarkers**

# Part I
# Big Data Analytics for Precision Medicine

# Hybrid Hierarchical Clustering Algorithm Used for Large Datasets: A Pilot Study on Long-Term Sleep Data

V. Gerla, M. Murgas, A. Mladek, E. Saifutdinova, M. Macas, and L. Lhotska

**Abstract**

Clustering is a popular analysis technique in a modern science full of unlabeled data, hidden dependencies and relations between elements in datasets. The presented study proposes a new hybrid hierarchical clustering method suitable for large datasets. It is based on the combination of effective simple methods. The proposed method was tested and compared with a widely used agglomerative clustering method. Two groups of datasets were used for testing. The first group contains data delivered from real biomedical data and related to a real problem of indication of sleep stages. The second group consists of artificially generated large data. Time, memory consumption, and mutual information were compared.

**Keywords**

Hybrid • Hierarchical • Clustering • Sleep • EEG

## Introduction

Cluster analysis represents one of the machine learning methods used to find data labels. There are two basic methods, *k*-mean clustering and hierarchical approaches. An algorithm for solving *k*-means problem was proposed by Lloyd in the paper [1], and it is still widely used [2, 3]. The second approach, hierarchical clustering, is an effective tool for data analysis due to its ability to utilize hierarchical structures of data, however, its disadvantage—huge memory requirements—prevents it from being used for large datasets [4].

V. Gerla (✉) · M. Murgas · A. Mladek · E. Saifutdinova · M. Macas · L. Lhotska
Czech Institute of Informatics, Robotics, Cybernetics, Czech Technical University in Prague, Jugoslavskych Partyzanu 1580/3, Prague, Czech Republic
e-mail: vaclav.gerla@cvut.cz

E. Saifutdinova
National Institute of Mental Health, Prague, Czech Republic

L. Lhotska
Faculty of Biomedical Engineering, Czech Technical University in Prague, Prague, Czech Republic

Cluster analysis is also an important tool in biomedicine research. It may help to reveal hidden information in huge amount of data valuable for identification of important states, trends, and patterns. The amount of stored large-scale and often heterogeneous data produced by various devices, sensors, networks, transactional applications, etc. is increasing exponentially [5]. The conventional technologies and methods, available to store and analyse the data do not work efficiently with such an amount of them.

There are new effective hierarchical schemes published recently to address these drawbacks and used for various type of datasets. These include, for example, spectral clustering algorithms combined with hierarchical method [6], hierarchical clustering linkage criterion called Genie [7] or new linkage method for hierarchical clustering named NC-link [4].

The aim of the study is to develop a combined approach exploiting the advantages of the above-mentioned approaches, k-means algorithm and hierarchical clustering. The paper investigates properties of proposed method in terms of time and memory consumption.



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_1

## Data

### Sleep EEG Dataset

To test the algorithms, two real datasets were used. Each dataset is formed by 30 s segments of full-night polysomnography. Each segment includes 19 EEG channels and was scored by American Association of Sleep Medicine [8] manual by a trained clinician. The data were obtained in the Bulovka Hospital, Czech Republic. Sample frequency is 250 Hz. Data were recorded using Fp, Fp2, F3, F4, C3, C4, P3, P4, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz, O1 and O2 electrodes on the surface of the head. The number of elements in the first dataset (*sleep_4*) is 840 and elements are labeled with five different markers—Wake, REM, NREM1, NREM2 and NREM3. The second dataset contains (*sleep_5*) is 876 and elements are marked as Wake, REM, NREM1 or NREM2.

For all 19 EEG channels the following features were computed: standard deviation, maximum and minimum values, 2nd and 3rd Hjorth parameters, Shannon entropy, power spectral density for typical EEG bands—delta, theta, alpha, beta and gamma. A total of 209 sleep related features were extracted for each 30 s data segments. The feature set is described in more detail in [9, 10]. After feature extraction a Principal Component Analysis (PCA) was used for both EEG datasets to reduce dimensionality to 10, which covered 90% of original data.

### Artificial Dataset

Clusters in the real dataset are not clear and relatively small. Therefore, artificial data were generated. Two datasets of N = 3000 (*random_3000*) and N = 15000 elements (*random_15000*). Each artificial dataset contains 3 clusters that consist of the same amount of data points in 5D space. Cluster elements are generated using three basic points: c1 = (1; 1; 1; 1; 1), c2 = (100; 100; 100; 100; 100) and c3 = (300; 300; 300; 300; 300). In the beginning N basic points of each kind are created and then the points are moved in all dimension by random number from interval $< 100; 0 >$.

## Hybrid Clustering

The presented hybrid algorithm is not purely hierarchical clustering, but the output is the desired dendrogram that denotes hierarchy in the dataset. The idea comes from the work of Olga Tanaseichuk et al. described in [11]. However our algorithm and its implementation is different from the Tanaseichuk's one. The final hybrid algorithm is schematically described in Fig. 1.

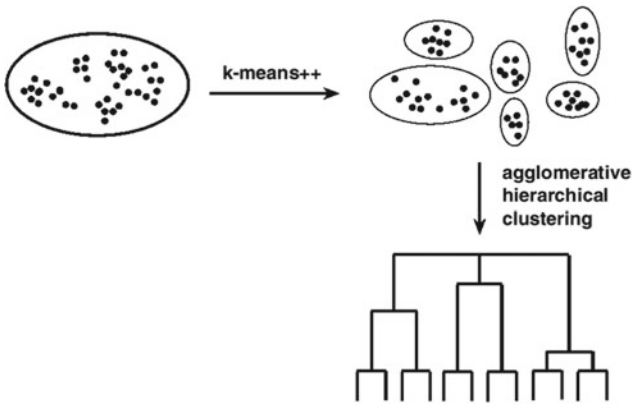



**Fig. 1** Work flow of the hybrid cluster algorithm

In the first step the *k*-means ++ cluster analysis is done in order to split dataset into *k*-clusters. In the second step, the clusters from the first step are hierarchically clustered in a way that they represent leaf clusters. For the second step we used the Ward's hierarchical clustering method used with Euclidean distance as the metric to calculate distances between objects [12].

In basic *k*-means algorithm the initialization of the centers is done uniformly and randomly. In this work we use a different approach of initialization, known as *k*-means ++ [13]. The biggest drawback of this method is its sensitivity to outliers but still, it is more reliable than running initialization many times. If we denote distance between data point $x$ and closest center $c$ as $D(x)$ the *k*-means ++ algorithm can be formally written as shown by Algorithm 2.

**Algorithm 1** k-means++ initialization algorithm

1: Choose uniformly at random first center $c_1$
2: **for** i=1 to $k$ **do**
3: Choose a new center $c_i$ with probability $\frac{D(x)^2}{\sum_{x \in \mathcal{X}} D(x)^2}$
4: **end for**

## Time and Memory Requirements

Before we compare algorithms on the experimental level we need to discuss their theoretical differences. For this purpose we use big $\mathcal{O}$ notation which represents the upper bound of complexity, i.e. the worst imaginable case. Basic Agglomerative Hierarchical clustering on $n$ data points needs in the beginning to store $\frac{1}{2} \cdot N^2$ of similarities which using $\mathcal{O}$ notation means $\mathcal{O}(N^2)$. Since in subsequent steps the similarity matrix becomes smaller and smaller, the maximum memory is required at the onset. Since we are restricted to

large datasets (e.g. $N = 15000$), we need to diminish this requirement.

Another characteristic of the algorithm in a need to track its time consumption. Although the most important parameter for us is the utilized memory, time consumption should not be excessive. In the case of Agglomerative Hierarchical clustering the time complexity is $\mathcal{O}(N^3)$ where $n$ is number of data points [14].

## Clustering Methods Comparison

Hybrid clustering method was compared with agglomerative hierarchical method. Each algorithm was run 5 times with respective setting to obtain correct results. Demanded memory and time consumption were recorded 5 times and averaged. The last characteristic of algorithm used for comparison is the mutual information (MI) [15]. MI is used to compare the ability of algorithms to split dataset into clusters. This property is used only for overall comparison of methods.

Since the real memory and time performance of the algorithm is hardware-dependent, only one computer (4-cores CPU of 2.4 GHz, 16 GB RAM) was used for all the tests.

## Results

The proposed hybrid clustering method and Ward's hierarchical method were tested on four datasets described in the Sect. "Data" (*random_3000*, *random_15000*, *sleep_4* and *sleep_5*). Memory and time consumption results for hybrid algorithm are shown in Figs. 2 and 3. The parameter $k$, that was changed, is from $k$-means ++ algorithm. We take $k$ from the range $<10; 490>$ and $<10; 100>$ for the artificial datasets and for EEG datasets, respectively.

As can be seen on Fig. 2, the memory consumption increases linearly with $k$ for the artificial datasets, while this

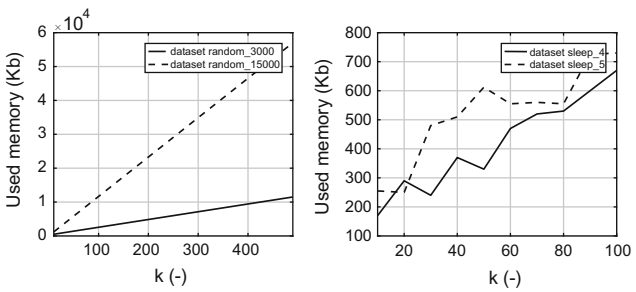



**Fig. 2** Memory consumption of hybrid algorithm for artificial datasets (left) and sleep EEG datasets (right). The value $k$-corresponds to the number of clusters for $k$-means ++ algorithm

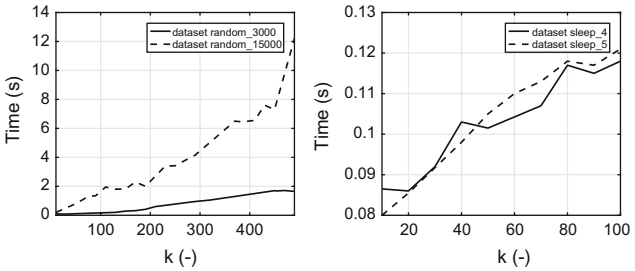



**Fig. 3** Time consumption of hybrid algorithm for artificial datasets (left) and sleep EEG datasets (right). The value $k$-corresponds to the number of clusters for $k$-means ++ algorithm

is not true for the EEG datasets. In addition, the memory consumption is not much of a difference for large values of $k$ for EEG datasets. This is probably due to the smaller number of used features and their regular distribution in the case of artificial datasets. A similar result can also be observed in Fig. 3.

It is non-trivial to tune the parameters $k$ in order to maximize the performance since the number of clusters is unknown. We expect 4–5 most important clusters for EEG datasets. These data may also contain a lot of unknown clusters consisting of various artifacts. We need to set $k$ to higher value to obtain correct result from linkage function. For the final comparison we decided to use $k = 30$. This was done on the basis of mutual information analysis (see below) over EEG datasets. The final mutual information significantly increased up to $k = 20$. For $k > 30$ there was only very gradual increase of mutual information.

Since the main task is to find an algorithm with lower memory usage, reasonable time consumption and approximately same mutual information measures as classical approach (agglomerative hierarchical clustering), we take used memory by algorithm as a main characteristic in the comparison of the methods. These results are presented in Table 1. The memory consumption is lower for hybrid clustering for each dataset. Also time consumption was tested. The results are shown in Table 2. Both algorithms need usually only from fractions of seconds to two seconds to create dendrogram and find clusters for each dataset.

The last parameter, that needs to be described is mutual information between the original labels of clusters and the results of cluster analysis. Table 3 consists an overview on mutual information measures of both algorithms. Higher mutual information means more accurate cluster analysis.

## Discussion

Proposed algorithm is based on the combination of k-means ++ and agglomerative hierarchical clustering. The most important information obtained from the experiments

**Table 1** Memory consumption comparison in Kb

| Dataset | Ward's method | Hybrid clustering |
|---|---|---|
| random_3000 | 35216 | 784 |
| random_15000 | 880568 | 3524 |
| sleep_4 | 3068 | 477 |
| sleep_5 | 2792 | 286 |

**Table 2** Time consumption comparison in seconds

| Dataset | Ward's method | Hybrid clustering |
|---|---|---|
| random_3000 | 0.15 | 0.13 |
| random_15000 | 2.05 | 0.49 |
| sleep_4 | 0.05 | 0.09 |
| sleep_5 | 0.05 | 0.09 |

**Table 3** Mutual information measurement comparison

| Dataset | Ward's method | Hybrid clustering |
|---|---|---|
| random_3000 | 1.59 | 1.58 |
| random_15000 | 1.59 | 1.59 |
| sleep_4 | 0.21 | 0.14 |
| sleep_5 | 0.48 | 0.43 |

with this method is that $k$ should not be too high, because with increasing $k$ we have linearly increasing memory consumption. The largest difference was in the memory consumption for the largest used dataset (*random_15000*). The amount of memory used by agglomerative hierarchical cluster analysis was 250 times more than hybrid clustering approach. By taking a look at time complexity of each algorithm we can see that MATLAB implementation of classical approach is the fastest method. On the other hand time used for both proposed methods is not too high and we can consider it reasonable (see Table 2).

In the initialization step of hybrid method we use $k$-means ++ initialization. Since in the initialization we evaluate distances between points and centers, and comparing distances between each other as in the clustering part of algorithm, the memory consumption is not higher than in the one with classic k-means. The same stands for time complexity, which means that time and memory complexity in the terms of big $\mathcal{O}$ notation does not change with $k$-means ++.

In the second phase of the Hybrid algorithm we use Agglomerative Hierarchical clustering in order to get hierarchy of subclusters. Since here the subclusters are represented with $k$ centroids, the time complexity of this step is $\mathcal{O}(k^3)$ and memory complexity is $\mathcal{O}(k^2)$. In the terms of the rules of big\$mathscrO$ notation and with assumption that $k \ll N$ we can state that time and memory requirements of this algorithm are $\mathcal{O}(k \cdot N \cdot d \cdot i)$, $\mathcal{O}((k+N) \cdot d)$ respectively.

## Conclusion

Hybrid clustering method was tested in the paper and compared with the agglomerative one on the artificial dataset and real biomedical data. We focused on finding an algorithm that would be more memory efficient as compared to the classical hierarchical clustering approach in the reasonable time and the cluster analysis itself would be similar. It is possible to claim that we succeed in finding a method which has lower memory consumption than agglomerative hierarchical clustering. On the other hand we pay for it by higher time consumption and slightly worse cluster detection.

Based on the results of the presented work it is possible to conclude that the lower memory consumption is balanced out by worsening other performance markers. It will be necessary to perform a much more detailed analysis in the future in order to conclude whether presented algorithm indeed performs worse than respective standard approaches. For example, for more accurate comparison of clustering algorithms it will be necessary to compare not only the mutual information analysis results, but also the final classification accuracy.

**Acknowledgements** This research has been supported by the project *Temporal context in analysis of long-term non-stationary multidimensional signal*, register number 17-20480S of the Grant Agency of the Czech Republic.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Lloyd S (2006) Least squares quantization in PCM. IEEE Trans Inf Theory 28:129–137
2. Yin Ch, Zhang S (2017) Parallel implementing improved k-means applied for image retrieval and anomaly detection. Multimed Tools Appl 76:16911–16927
3. Borgwardt S, Brieden A, Gritzmann P (2017) An LP-based k-means algorithm for balancing weighted point sets. Eur J Oper Res 263:349–355
4. Jeon Y, Yoo J, Lee J, Yoon S (2017) NC-link: a new linkage method for efficient hierarchical clustering of large-scale data. IEEE Access 5:5594–5608
5. Medvedev V, Kurasova O, Bernataviciene J, Treigys P, Marcinkevicius V, Dzemyda G (2017) A new web-based solution for modelling data mining processes. Simul Model Pract Theory 76:34–46. High-Performance Modelling and Simulation for Big Data Applications
6. Li L, Xiwei Ch, Dashi L, Yonggang L, Guandong X, Ming LHSC (2013) A spectral clustering algorithm combined with hierarchical method. Neural Netw World 6:499–521
7. Gagolewski M, Bartoszuk M, Cena A (2016) Genie: a new, fast, and outlier-resistant hierarchical clustering algorithm. Inf Sci 363:8–23
8. Iber C (2007) Sleep medicine american academy. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. American Academy of Sleep Medicine
9. Gerla V, Djordjevic V, Lhotska L, Krajca V (2009) System approach to complex signal processing task. Comput Aided Syst Theory-EUROCAST 2009:579–586
10. Gerla V (2012) Automated Analysis of Long-Term EEG Signals. PhD thesis. Czech Technical University in Prague
11. Tanaseichuk O, Hadj Khodabakshi A, Petrov D et al (2015) An efficient hierarchical clustering algorithm for large datasets. Austin J Proteomics Bioinf Genomics 2
12. Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J Classif 31:274–295
13. Arthur D, Vassilvitskii S (2007) K-means ++: the advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms, pp 1027–1035
14. Tan PN, Steinbach M, Kumar V (2006) Introduction to data mining. Pearson International EditionPearson Addison Wesley
15. James RG, Mahoney JR, Crutchfield JP (2017) Information trimming: sufficient statistics, mutual information, and predictability from effective channel states. Phys Rev E 95:060102

# Epileptic Seizure Prediction with Stacked Auto-encoders: Lessons from the Evaluation on a Large and Collaborative Database

R. Barata, B. Ribeiro, A. Dourado, and C. A. Teixeira

**Abstract**

The seizure prediction performance of algorithms based in stacked auto-encoders deep-learning technique has been evaluated. The study is established on long-term electroencephalography (EEG) recordings of 103 patients suffering from drug-resistant epilepsy. The proposed patient-specific methodology consists of feature extraction, classification by machine learning techniques, post-classification alarm generation, and performance evaluation using long-term recordings in a quasi-prospective way. Multiple quantitative features were extracted from EEG recordings. The classifiers were trained to discriminate preictal and non-preictal states. The first part of the feature time series was considered for training, a second part for selection of the "optimal" predictors of each patient, while the remaining data was used for prospective out-of-sample validation. The performance was assessed based on sensitivity and false prediction rate per hour (FPR/h). The prediction performance was statistically evaluated using an analytical random predictor. The validation data consisted of approximately 1664 h of interictal data and 151 seizures, for the invasive patients, and approximately 4446 h of interictal data and 406 seizures for the scalp patients. For the patients with intracranial electrodes 18% of the seizures were correctly predicted (27), leading to an average sensitivity of 16.05% and average FPR/h of 0.27/h. For the patients with scalp electrodes 20.69% of the seizures (84) on the validation set were correctly predicted, leading to an average sensitivity of 17.49% and an average FPR/h of 0.88/h. The observed performances were considered statistically significant for 4/19 invasive patients ($\approx 21\%$) and for 5/84 scalp patients ($\approx 6\%$). The observed results evidence the fact that, when applied in realistic conditions, the auto-encoder based classifier shows limited performance for a larger number of patients. However, the results obtained for some patients point that, in some specific situations seizure prediction is possible, providing a "proof-of-principle" of the feasibility of a prospective alarming system.

**Keywords**

EPILEPSIAE database • Seizure prediction • Machine learning • Stacked auto-encoders • Deep learning

## Introduction

Despite available drug and surgical treatment options, more than 30% of patients with epilepsy continue to experience seizures [1]. In these patients with pharmaco-resistant epilepsy, the apparent unpredictability of seizure occurrence

R. Barata · B. Ribeiro · A. Dourado · C. A. Teixeira (✉)
Centro de Informática e Sistemas (CISUC), Departamento de Engenharia Informática, Universidade de Coimbra, Coimbra, Portugal
e-mail: cteixei@dei.uc.pt



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_2

imposes a considerable limitation of their daily lives and results in a high risk of unforeseen endangering situations [2, 3]. A system that could warn the patient of an impending seizure or trigger an antiepileptic device to prevent seizure occurrence would dramatically improve the quality of life for especially these patients.

Although some promising results have been reported, suggesting the existence of a pre-seizure state and an consequently the capability to predict seizures [4], until recently prediction performances have not been prospectively evaluated on large, collaborative databases [3]. Thus, the question to which extent seizure prediction is possible remains unanswered [5–7]. In particular, the weaknesses of most studies are (1) that they optimize parameters retrospectively, (2) that they rely on short, selected data sets, and (3) that they lack rigorous statistical evaluation [4]. In [8, 9] report for the first time studies that make use of an appropriate database developed on behalf of the EPILEPSIAE project [10, 11]. The EPILEPSIAE database is used in our study.

The aim of this study is to investigate for the first time whether a deep-learning technique based on Stacked Auto-encoders (SA), and having as input EEG-based features are able to predict seizures quasi-prospectively on a large multiple-center database of long-term unselected recordings. Quasi-prospectively here refers to the fact that the data have actually been recorded, which renders a prospective study impossible. But as we only evaluate our results on unevaluated validation data, we are as close to a prospective study as possible.

It was referred by several authors that, to date, no single feature has shown predictive power [4, 12, 13]. Machine-learning methods can face at the same time several features computed from raw data, and collected from different cerebral sites. These methods try to classify the brain state based not only on a single feature, but also on the general behavior of the set of features, exploring their linear and non-linear interactions. This paper as others already published, such as [13], hypothesize that a patient-specific approach based on multiple EEG measures will achieve high sensitivity and specificity.

## Data & Methods

In this section, the database used as well as the methodology employed are described. In this paper, we distinguish "classifier" from "predictor". Classifier is the first part of the prediction system that discriminates the feature samples in four brain states (classes). A predictor is the full system composed by the classifiers plus the alarm generation procedure.

### Patient Characteristics and EEG Database

Long-term EEG recordings from 103 epilepsy patients (50 males; age range, 10–65 years; mean age: 35 years) suffering from medically intractable partial epilepsy were analyzed in this study. Data had been recorded in three different epilepsy units (Hospitais da Universidade de Coimbra, Portugal; Unité d'Épilepsie of the Pitié-Salpêtrière Hospital, Paris, France; Epilepsy Center, University Medical Centre of Freiburg, Germany) resulting in a total of almost 707 days (16,963 h) of EEG including 1062 seizures, and is part of the 275 patients containing in the EPILEPSIAE database [10, 11]. These 103 patients are those that have more than 5 seizures and enable the evaluation strategy implemented in this paper.

43% of the patients had temporal lobe epilepsy, lateralized to the right in 57%, to the left in 30% and bilateral in 13%. In 84 patients, EEG was recorded using 22–37 scalp electrodes; the average recording period was 149 h. In 19 patients, intracranial EEG with 14–121 recording sites was recorded using stereotactically implanted depth electrodes, subdural grids and/or strips; the average recording period was 232 h. EEG data were recorded using a Nicolet, Micromed, Compumedics, or Neurofile NT digital video EEG system at sampling rates of 256, 400, 512, 1024 or 2500 Hz.

### Feature Extraction

For the recorded EEG channels, 22 univariate features were extracted every five seconds of EEG with no overlap, i.e., consecutive five-seconds windows were considered for all the analyzed patients. The computed features are listed in Table 1. More details can be obtained in [14]. Features have to be computationally efficient, i.e., with potential for online implementation in low computational power environments. This is why we restricted ourselves to univariate linear features in this study.

### The Approach to a Quasi-prospective Study

The same feature was computed for all the electrodes and for all the patients. The evaluation methodology implemented in this paper encompasses the splitting of these time series into three parts. A first part containing the first three seizures, was used for optimizing the classifiers, i.e., for training. A second part containing the next two seizures was used for the selection of an "optimal" predictor among all the tested alternatives. The third and last part, containing at least one

**Table 1** EEG features

| | |
|---|---|
| Time domain | Mean |
| | Variance |
| | Skewness |
| | Kurtosis |
| | Energy |
| Frequency domain | MSE of estimated AR models |
| | Delta band rel. power (0.1–4 Hz) |
| | Theta band rel. power (4–8 Hz) |
| | Alpha band rel. power (8–15 Hz) |
| | Beta band rel. power (15–30 Hz) |
| | Gamma band rel. power (> 30 Hz) |
| | Spectral edge frequency (90%) |
| | Spectral edge power (90%) |
| | Decorrelation time |
| | Hjorth mobility |
| | Hjorth complexity |
| Time frequency | Energy of DB4 wavelet coefficients (6 decomposition levels) |

seizure, was used for quasi-prospective long-term out-of-sample evaluation, i.e., for validation. The average interictal duration of the validation data was approximately 88 h, and 54 h for the invasive, and scalp recordings, respectively. In total, 6111 h of interictal out-of-sample data were considered containing a total of 557 seizures.

## Classification Methodology

In this paper, the brain state is classified over time into one of two states: preictal and non-preictal; and the number of classifier's inputs are 22 times the number of channels, originating a very high dimensional space. The windows that are just before the seizure onset times are nominated as preictal windows. The non-preictal class encompass the ictal, postictal and interictal periods. The ictal period is the time frame where seizure occurs. The postictal period refers to epochs that are just after the seizure offset time. The interictal period relates to seizure-free epochs. The neurologists clinically defined the seizure onset and offset times. The exact time where the preictal state starts is unknown and can probably be patient-dependent. Thus, in a first approach, the search for appropriate predictors should include the consideration of a range of preictal times, or seizure occurrence periods (SOPs). In this work, four SOPs were considered: 10, 20, 30 and 40 min.

In this paper, we use machine-learning techniques to define decision boundaries between classes. The machine learning technique used was SA [15]. SA can be considered a deep machine learning algorithm, since there are based on the construction of a new representation of the data for posterior classification, a technique known as representation leaning. The neural networks developed contained between three and six layers, and the number of neuron varies between 1000 and 10 per layer, diminishing from the input to the output of the network. Their auto-encoders were trained according to a two-phase protocol: Greedy Layer-Wise Training procedure and that were then fine-tuned [16]. The first stage of the training is used to optimize each auto-encoder, individually, to compress and restructure the data in the best way possible, by introducing a bottleneck on the network [17]. The second stage of fine-tuning, is meant to lead the network to be more discriminative regarding the classes [16].

By making use of such a sophisticated technic we hope to overcome the obstacles faced when using shallow or classic classifiers. Moreover, due to the data dimensionality a system capable of automatically reduce the number of dimensions while likely retaining only the important information justify the use of SA.

## Seizure Prediction Method

If we consider that a single positive classifier output represents a prediction, we can have a prediction at each 5 s. As in practice a classifier will most likely not classify all of the samples correctly, a lot of false alarms could be issued. To reduce the number of false alarms the output of the classifier was smoothed using a method described in [18, 19]. This method is based on a sliding window with a size equal to the

considered preictal time, or SOP. Alarms are raised whenever the number of epochs classified as preictal in the sliding window is larger than a predefined threshold. The thresholds considered for each patient were 0.2, 0.4, 0.6 and 0.8, meaning that 20%, 40%, 60%, and 80% of the samples were classified as preictal inside the window, respectively. Once an alarm is raised, another one can only occur after a dead-time equal to the SOP. The main advantage of the algorithm implemented is that it is exactly known during which time period a seizure is to be expected. In the case of approaches that allow re-triggering, such as the one presented in [20], the alarms may prolong for long times and the patients do not know definitely when the seizure is to be expected.

## Quasi-prospective Performance Evaluation

The performance of the seizure prediction algorithm is evaluated using the seizure prediction characteristics, which characterizes the sensitivity of the seizure prediction algorithm, given the maximum rate of false alarms and two-time windows: the SOP and the time needed to perform any intervention, called the intervention time (IT). A correctly predicted seizure requires that the seizure onset to occur in the time window occurrence period. In this paper, SOP assumes one of the values defined for the preictal time, while the IT was fixed as 10 s. As specificity, we used the false prediction rate defined as the number of false alarms divided by the duration during which false alarms could be triggered, which is obtained by subtracting the time under false warning from the total interictal duration. Mathematically the FPR is given by [4]:

$$FPR = \frac{\#False\ Alarms}{Interictal\ Duration - (\#False\ Alarms \times SOP)}. \quad (1)$$

To statistically evaluate the results we use the analytic random predictor based on the binomial distribution [3, 21]. It quantifies critical sensitivities that could be obtained by chance given the time windows and the false prediction rate. No other information of the data is provided to the random predictor. Because we implemented a quasi-prospective evaluation protocol that resulted in the selection of a best predictor, we only tested a single predictor, i.e., a single degree of freedom is attained, leaving out the need for multiple testing corrections.

## Results and Discussion

For each patient four different preictal times were tested, as well as four different thresholds used on the alarm generation system, making 16 different predictors per patient. The predictor applied on the validation data was selected based on the performance on the testing set, and was the one with performance closest to 100% sensitivity, and 0/h FPR. This selection was based on a period containing two seizures that were not used for the training of the classifiers nor for any performance evaluation.

For the patients with intracranial electrodes the validation data consisted of 1664.42 h of recording and 151 seizures. For the patients with scalp electrodes the validation data consisted in 4445.57 h and 406 seizures.

We found out that for the patients with intracranial electrodes 18% of the seizures were correctly predicted (27), leading to an average sensitivity of 16.05% and average false positive rate of 0.27/h. For the patients with scalp electrodes 20.69% of the seizures (84) on the validation set were correctly predicted, leading to an average sensitivity of 17.49% and an average false prediction rate of 0.88/h.

We have compared these performances to those achieved by a random predictor [21]. We found that preictal changes can be identified above chance level in four out of 19 intracranially monitored patients ($\approx$ 21%) and in five out of 84 scalp monitored patients ($\approx$ 6%).

Regarding the influence of certain variables in the results, several important conclusions can be taken. For the patients with intracranial electrodes, the sensitivity of the female subjects was less than half of the male patients, 9.70% and 24.71%, respectively. Regarding the focal character of the seizures, the sensibility is considerably higher on patients with well-defined focalization. Moreover, three of the four patients with results above chance level have well defined focalization character, and all of these have focus on the frontal right lobe. Regarding the patients with scalp electrodes, the general panorama is that sleep stage influences sensitivity (paired t-test p-value $< 0.05$). For the scalp population, the preictal period also significantly influences sensitivity (paired t-test p-value = 0.01). Of the different preictal period durations tested, 30 min was the value that leaded to better results.

Comparing our results with the ones published in [9] that employed a realistic validation schema similar to the one implemented in this paper, similar results were obtained. The advantage of our approach relies on the automatic feature and channel selection accomplished by the deep-learning technique used.

## Conclusions

We approached seizure prediction to the best of our knowledge for the first time based on SA, applied in multicentre, long-term, unselected data. The result argues for the possibility to predict seizures, at least for some patients. Future work will be devoted to a true online assessment and evaluation of seizure prediction performance.

**Acknowledgements** The first author acknowledge the financial support received from CISUC.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Kwan P, Brodie MJ (2000) Early identification of refractory epilepsy. N Engl J Med 342:314–319
2. Cockerell OC, Johnson AL, Sander JW et al (1994) Mortality from epilepsy: results from a prospective population-based study. Lancet 344:918–921. https://doi.org/10.1016/S0140-6736(94)92270-5
3. Schulze-Bonhage A, Sales F, Wagner K et al (2010) Views of patients with epilepsy on seizure prediction devices. Epilepsy Behav 18:388–396. https://doi.org/10.1016/j.yebeh.2010.05.008
4. Mormann F, Andrzejak RG, Elger CE, Lehnertz K (2007) Seizure prediction: the long and winding road. Brain 130:314–333. https://doi.org/10.1093/brain/awl241
5. Lehnertz K, Litt B (2005) The first international collaborative workshop on seizure prediction: summary and data description. Clin Neurophysiol 116:493–505
6. Lehnertz K, Le Van Quyen M, Litt B (2007) Seizure prediction. In: Engel J, Pedley TA, Aicardi J (eds) Epilepsy A Compr. Lippincott Williams & Wilkins, Textb, pp 1011–1024
7. Stacey W, Le Van Quyen M, Mormann F, Schulze-Bonhage A (2011) What is the present-day EEG evidence for a preictal state? Epilepsy Res 97:243–251. https://doi.org/10.1016/j.eplepsyres.2011.07.012
8. Teixeira CA, Direito B, Alexandre Teixeira C et al (2014) Epileptic seizure predictors based on computational intelligence techniques: a comparative study with 278 patients. Comput Methods Programs Biomed 114:324–336. https://doi.org/10.1016/j.cmpb.2014.02.007
9. Direito B, Teixeira CA, Sales F et al (2017) A realistic seizure prediction study based on multiclass SVM. Int J Neural Syst 27:1750006. https://doi.org/10.1142/S012906571750006X
10. Klatt J, Feldwitch-Drentrup H, Ihle M et al (2012) The EPILEPSIAE database—an extensive electroencephalography database of epilepsy patients. Epilepsia 53:1669–1676. https://doi.org/10.1111/j.1528-1167.2012.03564.x
11. Ihle M, Feldwisch-Drentrup H, Teixeira CA et al (2012) EPILEPSIAE—a European epilepsy database. Comput Methods Programs Biomed 106:127–138
12. Feldwisch-Drentrup H, Schelter B, Jachan M et al (2010) Joining the benefits: combining epileptic seizure prediction methods. Epilepsia 51:1598–1606
13. Park Y, Netoff T, Parhi K (2010) Seizure prediction with spectral power of eeg using cost-sensitive support vector machines. J Med Device 4:27542
14. Teixeira CA, Direito B, Feldwisch-Drentrup H et al (2011) EPILAB: a software package for studies on the prediction of epileptic seizures. J Neurosci Methods 200:257–271. https://doi.org/10.1016/j.jneumeth.2011.07.002
15. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science (80-) 313:504–507
16. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press
17. Erhan D, Bengio Y, Courville A et al (2010) Why does unsupervised pre-training help deep learning? J Mach Learn Res 11:625–660
18. Teixeira CA, Direito B, Feldwisch-Drentrup H et al (2011) EPILAB: a software package for studies on the prediction of epileptic seizures. J Neurosci Methods 200:257–271. https://doi.org/10.1016/j.jneumeth.2011.07.002
19. Teixeira C, Direito B, Bandarabadi M, Dourado A (2012) Output regularization of svm seizure predictors: Kalman filter versus the "firing power" method. In: 2012 annual international conference on IEEE engineering medicine biology society (EMBC), pp 6530–6533
20. Snyder DE, Echauz J, Grimes DB, Litt B (2008) The statistics of a practical seizure warning system. J Neural Eng 5:392–401. https://doi.org/10.1088/1741-2560/5/4/004
21. Schelter B, Winterhalder M, Maiwald T et al (2006) Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction. Chaos 16:13108. https://doi.org/10.1063/1.2137623

# Deep Learning Techniques on Sparsely Sampled Multichannel Data—Identify Deterioration in ICU Patients

A. Chytas, K. Vaporidi, Y. Surlatzis, D. Georgopoulos, N. Maglaveras, and I. Chouvarda

**Abstract**

The focus of this paper is to recognize periods of time deviating from the norm using sparsely sampled multichannel signals. The case in question being the ICU, our domain of interest is patient deterioration. In many cases the recording and analyzing of frequently sampled streaming data that can carry more information is not always an option, while at the same time the availability for data recorded at large time intervals is a common occurrence. To address this issue, we examine whether Deep-learning methods can provide efficient results regarding the recognition of different states during the hospitalization, by utilizing hourly multichannel physiological recordings.

**Keywords**

ICU • Deterioration • Deep-learning • Autoencoders

## Introduction

Many Intensive Care Units (ICUs) lack the necessary software and hardware to continuously record streaming data, and thus they rely on manually recording vitals and other displayed values to the ICU database. And in the cases that it is possible, the analyzing of this continuous stream of can be a resource heavy task. Our approach is to utilize data that is most common to be found in clinical routine and develop methods to extract information out of them regarding the patient progress over various time-frames, using Deep learning techniques. Such techniques traditionally used in image processing can become applicable to our multichannel bio-signals, given the appropriate pre-processing.

Deep learning techniques have been successfully applied on ICU data in order to examine patient progress [1, 2]. Specific techniques that involve image inferring from time-series have also been used in medical data [3], Their approach involved the use of recurrence plots [4], for our dataset limitations we opted for a more straightforward representation of multichannel signals as images.

The paper structure is as follows; section "Data" is dedicated to the data and their acquisition, section "Methodology" describes the methods used for this research and section "Results" details the results. Sections "Compliance with Ethical Requirements" and "Conclusions" are referring to the ethical compliance and the paper's conclusions respectively.

## Data

Physiological data recordings for 2239 patients were gathered in PAGNI ICU, following all ethical procedures. The data were recorded manually at 1 h intervals, or in some cases where it was deemed necessary, more often. The

A. Chytas (✉) · N. Maglaveras · I. Chouvarda
Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: achillec@auth.gr

A. Chytas · N. Maglaveras · I. Chouvarda
Institute of Applied Biosciences (INAB), Centre for Research and Technology Hellas (CERTH), 6th km Charilaou-Thermi, GR 57001, Thessaloniki, Greece

K. Vaporidi · Y. Surlatzis · D. Georgopoulos
Department of Intensive Care Medicine, Medical School, University Hospital of Heraklion, Heraklion, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_3

**Table 1** Biosignals used in the analysis

| Name |
|---|
| Temperature (core) |
| Systolic pressure |
| Diastolic pressure |
| Respiratory rate |
| Heart rate |
| Oxygen saturation |
| Arterial pressure |

median duration of the recordings was 88 h. More than 50 signals with similar sampling rate, including drug flow administration data and ventilation related data (recordings, settings, etc.), were recorded, but their presence was not consistent, some patients were administrated insulin, patients in general were given different types of medicine. To bypass discrepancies in missing signals and to increase our data pool, we choose to utilize the bio-signals that were present to the majority of the patients, as depicted in Table 1.

## Methodology

### Preprocessing

Consecutive time periods that had none of the values recorded (e.g. when patient was admitted for surgery) were removed. Also, values that were outside of the possible physiological boundaries were considered invalid and were handled as if they were missing.

Even in those signals there was several missing values. To address the issue the multiple imputation by chained equations (MICE) [5] was applied and more specifically the Approximate Bayesian bootstrap [6, 7] since our missing values pattern was random.

The dataset was split into 8 h segments, and after the data cleaning and quality control we were left with approximately 70000 such segments.

### Autoencoders

Autoencoders are neural networks that try to reconstruct an original object to an approximate output [8, 9] often using backpropagation methods [10] They are trained to create a function that approximates the identity function so that the output is as similar as possible to the input. This can be achieved by using a different combination and number of hidden layers and the sizes of said layers. If the hidden layers unit size is set to a smaller amount compared to the input. the neural network is constrained into learning a compressed representation of the initial object [11].

For the main part of our analysis we will be using Autoencoders to extract features from our dataset that can differentiate between the changes in a patient's status in different time periods. Neural networks can not only be highly accurate on classification [12] but also, as previous studies have shown [13, 14], that is possible that by training neural networks you can extract high level information using unlabeled data. This unsupervised method can isolate and identify complex invariances. This part of the analysis was implemented with the python Keras framework.

For this approach to be used the data should be formatted much like an image. Yu et al. [3] were successful in inferring images based on signals and using CNNs on the transformed data. But in contrast to their approach, which required the use of recurrence plots [4] we decided to follow a more straightforward approach for representing our multichannel signals as images. Mainly due to our low sampling rate of our data, each 8 h segment was transformed as an $8 \times 7$ matrix which represented an image. Each column, which was a different signal was normalized [0, 1] to each one's normal limits.

The set was split randomly multiple times for cross-validation purposes into two sub-datasets that were used for training (75% of the data) and testing the neural networks respectively. Two layers were used for the finally utilized neural network. This method produced a set of features, smaller than the original, that are used to represent the original dataset (Fig. 1).

This approach would help as examine whether general patient patterns can emerge from our data. As an alternative approach, more focus on individual patient deterioration, the dataset was again split into two groups. In this procedure. which was repeated for each patient, the dataset used for training was data in its totality with the exception of the patient in question and the testing was the time-frames of the previously excluded patient.

For both cases, along with the dimensionality reduction method used in the previous step that provided a small set of features, those said features had afterwards subject to additional dimensionality reduction by using Principal Component Analysis (PCA). The components that were responsible for the 95% of the total variance were selected.

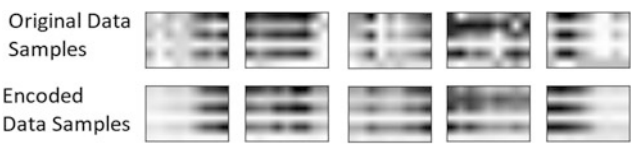



**Fig. 1** Dimensionality reduction applied on the time-frames

## Clustering

For the next and final part of the analysis we looked for the presence of clusters forming in our data-points. For our first scenario, we checked whether there would appear clusters among the dataset as a whole, and for the next scenario between the 8 h time-frames of individual patients. For both clustering cases, the bagged K-means [15] approach was applied. Regarding the whole dataset clustering we applied a hierarchical clustering afterwards.

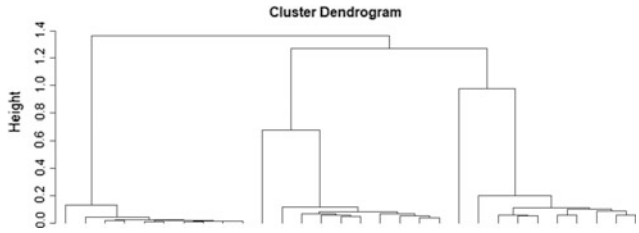



**Fig. 2** Clustering of the 8 h segments used for testing

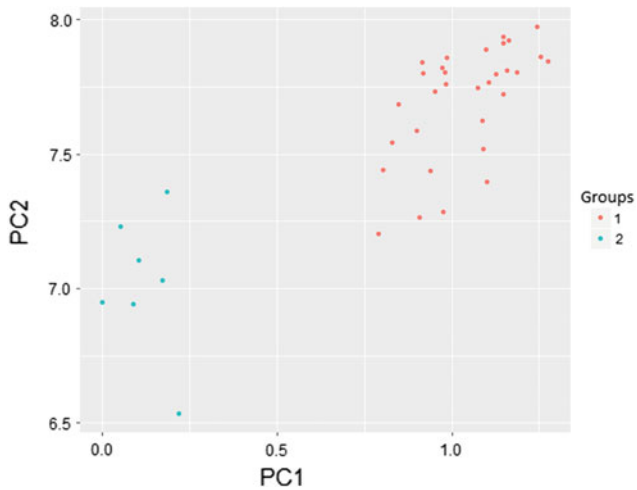



**Fig. 3** Patient-wise clustering, 2 visible clusters

## Results

Our first approach is to look at the dataset as whole, all patients included, and infer general patterns for total population. From the clustering of our dataset in its totality, all patients included, three major groups are distinct. Of course, that alone is not enough to provide meaningful insights about deterioration. The resulted hierarchy (Fig. 2) could imply different patient types which could also be proven useful, in a different way though.

Patient wise, in a small amount of cases two distinct clusters of 8 h segments appeared (Fig. 3). But that was not the case for every patient. In most patients, there could be distinguished a large cluster with a few 8 h segments being scattered around the said cluster (Fig. 4). While a few patients had either no visible groups forming or a single cluster with little to none deviating points. Although the results thus far have not been examined alongside with information regarding medical interventions and patient's events, our hypothesis is that patients spend most of their time in stable (for themselves) state. This stable condition is occasionally disrupted, either by improvement or deterioration. This could be inferred by the amount of data points that are deviating from the big "stable" cluster and sometimes are following opposing trajectories.

## Compliance with Ethical Requirements

Ethical approval for the processing of data has been obtained for the data that have been collected thus far and subsequently used in this research. Those said data have been

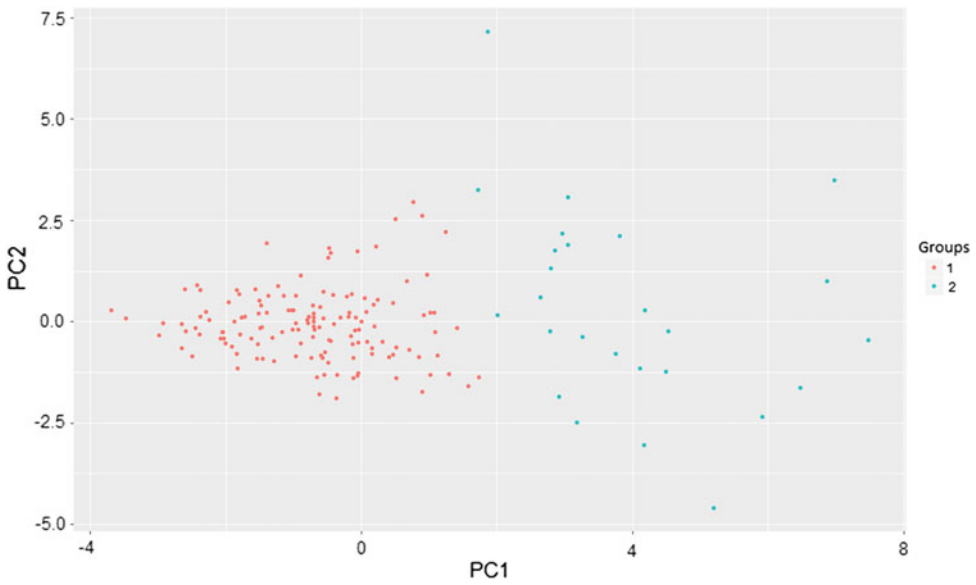



**Fig. 4** Patient-wise clustering, 1 visible cluster and additional scattered data points

fully anonymized, including time information, a procedure that took place in the premise of the PAGNI hospital. The researchers not directly affiliated with PAGNI have signed the necessary disclosure and have been authorized to process the anonymized data.

## Conclusions

The option of having densely sampled data available and ready for analysis is not always a realistic scenario in all cases. Sparsely sampled data in medical environments still exist (both by being actively produced and by have being recorded in earlier stages) and probably will continue to be a reality even in the so-called first world countries, and we can take advantage of their existence using state of the art processing methods. Deep learning techniques can provide additional insights on phenomena by inferring information that might be less apparent to the naked eye. Training neural networks with the data available can help to recognize changes in the patient norm, and facilitate an earlier indication of deterioration.

**Acknowledgements** This work leading this research has been partially supported by the E.C. funded program AEGLE under H2020 Grant Agreement No: 644906.

**Conflict of Interest** The authors declare that there is no conflict of interest.

## References

1. Beaulieu-Jones BK, Orzechowski P, Moore JH (2017) Mapping patient trajectories using longitudinal extraction and deep learning in the MIMIC-III critical care database. https://doi.org/10.1101/177428
2. Kam HJ, Kim HY (2017) Learning representations for the early detection of sepsis with deep neural networks. Comput Biol Med. https://doi.org/10.1016/j.compbiomed.2017.08.015
3. Yu et al (2016) Encoding physiological signals as images for affective state recognition using convolutional neural networks. In: 38th annual international conference of the IEEE engineering in medicine and biology society. Orlando, FL, pp 812–815
4. Eckmann J-P, Kamphorst SO, Ruelle D (1987) Recurrence plots of dynamical systems. Europhys Lett 4(9):973–977. https://doi.org/10.1209/0295-5075/4/9/004
5. Van Buuren S, Boshuizen HC, Knook DL (1999) Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 18(6):681–694
6. Rubin DB (1996) Multiple imputation after 18+ Years. J Am Stat Assoc 91(434):473–489. https://doi.org/10.1080/01621459.1996.10476908
7. Van Buuren S (2007) Multiple imputation of discrete and continuous data by fully conditional specification. Stat Methods Med Res 16(3):219–242. https://doi.org/10.1177/0962280206074463
8. Hinton GE, Zemel RS (1994) Autoencoders, minimum description length and helmholtz free energy. Adv Neural Inf Process Syst 3. https://doi.org/10.1021/jp906511z
9. Bengio Y (2009) Learning Deep Architectures for AI. Found Trends Mach Learn 2(1):1–127. https://doi.org/10.1561/2200000006
10. Liou CY, Cheng WC, Liou JW, Liou DR (2014) Autoencoder for words. Neurocomputing 139:84–96. https://doi.org/10.1016/j.neucom.2013.09.055
11. Vincent P, Larochelle H (2010) Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion pierre-antoine manzagol. J Mach Learn Res 11:3371–3408. https://doi.org/10.1111/1467-8535.00290
12. Krizhevsky A, Sutskever I, Geoffrey EH (2012) ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25(NIPS2012):1–9. https://doi.org/10.1109/5.726791
13. Erhan D, Courville A, Vincent P (2010) Why does unsupervised pre-training help deep learning? J Mach Learn Res 11:625–660. https://doi.org/10.1145/1756006.1756025
14. Le QV et al (2013) Building high-level features using large scale unsupervised learning. In: 2013 IEEE international conference on acoustics, speech and signal processing, pp 8595–8598. https://doi.org/10.1109/MSP.2011.940881
15. Leisch F (1999) Bagged clustering. Adapt Inf Syst Model Econ Manage Sci 51:11. https://doi.org/10.1126/science.1127647

# Convolutional Neural Networks for Early Seizure Alert System

T. Iešmantas and R. Alzbutas

**Abstract**

A general framework of a system for early seizure detection and alert is presented. Many studies have shown high potential of electroencephalograms (EEG) when there are used together with machine learning algorithms for seizure/non-seizure classification task. In this paper, mainly guidelines will be presented on how to use convolutional neural networks for the purpose of highly accurate classification of non-invasive EEG for patients with epilepsy. Convolutional neural networks can be pre-trained on a sample data as described in this paper and then implemented into an application or a device, which readjusts its parameters according to the patient-specific EEG patterns and thus can be further used as a seizure monitoring and alert system. The paper also demonstrated how transfer learning can be applied to create a patient-specific classifier with high accuracy.

**Keywords**

EEG • Epilepsy • Convolutional neural networks • Deep learning • Classification

## Introduction

Epileptic seizures detection based on EEG signal has seen some major developments over the last or this decade. It is mostly due to the application of machine learning algorithms and so called nonlinear measures of neurophysiological signals, like cross-correlation coherence, granger causality (see a nice review on this topic in [1]). Scalp EEGs were classified by support vector machines (SVM) at $\sim$96% of seizure detection rate [2] only by subdividing EEG spectra into 8 intervals. Convolutional neural networks achieved a zero-false-alarm seizure prediction on 20 patients out of 21 when analysed with phase-locking synchronicity, coherence or entropy of phase difference [3, 4]. SVM with spatiotemporal correlation structure features of EEG were also applied to classification of seizures of 19 patients and achieved 86% of seizure detection accuracy [5]. More information on various other classification results can be found in a survey by Nasehi and Pourghassem [6]. Above all, convolutional neural networks showed the best results. However, it was applied to intracranial EEG (i.e. invasive). Scalp EEG have considerably more noise due to various movement artefacts. It is difficult to find a method which would be able to clean all the artefacts, therefore for scalp EEG one should be aware of this variability (noise) as the classifier will be affected by it as well. On the other hand, high accuracy of machine learning methods enables the conception of efficient technologies for personal health monitoring. A device, able to produce an early warning of occurring seizure may decrease the harm of the seizure (which is not only physical injuries but also anxiety and depression). A therapeutic system capable of detecting and reacting to the onset of seizure may administer a local electrical, thermal, or neurochemical stimulus that halts the progression of seizure prior to the development of the symptoms. Or it could at least warn the patient or caring personnel to take some alleviating actions.

For such a system, based on deep learning methods, one needs not only to train it on a large sample of data, but also to

T. Iešmantas (✉) · R. Alzbutas
Faculty of Mathematics and Natural Sciences, Department of Applied Mathematics Kaunas, Kaunas University of Technology, K. Donelaicio g. 73, Kaunas, Lithuania
e-mail: tomas.iesmantas@ktu.lt



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_4

have ability to learn patient-specific seizure patterns. Training on number of patients would produce a classifier able to correctly identify highly varying seizure patterns. This classifier then should be able to readjust to the patient specific EEG data. This could be done by so called transfer learning.

In this paper, we consider a noisy sample of scalp EEG (https://www.physionet.org/pn6/chbmit/) and apply convolutional neural networks for seizure classification. Phase difference synchronicity is chosen as a bivariate feature, describing state of the brain. We investigate several architectures of the network for classification problem. Finally, we discuss the possibility of obtained classifier to apply for unseen patient and EEG pattern.

## Methods

### Convolutional neural network

Convolutional neural network (further, ConvNet) is a type of artificial neural network. ConvNet gained its fame as an image classifier and document recognition algorithm through the work of Yann Lecul [7–9]. The main architecture ingredients are convolutional layers, which are collections of learnable image filters, pooling layers, which are a kind of information aggregators or down-samplers and fully connected layers as in classical neural network (see an illustration of simple architecture in Fig. 1).

ConvNets can learn low-level features and high-level features in an integrated manner. Their main advantage is that they can learn optimal time-invariant local feature detectors and thus build representations that are robust to time shifts of specific feature motifs [3]. Therefore, it is a good classification method for time-dependent EEG or multivariate time series in general.

### Phase locking synchrony

The phases of two coupled nonlinear (noisy or chaotic) oscillators may synchronize even if their amplitudes remain uncorrelated. By synchronization, it is meant here that the following phase locking condition applies for any moment:

$$\varphi_{n,m}(t) = \left|n\varphi_x(t) - m\varphi_y(t)\right| < const.,$$

where $\varphi_x(t)$ and $\varphi_y(t)$ are phases of the signals. The most common scenario for the assessment of phase synchronicity entails the analysis of the pairs of signals. Thus, in typical experimental set-up, $n$ different channels of, say, EEG signals, are recorded and one studies the phase synchronicity between each of the pair of electrodes [1].

Through mathematical analysis of the spatiotemporal dynamics found in EEG recordings of patients with medically intractable epilepsy, researchers have discovered a preictal transition that precedes seizures for periods on the order of minutes to hours. This preictal dynamical transition is characterized by a progressive convergence (entrainment) of dynamical measures at specific anatomical areas in the neocortex and hippocampus [10].

Phase locking value (PLV, [11]) is one of the measures to capture the synchronization of different brain areas. The computation of PLV requires the Hilbert transform, arctangent, addition, sine and cosine, moving-average filtering and lastly, the PLV magnitude. Phase locking value is then defined as follows:

$$PLV = \left|\frac{1}{N}\sum_{n=1}^{N} e^{i\phi_n}\right|, \quad (1)$$

where $\phi_n$ is the relative phase at an nth time point; $N$ is length of a time frame over which averaging is performed.

Filtering is performed with finite impulse response filters with ranges of alpha, beta, gamma and delta brain waves. Then PLV is calculated for each type of waves.

### Pattern construction

Overall construction of patterns with bivariate features from EEG is described in detail in [3]. Here, it will be briefly presented.

Bivariate feature is a measure of relationship between two signals. First, we take a 5 s window and calculate PLV on six different frequency bands for every pair of channels: delta (< 4 Hz), theta (4–7 Hz), alpha (7–13 Hz), low beta (13–15 Hz), high beta (15–30 Hz), low gamma (30–45 Hz). This is done for each pair of channels. Then a spatio-temporal pattern is formed by stacking next to each other 12 consecutive 5 s PLV values. An example of such a pattern is in Fig. 2, where rows of the frame represents PLV values for each pair of channels and each frequency bands, while different columns are for PLV values obtained over different 5 s time windows. To increase the number of samples, every pattern was

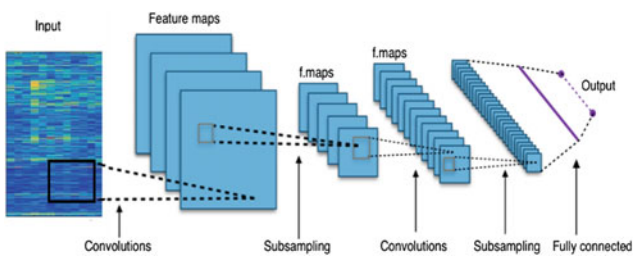



**Fig. 1** Architecture of typical ConvNet

constructed by sliding the frame every 5 s (hence, consecutive patterns has high overlapping).

A collection of such patterns is now a sample of images with labels "seizure" or "no seizure". Therefore, we have a binary classification problem. If the number of channels is N, then number of dimensions of one pattern is N x (N−1)/2 × 6 × 12. In data that we analysed, there were 21 channels (there was 23 channels, but 2 was dropped, because it was simply an inverted signal of other two channels), thus one pattern has dimension 21 * 20/2 * 6 * 12 = 15120.

This EEG data transformation into an image enables application of deep learning algorithms that are traditionally used for image classification. In addition, the above pattern construction is not the only way, and it would be interesting to see how (if at all) the classification results would differ with differently constructed images/patterns.

In addition, the division of EEG frequency into 6 bands is arbitrary and different segmentations might be considered in the future.

The data sample consists of 916 h of continuous scalp EEG signal sampled at 256 Hz. The data set was recorded from 23 pediatric patients at Children's Hospital Boston and one adult patient at Beth Israel Deaconess Medical Center. It was judged that patients experienced 173 events that were identified as clinical seizures.

## Classification Results

For classification problem, we assumed two classes: "1" for seizure and "−1" for non-seizure class. A pattern (constructed as in part C) does not necessarily cover an entire seizure, thus a seizure can be represented by several patterns if it is a long lasting event; the label is the same, no matter which phase of the seizure is covered by the pattern. The cost function that we used was squared error loss:

$$L(y, \tilde{y}) = \sum_{j=1}^{K} \left(y_j - \tilde{y}_j\right)^2; \tag{2}$$

where $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_K)^T$ is the output of the last fully connected *softmax* layer and $y$ is the true labels.

One of the problems that seizure classification has it is high imbalance of classes. In our case, we have only one seizure label for 100 non-seizure labels. This high imbalance of classes, if not treated carefully, would result to high recognition of non-seizure periods, while most seizures would be labelled incorrectly. As a way out, oversampling of "seizure" class patterns was used, i.e. equalization of class ratios was achieved by oversampling the "seizure class" (down-sampling of "non-seizure class" was not used). It proved a sufficient strategy to avoid high imbalance problems.

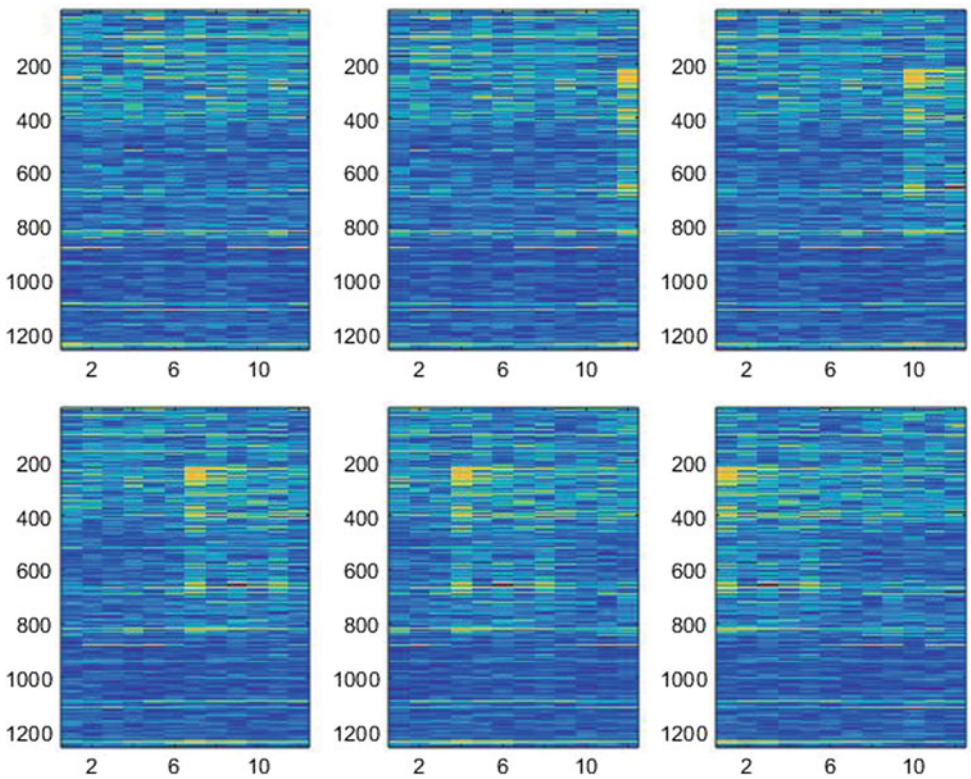



**Fig. 2** Several examples of seizure patterns. One can clearly identify a seizure occurring and developing in time (from top left to bottom right is the direction of time)

Another issue that we encountered was slow convergence when batch size is large (e.g. 100 patterns at once). However, this problem disappeared when batch size was reduced to 10. The cost function was very noisy, but satisfactory learning occurred after only one epoch (see Fig. 3). This is particularly useful in a standalone application with learning module, as it would require less memory for learning incoming new EEG patterns.

To evaluate the classifiers, we separated the sample into three parts: 20% of consecutive patterns for 22 patients was left as a testing sample (i.e. to see how trained classifier performs for the same patients but with unseen data), 80% of first consecutive patterns was used for network architecture selection by 10-fold cross validation, and 1 patient was left to see how classifier performs on completely unseen patients.

The first architecture of ConvNet, which we considered was this: a pattern is filtered with 5 convolutional filters of dimensions $1 \times 5$; resulting images were then pooled along the time axis by factor of two; then 5 filters of dimension $1260 \times 3$ were applied with resulting images pooled again along the time axis by factor of two. No fully connected layer was applied except for the output layer with two nodes for each label "seizure" and "no seizure".

The classification results for this architecture were as follows: 91% of all tested seizure periods were identified. For two patients two seizures were completely missed. For the non-seizure periods, convolutional network predicted 0.7 false positives per one hour on average.

We also considered other architectures of our ConvNet. One of the choices was to increase the number of filters. However, this completely failed, because now three seizures for three different patients were completely missed. One reason might be that overfitting occurred. This will be investigated in the future.

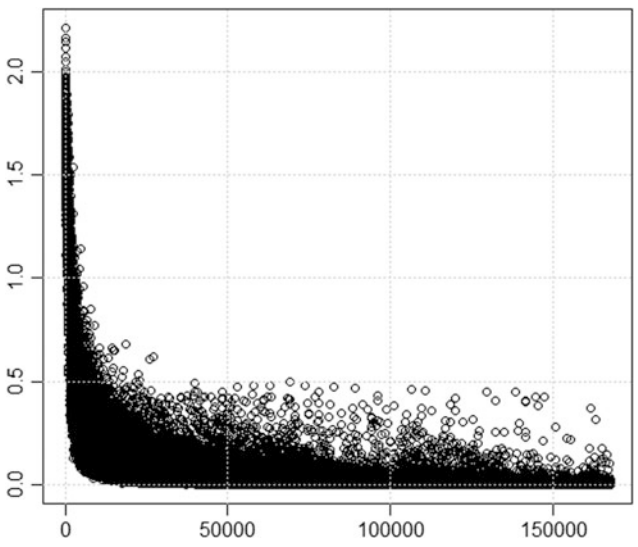



**Fig. 3** Cost function after one epoch

Another architecture was as follows: 5 filters of dimensions $3 \times 5$, followed by pooling layer with factor two in both dimensions, as opposed to the previous architecture; then a layer with 5 filters of dimension $628 \times 3$ and pooling along both dimensions. For seizure periods, we had accuracy of detection the same as for the first architecture, i.e. 91%. In addition, we obtained better results with 0.4 false positives per hour on average.

This increase in accuracy of non-seizure period detection might be due to the fact, that filters with dimensions $3 \times 5$ enables an interaction between different channel pair bivariate features. This fact should be further investigated whether it is possible to push even further classification accuracy.

## Transfer Learning

Our considered ConvNet architectures were sufficiently simple to cause no difficulties of calculations during the forward pass of data. This is an attractive feature of an algorithm when one considers implementation of it as a standalone application. On the other hand, the training of the network is not such a light task and requires more resources so that a standalone application, with additional learning module, but without large computational resources, is somewhat less realistic.

In case of no learning module, the application would simply observe the patterns of patients' EEG and produce labels "seizure" or "no seizure". To do this accurately, one should implement and test many other ConvNet architectures or pre-train the network on very large number of dataset, so that high variety of seizure patterns would be observed and learned. However, considering the complexity of brain and specificity of patients, this would be difficult or almost impossible to achieve completely.

We tried our classifier on completely new data sample, i.e. we checked our classifier on a patient whose EEG data were not used for training. However, only one out of seven seizures were identified with 0 false alarms. That's clearly a poor result.

Next, we retrained our classifier on 4000 samples of that patient. During this period, there were three episodes of seizures. Then the retrained classifier identified all of the remaining 4 seizures with 0 false alarms, which is excellent case.

So, on the other hand, if one would be able to implement a cheap (in terms of computational resources) learning module into a standalone application, the ConvNet, even pre-trained on small amount of data, would be able to readjust its parameters for patient-specific brain patterns. In

this case various techniques could be also applied for transfer learning in seizure classification.

## Conclusions

Considering analysis of scalp EEG of patients with epilepsy, even though non-invasive EEGs are plagued with high level of noise due to various artefacts, convolutional neural networks show high potential for automatic seizure detection systems. There are many problems to solve before a standalone application for seizure early alert can be developed: small data samples for training – one possibility would be to use data augmentation techniques as in image classification; high variability of the EEG measurements due to patient-specific artefacts and natural variability between patients – because of this the application with low-resource learning module integrated together is more realistic as compared to the application with only pre-trained classifier.

We believe that the aforementioned problems could be at least partially overcome – convolutional neural network architecture as well as data augmentation might be the key.

**Acknowledgements** T. Iesmantas was supported by the postdoctoral fellowship grant, provided by the Kaunas University of Technology, Faculty of Mathematics and Natural Sciences. T. Iesmantas was partially supported by the postdoctoral fellowship grant, provided by the Kaunas University of Technology, Faculty of Mathematics and Natural Sciences. In addition, part of the research presented in this paper was based upon work from COST Action (ENJECT TD 1405), supported by COST (European Cooperation in Science and Technology).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Pereda E, Quian Quiroga R, Bhattacharya J (2005) Nonlinear multivariate analysis of neurophysiological signals. Prog Neurobiol 77:1–37
2. Shoeb A, Guttag J (2010) Application of machine learning to epileptic seizure detection. In: Proceedings of the 27th international conference on machine learning, Haifa, Israel, pp 975– 982
3. Mirowski P, Madhavan D, LeCun Y, Kuzniecky R (2009) Classification of patterns of EEG synchronization for seizure prediction. Electroencephalogr Clin Neurophysiol 120 (11):1927–1940
4. Mirowski P, Le Cun Y, Madhavan D, Kuzniecky R (2008) Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG. IEEE Workshop on Machine Learning for Signal Processing
5. Williamson JR, Bliss DW, Browne DW, Narayanan JT (2012) Seizure prediction using EEG spatiotemporal correlation structure. Epilepsy Behav 25(2):230–238
6. Nasehi S, Pourghassem H (2012) Seizure detection algorithms based on analysis of eeg and ecg signals: a survey. Neurophysiology 44(2):174–186
7. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1(4):541–551
8. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1990) Handwritten digit recognition with a back-propagation network. In: Touretzky David (ed) Advances in neural information processing systems 2 (NIPS*89). Denver, CO
9. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. In: Proceedings of the IEEE, Nov 1998
10. Myers M, Padmanabha A, Hossain G, de Jongh Curry AL, Blaha CD (2016) Seizure prediction and detection via phase and amplitude lock values. Front Hum Neurosci 10(80)
11. Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. Hum Brain Mapp 8:194–200

# Prediction of Cardiac Arrest in Intensive Care Patients Through Machine Learning

E. Akrivos, V. Papaioannou, N. Maglaveras, and I. Chouvarda

**Abstract**

Cardiac arrest is a critical health condition characterized by absence of traceable heart rate, patient's loss of consciousness as well as apnea, with inhospital mortality of ~80%. Accurate estimation of patients at high risk is crucial to improve not only the survival rate, but also the quality of life as patients who survived from cardiac arrest have severe neurological effects. Existing research has focused on demonstrating static risk scores without taking account patient's physiological condition. In this study, we are implementing an integrated model of sequential contrast patterns using Multichannel Hidden Markov Model. These models can capture relations between exposure and control group and offer high specificity results, with an average sensitivity of 78%, and have the ability to identify patients in high risk.

**Keywords**

Cardiac arrest • Prediction • MC-HMM • Sequential pattern recognition • Classification

## Introduction

Cardiac arrest is defined as interruption of mechanical activity of heart, which is confirmed by absence of traceable heart rate, patient's loss of consciousness, as well as apnea, according to the Utstein style. Cardiac arrest is defined as inpatient when it occurs in a hospitalized patient who had pulse on admission to the hospital [1]. Common causes of cardiac arrest are ventricular fibrillation (VF), ventricular tachycardia (VT), asystole and electrical activity of the heart without pulses. About 200,000 cases of inpatient cardiac arrest are reported each year in U.S.A. (United States of America) [2]. Cardiac arrest occurs in 1–5 per 1000 hospitalized patients and ~20% survive until their discharged [1, 3, 4]. Generally, patients at high risk of cardiac arrest have co-morbidities, which affect their health outcome and recovery after cardiac arrest [2]. Studies have shown that clinical signs of deterioration, such as hemodynamic instability and respiratory distress, of patients within a period of eight hours prior to cardiac arrest could be used to avoid cardiac arrest in 84% of these [1]. However, the recognition of the causes of cardiac arrest, has been shown to increase the survival rate of patients within an hour of episode by about 29% and by 19% until their discharge [3]. Therefore, early and accurate detection of patients at-risk is critical to improve health outcome and survival rate.

Increasing use of electronic health records (EHR) leads to greater accessibility and availability of medical data. The

E. Akrivos (✉) · N. Maglaveras · I. Chouvarda
Lab of Computing Medical Informatics and Biomedical Imaging Technologies, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: e_akrivos@icloud.com

E. Akrivos
Department of Internal Medicine, 424 Military Hospital of Thessaloniki, Thessaloniki, Greece

V. Papaioannou
Alexandroupolis University Hospital/Intensive Care Unit, Alexandroupoli, Greece

N. Maglaveras
McCormick School of Engineering & Applied Sciences, Department of Electrical Engineering & Computer Science, Northwestern University, Evaston, IL 60201, USA



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_5

Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC II) database was developed from medical data of over 30,000 patients during 2001–2008 from Beth Israel Deaconess Medical Center in Boston. MIMIC II is the most extensive resource of intensive care unit (ICU) medical data and it is available to the public [5, 6].

Recent research used measurements of vital signs, such as blood pressure, respiratory rate, temperature and healthcare professional's opinion to model early warning scores to identify patients at high risk of cardiac arrest [7–9]. However, these researches could not predict the accurate time of cardiac arrest. DYNACARE is a model based on dynamic time series attending to predict the time of cardiac arrest [10].

The present study proposes an approach that discovers sequential contrast patterns from commonly observed measurements, such as blood pressure, respiratory rate and heart rate, transforming the classical time series data to a sequence of patterns for implementation of a classifier for cardiac arrest. Following, the classifier is used to predict the likelihood of a sequence of patterns to belong in cardiac arrest class. This method has been used for the prediction of Sepsis [11], but to our knowledge it is now applied to cardiac arrest prediction for the first time.

## Materials and Methods

The study was conducted with data from MIMIC-II for adult patients (age 18+ on ICU admission) aged up to 90 years who were hospitalized in the Cardiological ICU and experienced a recorded cardiac arrest episode according to ICD-9 (International Classification of Diseases) 427.5 for cardiac arrest. The study focused on different types of variables, such as demographic data, vitals signs, medication and laboratory measurements. Patient data was discretized in 2-hour bins. An additional requirement for each patient was to have at least 36 measurements (3 days of hospitalization) to ensure sufficient data points. There were 698 patients with a cardiac arrest diagnosis from 27,542 of MIMIC-II database, from which only 162 met the minimum data criteria. Patients who have been diagnosed with highrisk heart diseases for cardiac arrest and have not occured an event of cardiac arrest, were selected as control group, with diseases such as coronary heart disease, myocardial infarction, major heart disease, valvular heart disease, congenital heart disease and heart rhythm abnormalities such as Brugada syndrome and long QT [12–22]. The selected ICD-9 codes for these diseases was 414.01, 410.90, 429.3, 424.0, 424.1, 746.0–746.9, 746.89 and 426.82. Similar data criteria to patients with cardiac arrest were also used in control population, with a final number of control population 5,278 patients.

### Data Preparation and Preprocessing

The first step was extraction of data from Mimic Database in flat files. Quality inspection revealed a number of missing values in different data fields. Missing data was processed using the Multiple Imputation method and predictive mean matching (PMM) algorithm [23]. In order to increase similarity of considered cases, medication, demographic data and laboratory measurements were used as coefficients for PMM to predict missing values of heart rate, systolic blood pressure, diastolic blood pressure, respiratory rate, $PO_2$ and $PCO_2$.

Following, the quantization of measurements and mapping to specific states was necessary for sequential pattern analysis, since pattern discovery methods are more effective on symbolic data types. Frequent sequence patterns methods [24–26] is used to identify patterns and frequency support in sequences between the two classes of sequence data.

### Mining Sequential Contrast Patterns

Emerging patterns (EPs) are described as patterns that satisfied specific user-defined frequency rules for different classes of data. This means that in a categorized data in two categories, positive (cardiac arrest group) and negative (control group), the patterns must have a high frequency support in the positive category and a low frequency support in the negative category. Since EPs have these characteristics, they are considered to be distinct patterns and have the ability to distinguish the contrast between the two categories (also known as growth rate of EP). Therefore, the strength of EPs is expressed by the ratio of frequency in both classes.

Extending the above description, a sequence pattern $S_p$ can be characterized as sequential contrast pattern if satisfy the conditions (a) and (b) depicted below in Eqs. (1) and (2)

(a) Positive support:

$$counts_{S_p}(D^+, g) \geq \alpha \tag{1}$$

(b) Negative support:

$$counts_{S_p}(D^-, g) \leq \beta \tag{2}$$

where $D^+$, $D^-$ two different datasets with labels, such as positive sequences and negative sequences, respectively, g is the gap-constraint, $counts_{S_p}(D, g)$ the frequency support of a

**Table 1** Contrast patterns for heart rate sequences

| Heart Rate patterns | Pattern-id |
|---|---|
| Tachyc < N1hr | HR1 |
| Tachyc < Nhr | HR2 |
| Tachyc < N1hr < N1hr | HR3 |
| Tachyc < Tachyc | HR4 |

**Table 2** Contrast patterns for systolic blood pressure sequences

| Systolic BP patterns | Pattern-id |
|---|---|
| Nbpsys < HypotensS | SB1 |
| HypotensS < Nbpsys | SB2 |
| Nbpsys < Nbpsys < HypotensS | SB3 |

**Table 3** Contrast patterns for respiratory rate sequences

| Respiratory Rate patterns | Pattern-id |
|---|---|
| Bradypnoea-FP-FP | RR1 |
| FP-FP-FP | RR2 |
| FP-FP-Bradypnoea | RR3 |

sequence pattern $S_p$, $\alpha$ and $\beta$ thresholds for frequency support in two datasets. Thus, discovered patterns lead to mining sequential contrast patterns, given the above characteristics, which must satisfy (a) and (b) condition [11].

In the present study, using the above description, after descretization of variables based on normal value's cut offs, resulted in contrast patterns for three variables, where $\alpha = 0.7$, $\beta = 0.5$ and $g = 2$. Variables with contrast patterns were heart rate, systolic blood pressure and respiratory rate. Tables 1, 2 and 3 show the contrast patterns and their unique identification name with which they were replaced.

A sliding window with length equivalent to the longest pattern (length = 3) was used to transform the discrete sequences of data to sequences of contrast patterns, with purpose to use these as input data to HMM, instead of ordinary time series sequences. Table 4 shows the above transformation from a patient's sequences.

## Multichannel Hidden Markov Model

Multi Channel Hidden Markov Models (MC-HMM) are an extension of the conventional form of Hidden Markov Models (HMMs) for multiple variable or channel data sequences. MC-HMM has been used on applications such as speech recognition, activity recognition, anomalous trading activities, medical events, disease interactions and fault diagnosis [27–30]. In the present study, MC-HMM was used to model interactions between multiple clinical measurements, which are represented as sequential contrast patterns. According to the theory of MC-HMM, each discrete state for each channel is individually transformed into a three-state mode, based on the markov property. Therefore, it appears that the probability of transition and emission for each state can be mapped as a mutation of the three unique states that correspond to each channel. Two MC-HMM's constructed for the two classes of data. The first MC-HMM was trained by expectation maximization (EM) algorithm for patients who belong in cardiac arrest class, while the second one was trained for patients of control group.

## Results

For prediction of cardiac arrest, the 8-fold cross validation method performed. For each dataset, cardiac arrest and control dataset respectively, 7 folds randomly selected used as training datasets for each model respectively and 1 fold for test set. Thus, each model was trained to find the sequences that belong to their class. Test set from cardiac arrest patient's data was containing only the sequences from observational window before the onset of cardiac arrest, for the classification purposes. Test sets from the two datasets were merged and likelihood for each patient's sequence computed for the two models. If the likelihood of the sequence patterns of the cardiac arrest patient's model was greater than the control pattern then the patient was considered to belong in class with patients at higher risk of

**Table 4** Example of transformation from discrete patient sequence to sequence of contrast pattern

| Discrete sequences | Variable | Contrast pattern ID |
|---|---|---|
| N1hr-N1hr-Tachyc-N1hr-Nhr-N1hr-….-Nhr-Tachyc-N1hr-N1hr-N1hr-Nhr-Nhr-N1hr | HR | HR1-HR3-HR2-HR3-X-X-X-X |
| HypotensS-Nbpsys-Nbpsys-Nbpsys-…-Nbpsys-HypotensS-Nbpsys-Nbpsys-Nbpsys-Nbpsys | SysBP | SB2-SB3-SB1-SB2-X-X-X-X |
| Bradypnoea-FP-RRnorm-FP-RRnorm-…-RRnorm-RRnorm-FP-RRnorm-FP-Bradypnoea-FP-RRnorm | RR | RR1-RR3-X-X-X-X-X-X |

**Table 5** Statistical results for 8 fold cross validation prediction

| Sensitivity (mean ± SD) | Specifity (mean ± SD) |
|---|---|
| 0.78 (±0.04) | 0.43 (±0.02) |

cardiac arrest and were categorized respectively. Table 5 shows statistical results from prediction.

## Discussion

The study's results have evidence that integrating MC-HMM models with sequential contrast patterns as input data can perform well to predict cardiac arrest. A limitation in selection criteria of cardiac arrest group, which led to the reduced sample, was that patients with respiratory cardiac arrest were excluded. By discovering patterns, based on the contrast of their frequencies on data of two populations, intervention and control respectively, it is possible to interpret the difference between the two data populations. In particular, the use of $\alpha$ and $\beta$ thresholds to calculate the growth rate of a pattern is equivalent to the odds ratio, which is used in medical research to find relationships between an exposure and an outcome. However, false positive rate is high. This issue was the result of the restricted design of population groups with ICD-9 code. In the present study, cardiac arrest diagnosis was one of the criteria for patient selection, while VT and VF diagnosis were criteria for control group. This issue poses the problem of semantically defining and selecting the correct cases within a rich database. Furthermore, the significance of a contrast pattern is determined by the growth rate, which if it is too high it creates few patterns, and if it is too low it creates patterns without significance. Therefore, creating an algorithm for optimal selection of the threshold value for pattern development is necessary in order to find important contrast patterns.

## Conclusion

In this study, an attempt was made to model cardiac arrest by using an integrated framework, which was previously successfully tested to predict septic shock [11]. However, while the results are promising, it became obvious that the complexity of cardiac arrest mechanism poses many difficulties in modeling. Thus, the present study demonstrates the importance of using sequential contrast patterns to capture relations between groups.

**Conflict of Interest** The authors declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

## References

1. Sandroni C, Nolan J, Cavallaro F, Antonelli M (2007) In-hospital cardiac arrest: incidence, prognosis and possible measures to improve survival. Intensive Care Med 33(2):237–245
2. Graham R, McCoy, MA, Schultz AM (2015) Committee on the treatment of cardiac arrest: current status and future directions, Board on Health Sciences Policy, Institute of Medicine Strategies to improve Cardiac Arrest Survival: A Time to Act. Washington (DC). National Academies Press (US), 29 Sept 2015
3. Bergum D, Haugen BO, Nordseth T, Mjølstad OC, Skogvoll E (2015) Recognizing the causes of in-hospital cardiac arrest-A survival benefit. Resuscitation. 97:91–96
4. Nolan JP, Soar J, Smith GB, Gwinnutt C, Parrott F, Power S et al (2014) Incidence and outcome of in-hospital cardiac arrest in the United Kingdom National Cardiac Arrest Audit. Resuscitation. 85 (8):987–992
5. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG et al (2000) Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation 101(23):E215–E220
6. Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LW, Moody G et al (2011) Multiparameter intelligent monitoring in intensive care ii: a public-access intensive care unit database. Crit Care Med 39(5):952–960
7. Smith AF, Wood J (1998) Can some in-hospital cardiorespiratory arrests be prevented? A prospective survey. resuscitation. 37 (3):133–137
8. Hodgetts TJ, Kenward G, Vlachonikolis IG, Payne S, Castle N (2002) The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team. Resuscitation. 54(2):125–131
9. McBride J, Knight D, Piper J, Smith GB (2005) Long-term effect of introducing an early warning score on respiratory rate charting on general wards. Resuscitation. 65(1):41–44
10. Ho JC, Park Y, Carvalho CM, Ghosh J (2013) DYNACARE: dynamic cardiac arrest risk estimation. J Mach Learn Res 31:333–341

11. Ghosh S, Li J, Cao L, Ramamohanarao K (2017) Septic shock prediction for ICU patients via coupled HMM walking on sequential contrast patterns. J Biomed Inf 66:19–31
12. Longo DL et al (2015) Cardiovascular collapse, cardiac arrest, and sudden cardiac death. In: Harrison's Principles of Internal Medicine, 19th edn, New York
13. Sudden cardiac arrest. http://www.nhlbi.nih.gov/health/health-topics/topics/scda/
14. Podrid PJ. Overview of sudden cardiac arrest and sudden cardiac death. http://www.uptodate.com/home
15. American Heart Association. Heart attack or sudden cardiac arrest: How are they different? http://www.heart.org/HEARTORG/Conditions/More/MyHeartandStrokeNews/Heart-Attack-or-Sudden-Cardiac-Arrest-How-Are-They-Different_UCM_440804_Article.jsp-Vi55p36rTIU
16. Neumar RW, Shuster M, Callaway CW, Gent LM, Atkins DL et al (2015) Part 1: executive summary: 2015 American heart association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation. 132(18 Suppl 2): S315-167
17. Arrhythmia. National Heart, Lung, and Blood Institute. http://www.nhlbi.nih.gov/health/health-topics/topics/arr
18. Fuster V et al (2011) Sudden cardiac death in hurst's the heart, 13th edn. The McGraw-Hill Companies, New York
19. Goldberger AL et al (2013) Sudden cardiac arrest and sudden cardiac death in clinical electrocardiography: a Simplified Approach, 8th edn. Saunders Elsevier, Philadelphia
20. Association. AH. Ejection fraction heart failure measurement. http://www.heart.org/HEARTORG/Conditions/HeartFailure/SymptomsDiagnosisofHeartFailure/Ejection-Fraction-Heart-Failure-Measurement_UCM_306339_Article.jsp-Vi58RH6rTIU
21. Riggin EA. Allscripts EPSi. Mayo Clinic, Rochester, Minn
22. Rohren CH (expert opinion). Mayo Clinic, Rochester, Minn
23. Vink G, Frank LE, Pannekoek J, van Buuren S (2014) Predictive mean matching imputation of semicontinuous variables. Stat Neerl 68(1):61–90
24. Klema J, Novakova L, Karel F, Stepankova O (2008) Sequential data mining: A comparative case study in development of atherosclerosis risk factors. Syst Man Cybern Part C Appl Rev IEEE Trans 38(1):3–15
25. Baralis E, Bruno G, Chiusano S, Domenici VC, Mahoto NA, Petrigni C (2010) Analysis of medical pathways by means of frequent closed sequences. In: Knowledge-based and intelligent information and engineering systems, pp. 418–425
26. Berlingerio M, Bonchi F, Giannotti F, Turini F (2007) Time-annotated sequences for medical data mining. In: Seventh IEEE international conference on data mining workshops. IEEE, pp 133–138
27. Audhkhasi K, Osoba O, Kosko B (2013) Noisy hidden Markov models for speech recognition. In: The 2013 international joint conference on neural networks (IJCNN)
28. Cao L, Ou Y, Yu PS (2012) Coupled behavior analysis with applications. IEEE Trans Knowl Data Eng 24(8):1378–1392
29. Masoudi S, Montazeri N, Shamsollahi MB, Ge D, Beuche A, Pladys P, et al (2013) Early detection of apnea-bradycardia episodes in preterm infants based on coupled hidden Markov model. IEEE international symposium on signal processing and information technology
30. Zhou H, Chen J, Dong G, Wang H, Yuan H (2016) Bearing fault recognition method based on neighbourhood component analysis and coupled hidden Markov model. Mech Syst Signal Process 66–67:568–581

# Part II
# Scientific Challenge—Lung Sounds Analysis

# A Respiratory Sound Database for the Development of Automated Classification

B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, and N. Maglaveras

**Abstract**

The automatic analysis of respiratory sounds has been a field of great research interest during the last decades. Automated classification of respiratory sounds has the potential to detect abnormalities in the early stages of a respiratory dysfunction and thus enhance the effectiveness of decision making. However, the existence of a publically available large database, in which new algorithms can be implemented, evaluated, and compared, is still lacking and is vital for further developments in the field. In the context of the International Conference on Biomedical and Health Informatics (ICBHI), the first scientific challenge was organized with the main goal of developing algorithms able to characterize respiratory sound recordings derived from clinical and non-clinical environments. The database was created by two research teams in Portugal and in Greece, and it includes 920 recordings acquired from 126 subjects. A total of 6898 respiration cycles were recorded. The cycles were annotated by respiratory experts as including crackles, wheezes, a combination of them, or no adventitious respiratory sounds. The recordings were collected using heterogeneous equipment and their duration ranged from 10 to 90 s. The chest locations from which the recordings were acquired was also provided. Noise levels in some respiration cycles were high, which simulated real life conditions and made the classification process more challenging.

**Keywords**

Respiratory sounds • Adventitious sounds • Automated classification • Database

## Introduction

Respiratory diseases cause an immense health, economic and social burden and are the third leading cause of death worldwide [1] and a significant burden for public health systems [2]. Therefore, significant research efforts have been dedicated to improving early diagnosis and routine monitoring of patients with respiratory diseases to allow for timely interventions [3]. A great amount of research has been focused in the auscultation and characteristics of respiratory sounds (RS), as they are directly related to movement of air, changes within the lung tissue, and position of secretions within the tracheobronchial tree, which make them valuable indicators of respiratory health and respiratory disorders [4].

B. M. Rocha (✉) · L. Mendes · R. P. Paiva · P. Carvalho
Centre for Informatics and Systems, University of Coimbra, DEI, Polo 2, Pinhal de Marrocos, Coimbra, Portugal
e-mail: bmrocha@dei.uc.pt

D. Filos · I. Vogiatzis · E. Perantoni · E. Kaimakamis · P. Natsiavas · I. Chouvarda · N. Maglaveras
Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

A. Oliveira · C. Jácome · A. Marques
Lab3R—Respiratory Research and Rehabilitation Laboratory, School of Health Sciences, University of Aveiro, Aveiro, Portugal

A. Oliveira · A. Marques
Institute for Biomedicine (iBiMED), University de Aveiro, Aveiro, Portugal



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_6

Respiratory sounds are generally classified as normal or adventitious. Auscultation-based diagnosis and monitoring of respiratory conditions rely heavily on the presence of adventitious sounds and on the altered transmission characteristics of the chest wall. Adventitious sounds are RS superimposed on normal respiratory sounds which can be discontinuous (crackles) or continuous (wheezes). Crackles are discontinuous, explosive, and non-musical adventitious RS that occur frequently in cardiorespiratory diseases [5]. They are usually classified as fine and coarse crackles based on their duration, loudness, pitch, timing in the respiratory cycle, and relationship to coughing and changing body position [6]. Wheezes are musical RS that usually last more than 250 ms. They are a common clinical sign in patients with obstructive airway diseases such as asthma and chronic obstructive pulmonary disease (COPD) [7].

In the first edition of the ICBHI Scientific Challenge, participants were asked to develop algorithms that characterize sound recordings collected from clinical and non-clinical (such as in-home visits) environments. The goal was to classify, for each respiratory cycle of a short recording (10–90 s), acquired at a single location, whether the respiratory cycle contained crackles, wheezes, or both.

To develop solutions for the challenge, participants had access to a respiratory sound database containing various events (e.g., noise, cough, wheezes, crackles) collected from healthy people and patients with different respiratory conditions (e.g., COPD, asthma), providing a variety of signal sources. Data included not only clean respiratory sounds but also noisy recordings, providing authenticity to the challenge. Data were recorded from different locations, depending on the individual protocols used for each data set. The database was annotated by health professionals. The ICBHI challenge process has been supported by a dedicated web application.[1] Users registered in the contest, accessed the provided datasets, submitted their source code, and communicated in forums provided by this web platform. Furthermore, the platform automatically informed end-users of their submissions' ranking and evaluation. Finally, the platform actively enforced rules of the contest, e.g., the number of submissions allowed in each contest phase.

The automatic detection or classification of adventitious RS has been the subject of many studies in the last decades. Pramono et al. [8] summarized the most relevant methods employed in those studies. Algorithms developed to detect or classify events usually involve two steps; adventitious RS are no exception. The first step is to extract the relevant features that will be used as detection or classification variables. The second step is to use detection or classification techniques on the data, based on the features extracted. The most common features employed in the literature include Mel-frequency cepstral coefficients (MFCCs), spectral features, energy, entropy, and wavelet coefficients. Machine learning algorithms proposed in the literature include empirical rule-based methods, support vector machines (SVMs), artificial neural networks (ANNs), Gaussian mixture models (GMMs), k-nearest neighbors (k-NNs), and logistic regression models.

Most prior attempts on automated classification of respiratory sounds have been limited by the small number of patients employed in the studies. It is possible to achieve very good classification results because the algorithm can be custom designed and fit carefully to match the data and the features collected from a small number of patients. However, as the number of patients is increased to several dozen or several hundred, the features learned from small datasets typically fail to generalize [9].

This paper is structured as follows: in the Challenge data section, we describe the data collection process, as well as the structure of the challenge; future uses of the database are discussed in the Conclusion section.

## Challenge Data

### Data Collection

The ICBHI Scientific Challenge database contains audio samples, collected independently by two research teams in two different countries, over several years. The database consists of a total of 5.5 h of recordings containing 6898 respiratory cycles, of which 1864 contain crackles, 886 contain wheezes, and 506 contain both crackles and wheezes, in 920 annotated audio samples from 126 subjects.

#### School of Health Sciences, University of Aveiro (ESSUA)

Most of the database consists of audio samples recorded by the ESSUA research team at Respiratory Research and Rehabilitation Laboratory (Lab3R), ESSUA and at Hospital Infante D. Pedro, Aveiro. Sounds from several studies conducted by this research team were included in the database. All the recordings followed the computerized RS analysis guidelines for short-term acquisitions [10], collecting sounds from seven chest locations: trachea; left and right anterior, posterior, and lateral. Sounds were collected in clinical and non-clinical (home) settings. The acquisition of RS was performed on subjects of all ages, from infants to adults and elderly people. Subjects included patients with lower respiratory tract infections, upper respiratory tract infections, COPD, asthma, and bronchiectasis.

In some studies, the sounds were collected sequentially with a digital stethoscope (Welch Allyn Master Elite Plus

[1] https://bhichallenge.med.auth.gr/.

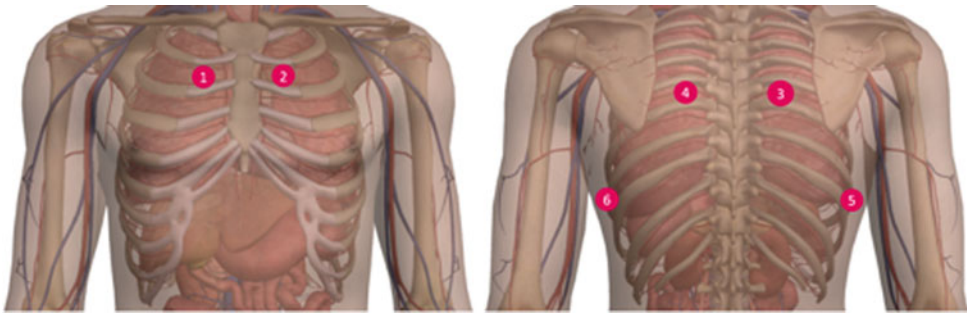



**Fig. 1** Chest locations for the recording of respiratory sounds

Stethoscope Model 5079-400). In other studies, the sounds were collected using either seven stethoscopes (3 M Littmann Classic II SE) with a microphone in the main tube or seven air-coupled electret microphones (C 417 PP, AKG Acoustics) located into capsules made of Teflon. Respiratory sounds were annotated using the Computerised Lung Auscultation – Sound System (CLASS) [11].

### Aristotle University of Thessaloniki (AUTH)

Respiratory sounds were acquired at the Papanikolaou General Hospital, Thessaloniki and at the General Hospital of Imathia (Health Unit of Naousa), Greece. Sounds were collected sequentially from six chest locations, as shown in Fig. 1. The acquisition of RS was performed on adult and elderly patients. All patients had COPD with comorbidities (e.g. heart failure, diabetes, hypertension).

These recordings were acquired as part of the European project WELCOME (Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Comorbidities) project and were annotated using Audacity[2] 2.0.6 a free, open source, cross-platform software for recording and editing sounds.

## Data Annotation and Curation

### ESSUA

Sounds annotation by respiratory experts is the most common and reliable method to assess the robustness of algorithms to detect adventitious RS [12]. Two respiratory physiotherapists and one medical doctor, with experience in visual-auditory crackles/wheezes recognition, independently annotated the sound files in terms of presence/absence of adventitious sounds and identification of breathing phases. Nevertheless, as annotation is a time–consuming process, being difficult to conduct in a large amount of sound files, in part of ESSUA database, only one respiratory physiotherapist annotated the files. For the annotation, the Respiratory Sound Annotation Software was used (Fig. 2) [13].

### AUTH

Respiratory sound annotations were performed by three experienced physicians, two specialized pulmonologists and one cardiologist. Annotations discriminated the following sounds: normal (respiratory sound), fine crackles, coarse crackles, wheezing, speech, cough, artifact. Figure 3 reproduces a sample of the annotation process. Figure 4 shows an example of an annotated sound recording.

## Previous Uses of Data for Classification

Part of the ESSUA database has been used previously for the detection of crackles. Pinho et al. [14] developed an algorithm for automatic crackle detection and characterization and evaluated its performance and accuracy against a multi-annotator gold standard. The developed algorithm was based on three main procedures: (i) extraction of a window of interest of a potential crackle (based on fractal dimension and box filtering techniques); (ii) verification of the validity of the potential crackle considering computerized RS analysis established criteria; and (iii) characterization and extraction of crackle parameters. The paper reported a performance of 89% sensitivity and 95% precision.

Part of the AUTH database has been used previously for the detection of wheezes, crackles, and cough. Mendes et al. proposed a method for the detection of wheezes based on their distinct signature in the spectrogram space (WS-SS). In addition to this feature, 29 musical features were computed using the MIR Toolbox [15]. The paper reported a performance of 91% sensitivity and 99% specificity.

Mendes et al. [16] proposed a method for the detection of crackles using a multi-feature approach. 35 features were extracted, including 31 musical features, a wavelet-based feature, entropy and Teager energy. WS-SS was also extracted to improve the robustness of the method against the presence of wheezes. The paper reported a performance of 76% sensitivity and 77% precision. Rocha et al. [17] proposed a method for the detection of explosive cough events based on a combination of spectral content

[2] http://audacity.sourceforge.net/.

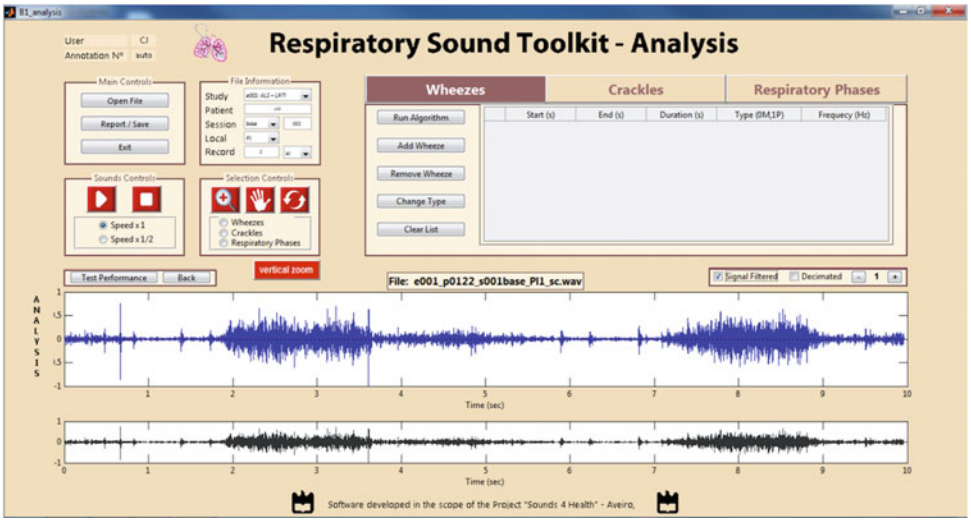



**Fig. 2** Respiratory sound annotation software

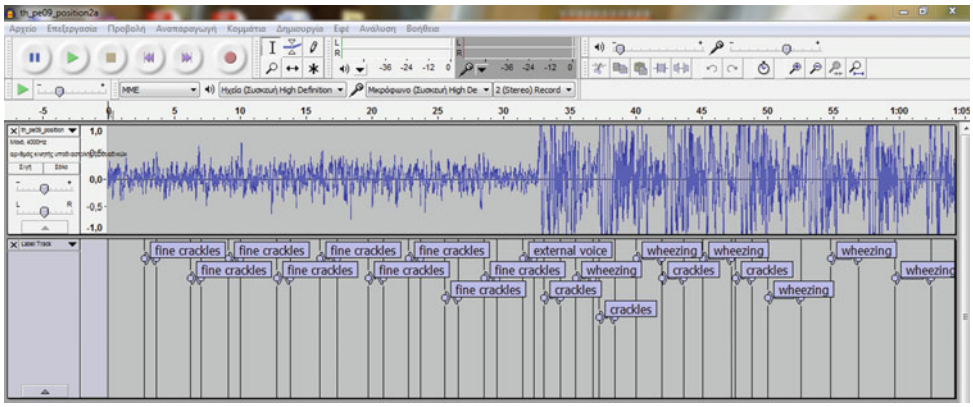



**Fig. 3** A sample of the respiratory sound annotation process

descriptors and pitch-related features. The paper reported a performance of 92% sensitivity and 85% specificity.

## Preparation of the Database for the ICBHI Challenge

The challenge was structured in two phases: unofficial and official. During each phase, data from the two aforementioned databases were divided into training (60%) and testing (40%) sets.

The data included in each of the train/test sets were derived from mutually exclusive populations and thus the recordings from the same subject could not be present in both the training and testing sets. Furthermore, the data included in the database were anonymized and no personal information were provided.

During the official phase of the challenge, the training set included 2063 respiratory cycles from 539 recordings derived from 79 subjects, while the testing set included 1579 respiration cycles from 381 recordings derived from 49 patients. Table 1 provides further details about the distri-

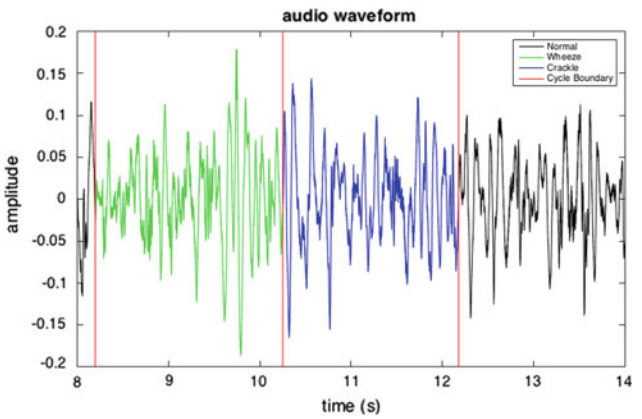



**Fig. 4** A segment including three respiratory cycles: the first contains wheezes (green), the second contains crackles (blue), and the third is normal (black). Respiratory cycle boundaries are represented by vertical lines (red)

**Table 1** Summary of the training and testing sets used in the official phase of the ICBHI challenge

| Database | Testing set | | | Training set | | |
|---|---|---|---|---|---|---|
| | ESSUA | AUTH | All | ESSUA | AUTH | All |
| #patients | 38 | 11 | 49 | 72 | 7 | 79 |
| #recordings | 317 | 64 | 381 | 507 | 32 | 539 |
| #wheezes | 588 | 61 | 649 | 459 | 42 | 501 |
| #crackles | 273 | 112 | 385 | 1104 | 111 | 1215 |
| #crackles + wheezes | 106 | 37 | 143 | 335 | 28 | 363 |
| #normal | 1216 | 363 | 1579 | 1740 | 323 | 2063 |

bution of the adventitious RS between the datasets.

## Conclusion

The creation of this database and the related scientific challenge constitute an initial but decisive step towards leveraging computational lung auscultation, and also towards highlighting the complexity of the RS classification problem. The availability of the database after the challenge (details will be posted on the challenge's website), along with the challenge's approaches and results, will set the basis to ensure the continuation of efforts, hopefully inspiring and facilitating future relevant competitions.

**Acknowledgements** The authors would like to thank the health professionals and the patients who have agreed to participate in the data collection process. This work was financially supported by the EU project WELCOME (FP7-ICT-2013-10/611223) and the FEDER/COMPETE/FCT project UID/BIM/04501/2013. Finally, the authors would like to thank IFMBE for endorsing and supporting this scientific challenge.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. World Health Organization (2015) The top 10 causes of death
2. Gibson GJ, Loddenkemper R, Lundbäck B, Sibille Y (2013) Respiratory health and disease in Europe: the new European Lung White Book. Eur Respir J 42:559–563
3. Marques A, Oliveira A, Jácome C (2014) Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review. Respir Care 59(5):765–776
4. Earis J, Cheetham B (2000) Current methods used for computerized respiratory sound analysis. Eur Respir Rev 10(77):586–590
5. Piirila P, Sovijarvi AR (1995) Crackles: recording, analysis and clinical significance. Eur Respir J 8(12):2139–2148
6. Sarkar M, Madabhavi I, Niranjan N, Dogra M (2015) Auscultation of the respiratory system. Ann Thorac Med 10(3):158
7. Sovijarvi ARA, Malmberg LP, Charbonneau G, Vanderschoot J, Dalmasso F, Sacco C, Rossi M, Earis JE (2000) Characteristics of breath sounds and adventitious respiratory sounds. Eur Respir Rev 10:591–596
8. Pramono RXA, Bowyer S, Rodriguez-Villegas E (2017) Automatic adventitious respiratory sound analysis: a systematic review. PLoS ONE 12(5):e0177926
9. Chamberlain D, Kodgule R, Ganelin D, Miglani V, Fletcher RR (2016) Application of semi-supervised deep learning to lung sound analysis. In: 38th annual international conference of IEEE engineering in medicine and biology society, pp 804–807
10. Rossi M, Sovijarvi ARA, Piirila P, Vannuccini L, Dalmasso F, Vanderschoot J (2000) Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings. Eur Respir Rev 10:611–615
11. Machado A, Oliveira A, Jácome C, Pereira M, Moreira J, Rodrigues J, Aparício J, Jesus LMT, Marques A (2017) Usability of Computerized Lung Auscultation–Sound Software (CLASS) for learning pulmonary auscultation. Med Biol Eng Comput 1–11
12. Guntupalli KK, Alapat PM, Bandi VD, Kushnir I (2008) Validation of automatic wheeze detection in patients with obstructed airways and in healthy subjects. J Asthma 45 (10):903–907
13. Dinis J, Guilherme C, Rodrigues J, Marques A (2013) Respiratory sound annotation software. In: International conference on health informatics, pp 183–188
14. Pinho C, Oliveira A, Jácome C, Rodrigues J, Marques A (2015) Automatic crackle detection algorithm based on fractal dimension and box filtering. Procedia Comput Sci 64:705–12
15. Lartillot O, Toiviainen PA (2007) Matlab toolbox for musical feature extraction from audio. In: International conference on digital audio effects, pp 237–244
16. Mendes L, Vogiatzis IM, Perantoni E, Kaimakamis E, Chouvarda I, Maglaveras N, Henriques J, Carvalho P, Paiva RP (2016) Detection of crackle events using a multi-feature approach. In: 38th Annual International Conference of IEEE Engineering in Medicine Biology Soceity, pp 3679–83
17. Rocha BM, Mendes L, Couceiro R, Henriques J, Carvalho P, Paiva RP (2017) Detection of explosive cough events in audio recordings by internal sound analysis. In: 39th annual international conference of IEEE engineering in medicine biology, pp 2761–2764

# Hidden Markov Model Based Respiratory Sound Classification

N. Jakovljević and T. Lončar-Turukalo

**Abstract**

This paper presents a method based on hidden Markov models in combination with Gaussian mixture models for classification of respiratory sounds into normal, wheeze and crackle classes. Input features are mel-frequency cepstral coefficients extracted in the range between 50 Hz and 2000 Hz in combination with their first derivatives. The audio files are preprocessed to remove noise using spectral subtraction. Our best score achieved in the official ICHBI Challenge second evaluation phase is 39.56.

**Keywords**

Respiratory sounds • Crackles • Wheezes • Hidden Markov models • Spectral subtraction

## Introduction

Auscultation is a common, fast and noninvasive way to diagnose patients with lung diseases. Respiratory sounds according to their acoustic properties can be classified into normal and abnormal [1, 2]. Frequency content of normal respiratory sounds depends on stethoscope position and does not contain tonal (musical) components [2]. For example, lung or vesicular sounds are dominated by frequencies below 100 Hz, whereas in the tracheal sounds frequencies from 100 to 1500 Hz are more distinctive. Abnormal sounds consist of both normal and adventitious respiratory sound. Adventitious crackle sounds are discontinuous, nontonal lung sounds with a duration of less than 20 ms [2]. They are normally heard during inspiration and sometimes during expiration [2]. The crackle sounds' frequency range is 60–2000 Hz, with their major contribution below 1200 Hz [2]. Wheezes are continuous tonal lung sounds with the dominant frequency above 400 Hz, and with a duration longer than 100 ms [2].

The most comprehensive evaluation of different classification algorithms over healthy and asthmatic respiratory sound databases is presented in [3]. The best performance in [3] is obtained by the model based on Gaussian mixture models (GMM) in combination with mel-frequency cepstral coefficients (MFCCs). For these reasons this model has been selected as the baseline model. The functionality of this model has been enriched with the information about the frame position in a sequence, leading to hidden Markov model (HMM) instead of GMM. As hidden Markov models were the backbone in automatic speech recognition for many years [4], theoretical foundations have been developed, and many practical considerations are well defined. A respiration cycle varies in duration and acoustical content, just as in speech, which suggests that HMM is an appropriate tool to model it.

## Methods

### Preprocessing

The dataset contains audio recordings sampled at 44.1 kHz and 4 kHz. Even though a majority of the recordings is

N. Jakovljević (✉) · T. Lončar-Turukalo
Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, Novi Sad, Serbia
e-mail: jakovnik@uns.ac.rs



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_7

sampled at 44.1 kHz, downsampling to 4 kHz is performed as the frequency content of both wheeze and crackle is in the range of 60–2000 Hz [2]. An additional benefit is a significant reduction in computational complexity of feature extraction.

To remove sounds caused by heartbeats, the signal components at low frequencies have to be suppressed. We have evaluated the performance of two different filters. The first one is the low order bandpass filter with the transfer function:

$$H_1(z) = \frac{1 - z^{-2}}{1 - 0.9z^{-2}} \tag{1}$$

The additional benefit of this filter are the reduced effects of sudden changes in signal which can appear at the edges of clipped segments if only a high pass filter was applied.

The second filter is the high pass finite impulse response filter with cutoff frequency $f_c$ = 100 Hz and constant group delay $\tau_g$ = 1024 samples obtained by Hann window function. In this way components at frequencies below 96 Hz are attenuated by at least 54 dB, i.e. heartbeat sound is suppressed more than in the case of the first filter.

## Noise Suppression

Many sound files in the dataset contain stationary noise, thus the following step in this algorithm is noise suppression. The implemented noise suppression is based on spectral subtraction [5], which is performed on the signal which is segmented into 30 ms long frames shifted by 15 ms using Hann window function. For each frame discrete Fourier transform (DFT) is performed and each magnitude spectrum is decreased by the estimated noise magnitude spectrum, i.e.:

$$|X_d(k,t)| = |X(k,t)| - |D(k)| \tag{2}$$

where $|X(k,\ t)|$, $|D(k)|$ and $|X_d(k,\ t)|$ are the magnitude spectra of the original signal, the noise, and the denoised signal at time $t$ respectively, where $k$ denotes the frequency bin. The noise magnitude spectrum $|D(k)|$ is estimated as the mean value of $|X(k,\ t)|$ over 1% of the frames with minimum energy in the audio signal, excluding invalid frames with zero energy.

The problem of the negative values of $|X_d(k)|$ has been solved using two approaches. The first approach, referred to as SS1, sets the negative magnitude values to 1% of $|X(k,t)|$, i.e.:

$$|X_d(k,t)| = \begin{cases} |X(k,t)| - |D(k)| & |X(k,t)| > |D(k)| \\ 0.01 \cdot |X(k,t)| & \text{else} \end{cases} \tag{3}$$

The second approach, referred to as SS2, additionally reduces the musical noise level introduced by magnitude spectrum subtraction. The musical noise is caused by sudden drops of magnitude at a certain frequency bin in successive frames. Relying on the assumption that breath sound should be dominant in the signal, for each $k$ the estimated noise level $|D(k)|$ has been iteratively reduced by 10%, until in at least 60% of frames $|X(k,\ t)| > |D(k)|$ is fulfilled. The denoised magnitude spectrum is obtained by:

$$|X_d(k,t)| = \begin{cases} |X(k,t)| - |D(k)| & |X(k,t)| > |D(k)| \\ |X(k,t)|^2 & \text{else} \end{cases} \tag{4}$$

where instead of linear scaling of critical components, quadrature scaling is introduced, further suppressing small magnitudes in $|X_d(k,\ t)|$. It should be noted that $|X(k,\ t)|$ has to be range normalized to accommodate quadrature scaling.

To suppress sudden drops of magnitude, $|X(k,\ t)|$ is monitored in 5 successive frames. If $|X(k,\ t)| < |D(k)|$ in at least 3 of 5 adjacent frames, the frequency bin is marked as noise. An entire frame is considered as corrupted by noise and set to zero ($|X_d(k,t)| = 0$, for each $k$) if more than 70% of the bins are marked as noise.

In the synthesis step, the phase spectrum is approximated with the phase spectrum of the noisy signal, thus the spectrum of denoised signal is:

$$X_d(k,t) = |X_d(k,t)|e^{j\ \arg\{X(k,t)\}} \tag{5}$$

and the reconstructed signal is the sum of overlapping segments obtained by inverse DFT of $X_d(k,t)$.

## Feature Extraction

The MFCCs are estimated every 10 ms using 30 ms long windows. The frequency range [50, 2000 Hz] is divided into 16 equal-width overlapped channels in mel-frequency domain. The discrete cosine transform is performed on the logarithm of 16 energy coefficients calculated for each channel.

$$C_n = \sum_{k=1}^{16} \log(E(k)) \cos\left(\frac{n\pi}{16}\left(k - \frac{1}{2}\right)\right) \tag{6}$$

for $n$ = 0, 1, … 15, where $C_n$ is the $n$th MFCC and $E(k)$ is the energy at the $k$th channel. The coefficient $C_0$, which represents signal energy in the selected frequency band, is discarded from further steps, since in some signals it significantly correlates with heartbeat sound.

The cepstral mean and variance normalization per record is applied to remove variations caused by the remaining noise and it is defined by:

$$\bar{C}_n(t) = \frac{C_n(t) - \bar{C}_n}{S_n} \tag{7}$$

where:

$$\bar{C}_n = \frac{1}{T}\sum_{t=1}^{T} C_n(t) \tag{8}$$

$$S_n = \frac{1}{T}\sum_{t=1}^{T} (C_n(t) - \bar{C}_n)^2 \tag{9}$$

and $T$ is the duration of signal in frames.

Additionally, to track feature dynamics and to decorrelate successive feature vectors, first time derivatives of MFCCs are introduced, increasing the cardinality of the feature vector to $d = 30$.

## Modeling

By visual inspection we have found that the same sound class varies in acoustic content depending on recording location, thus a respiration cycle for each location (trachea, anterior left/right, lateral left/right, posterior left/right) and sound class (normal, crackle, wheeze, and both crackle and wheeze) is represented as a sequential HMM with $S$ states (see Fig. 1).

An HMM is described by its initial state probabilities ($\mathbf{\Pi}$), state transition matrix ($\mathbf{A}$), and emitting probability density function for each state ($b_s$). A state emitting probability density function (pdf) for a given $d$-dimensional observation $\mathbf{o}$ is defined by:

$$b_s(\mathbf{o}) = \sum_{i=1}^{M} w_i \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_i|^{1/2}} e^{-\frac{1}{2}(\mathbf{o}-\boldsymbol{\mu}_i)^T \mathbf{\Sigma}_i^{-1}(\mathbf{o}-\boldsymbol{\mu}_i)} \tag{10}$$

where $w_i$, $\boldsymbol{\mu}_i$ and $\mathbf{\Sigma}_i$ are weight, mean and covariance matrix of the $i$-th mixture component, respectively. Although each state can have a different number of mixture components, it is common that the number is the same for all states.

In case of sequential model only one state can be the first one, so in the vector $\mathbf{\Pi}$ only one value is equal to 1 and the others are 0, and each row in the state transition matrix $\mathbf{A}$ contains at most 2 nonzero elements.

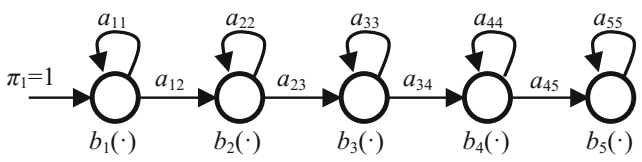



**Fig. 1** Sequential HMM with $S = 5$ states

The standard criterion for HMM parameter estimation is the maximization of the likelihood that the models will generate the training sequence [4]. The optimization is usually performed using expectation maximization algorithm (Baum-Welch estimation). For an efficient estimation procedure, the initial values of model parameters should be carefully set. In this study, the initial parameters were obtained by the time equidistant partition of the observation sequence between states, and for each state the sample mean $\boldsymbol{\mu}_s$ and the covariance matrix $\mathbf{\Sigma}_s$ were calculated. In case of several mixture components per state, means ($\boldsymbol{\mu}_i$) were obtained by random sampling from normal distribution $N(\boldsymbol{\mu}_s, \mathbf{\Sigma}_s)$, and covariance matrices ($\mathbf{\Sigma}_i$) by assigning the corresponding sample covariance matrix ($\mathbf{\Sigma}_i = \mathbf{\Sigma}_s$). The initial transition probabilities (Fig. 1) were set to 0.5, with stay probability corresponding to the last HMM state, except for $a_{SS}$, which was initialized to 1.

The existing model parameters are used to calculate probabilities that the model will be in the state $s$ at time $t$ and will generate the observation ($\mathbf{o}_t$) using the $m$-th mixture component. These probabilities are used to update the values of the transition probabilities, means and covariance matrices of the model. In our experiments these parameters converged in 6–12 iterations.

During the test phase, an unknown observation sequence, denoted $\mathbf{O} = [\mathbf{o}_1, \mathbf{o}_2, \ldots \mathbf{o}_T]$, is aligned with all HMMs ($\lambda_c$), and the classification decision is based on the maximum likelihood criterion, i.e.

$$\hat{c} = \arg \max_{1 \le c \le C} p(\mathbf{O}|\lambda_c) \tag{11}$$

$$p(\mathbf{O}|\lambda_c) = \pi_1 b_1(\mathbf{o}_1) \sum_{s(2),\ldots,s(T)} a_{1s(2)} \prod_{t=2}^{T} b_{s(t)}(\mathbf{o}_t) a_{s(t)s(t+1)} \tag{12}$$

where $s(t)$ represents the state at time $t$, and $C$ the number of classes. Having in mind computational complexity, the log probabilities are used instead of probabilities themselves.

## Database

For training and evaluation, the official ICBHI Challenge respiratory sound database released in September 2017 was used [6]. The details on data acquisition and ethical considerations are provided [6]. The number of attempts for the official scoring was limited, therefore many of experiments were evaluated only on a validation set. The official training set was divided into 10 folds. The validation set in each fold contains at least one sound class for every possible recording

**Table 1** Sensitivity (Se), specificity (Sp), and score evaluated on the validation set, and score on the official test for different preprocessing procedures (PP), the number of states (S), the number of mixture components per state (M), and covariance matrix type (CMT)

| PP | S | M | CMT | Validation set | | | Official |
|---|---|---|---|---|---|---|---|
| | | | | Se | Sp | Score | Score |
| T2 | 1 | 4 | full | 0.4381 | 0.4533 | 44.57 | n/a |
| T2 | 1 | 8 | full | 0.4252 | 0.5136 | 46.94 | n/a |
| T2 | 1 | 16 | full | 0.3517 | 0.6115 | 48.16 | n/a |
| T2 | 1 | 32 | full | 0.2089 | 0.7671 | 48.80 | n/a |
| T2 | 1 | 64 | full | 0.0917 | 0.8702 | 48.09 | n/a |
| T1 | 5 | 1 | full | 0.4093 | 0.5326 | 47.09 | 39.32 |
| T1 | 5 | 1 | diag. | 0.4079 | 0.4091 | 40.85 | 39.02 |
| T1 | 6 | 1 | full | 0.4232 | 0.5669 | 49.50 | 39.37 |
| T2 | 6 | 1 | full | 0.4102 | 0.5267 | 46.85 | 36.98 |
| Class. ensemble | | | | n/a | n/a | n/a | 39.56 |

location. All respiratory cycle instances from an audio file were in the same (train/validation) set.

## Evaluation Criterion

The performances of classifiers were evaluated using officially proposed scores [7] i.e. sensitivity (*Se*), specificity (*Sp*), and overall score, compactly written as:

$$Se = \frac{C_c + C_w + C_b}{T_c + T_w + T_b},\ Sp = \frac{C_n}{T_n},\ Score = \frac{Se + Sp}{2} 100\% \tag{13}$$

where $C_i$ and $T_i$ ($i = c,\ w,\ b$) are the number of correctly recognized instances of class $i$, and the total number of instances of class $i$ in the test (or validation) set, respectively. Indices $c$, $w$, $b$, and $n$ stand for classes: crackle, wheeze, both crackle and wheeze, and normal, respectively.

## Results and Discussion

The selected results are summarized in Table 1. The classifiers differ by the preprocessing procedure, the number of states and mixture components per state and the type of the covariance matrix. In the first preprocessing procedure (T1), proposed in the first phase of ICBHI Challenge, the input signal is filtered through the bandpass filter $H_1(z)$ and noise suppression is based on the SS1 method. The second preprocessing procedure (T2) includes downsampling to 4 kHz, filtering by the high pass FIR filter and noise suppression based on SS2. It should be noted that the features are extracted in the frequency range [50, 2000 Hz] independently of the preprocessing procedure. Our initial experiments for the simpler models on reduced dataset have shown that there is no significant difference between these preprocessing procedures, but a difference has been noted on the extended dataset (see last two rows in Table 1).

The baseline system based on GMM has shown slightly inferior performance to the HMM based systems. It can be noted that with the increasing number of mixture components the overall score is improving, as the result of higher specificity. However, sensitivity is decreasing, indicating that the classifier could not resolve adventitious sound types.

Introducing HMM, i.e. taking into consideration the position of the frame in a sequence, increases the accuracy of the model without a significant increase of its complexity.

As the used features are correlated, modeling data with full covariance matrix increases the overall score by increasing the specificity, without degradation in sensitivity (Table 1, rows 6 and 7). The difference of the scores obtained on the validation set (6.24) is higher than the difference of the official test set scores (0.30).

The overall discrepancies of scores obtained in cross-validation using the publicly available dataset, and the official test set (Table 1, columns 7 and 8) are noticeable. One plausible reason for the score discrepancies might be the correlation of the recordings in the publicly available dataset (recordings from the same subject might be present in both training and validation set), whereas the test set strictly comprises a disjunct set of subjects [7].

To increase the overall score, we have tried with an ensemble of classifiers trained over the 10 different folds. All classifiers which had the same model complexity (28 models with 5 states and 1 Gaussian per state) were trained with a single learning method. The final decision was made by simple majority voting by the classifiers. This approach has achieved our best official score of 39.56, that represents a minor increase in the score (0.24) at the expense of 10 times greater computational complexity.

The presented results are modest in comparison with the results published in [1, 3, 8], where both less extensive databases and a smaller number of the adventitious sound classes are used. There are several challenging issues regarding the database used in this study: different types of noise, multiple recording locations, and small numbers of samples for different classes.

## Conclusions

This study shows that MFCCs in combination with HMM can be used for classification or respiratory sounds into 4 categories: normal, crackle, wheeze, and both crackle and wheeze. The performances of the examined classifiers are modest because they were evaluated on real data under varying levels of different types of real noise. We assume that advanced noise suppression techniques can improve the overall score.

**Acknowledgements** This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, TR 32035 and TR 32040. We acknowledge the support of the COST Action ENJECT TD1405 in the form of ITC grant awarded to the first author.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Reichert S, Gass R, Brandt C, Andres E (2008) Analysis of respiratory sounds: state of the art. Clin Med Circ Respirat Pulm Med 2:45–58
2. Sarkar M, Madabhavi I, Niranjan N, Dogra M (2015) Auscultation of the respiratory system. Ann Thorac Med 10(3):158–168. https://doi.org/10.4103/1817-1737.160831
3. Bahoura M (2009) Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. Comp Bio Med 39(9):824–843
4. Gales M, Young S (2008) The application of hidden Markov models in speech recognition. Found. Trends Signal Process 1 (3):195–304. https://doi.org/10.1561/2000000004
5. Berouti M, Schwartz M, Makhoul J (1979) Enhancement of speech corrupted by acoustic noise. In: Proceedings of. IEEE international conference on acoustics, speech and signal processing, pp 208–211. https://doi.org/10.1109/ICASSP.1979.1170788
6. Rocha BM, Filos D, Mendes L et al (2017) A respiratory sound database for the development of automated classification. In: Proceedings of international conference on biomedical and health informatics. Thessaloniki, Greece (in press)
7. ICBHI Challenge. https://bhichallenge.med.auth.gr/rules
8. Kochetov K, Putin E, Azizov S, Skorbogatov I, Filchenkov A (2017) Wheeze detection using convolutional neural networks. In: Proceedings of EPIA conference on artifical intelligence. Porto, Portugal, pp 162–173. https://doi.org/10.1007/978-3-319-65340-2

# An Automated Lung Sound Preprocessing and Classification System Based OnSpectral Analysis Methods

Gorkem Serbes, Sezer Ulukaya, and Yasemin P. Kahya

**Abstract**

In this work, respiratory sounds are classified into four classes in the presence of various noises (talking, coughing, motion artefacts, heart and intestinal sounds) using support vector machine classifier with radial basis function kernel. The four classes can be listed as normal, wheeze, crackle and crackle plus wheeze. Crackle and wheeze adventitious sounds have opposite behavior in the time-frequency domain. In order to better represent and resolve the discriminative characteristics of adventitious sounds, non-linear novel spectral feature extraction algorithms are proposed to be employed in four class classification problem. The proposed algorithm, which has achieved 49.86% accuracy on a very challenging and rich dataset, is a promising tool to be used as preprocessor in lung disease decision support systems.

**Keywords**

Adventitious pulmonary sounds • Respiratory sounds • Wheeze • Crackle • Respiration cycle

## Introduction

Pulmonary diseases affect the comfort of the patients in their daily lives and some of the patients need to be tracked continuously for attacks and severe conditions. Low cost computerized pulmonary medical decision support systems may be employed to be used at home to monitor the status of the patient and this may reduce frequent hospital visits. However, in order to be used in clinical settings computerized systems must be validated in an objective way and should meet high accuracy requirements.

In literature, there are diverse approaches to explain pulmonary sound generation mechanisms [1, 2]. Although the generation mechanisms are varied, the indicators of dysfunctions are extensively studied and the most well known indicators are wheeze and crackle sounds in the computerized analysis area [3]. Wheezes are the oscillatory type of adventitious sounds whereas crackles are the transient type. These sounds represent opposite time-frequency (TF) characteristics and added onto normal (vesicular) breath sounds. The time and frequency content of normal, wheeze and crackle sounds are largely overlapped. Therefore, there is a need for an automatic system to provide fine TF resolution to resolve overlapped components of the pulmonary sounds. Moreover, an automatic system must deal with adventitious types of sounds without using prior information since some of them may be absent or consecutively present in the recorded pulmonary sound [1, 2].

G. Serbes
Biomedical Engineering Department, Yildiz Technical University, Istanbul, Turkey

S. Ulukaya (✉) · Y. P. Kahya
Electrical and Electronics Engineering Department, Boğaziçi University, Istanbul, Turkey
e-mail: sezer.ulukaya@gmail.com

S. Ulukaya
Electrical and Electronics Engineering Department, Trakya University, Edirne, Turkey



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_8

In order to be used in medical decision support systems, detection and classification of pulmonary adventitious sounds with high accuracy is crucial. In literature, studies are mainly focused on wheeze/non-wheeze [4–6], crackle/non-crackle classification [7–9] and crackle-wheeze-normal classification [10–12]. There are also six-class [13] and nine-class [14] classification studies in literature. Wheeze and crackles may overlap or located consecutively in a breath cycle; moreover, it is reported that asthma and chronic obstructive pulmonary disease (COPD) may overlap on 15% of the obstructive lung disease population [15]. In the presence of these cases, to the best of our knowledge, there is no classification study which deals with wheeze and crackle sounds together in a breath cycle. In this paper, a novel time-frequency analysis based pulmonary sound classification system is proposed to classify pulmonary sound recordings into normal, crackle, wheeze and crackle-wheeze groups. Crackles are transient waveforms which typically last 20 ms with a scattered frequency content [16]. On the other hand, wheezes are oscillatory waveforms which typically last more than 80 ms [17]. Crackles are associated with pulmonary diseases such as COPD, chronic bronchitis, pneumonia, fibrosing alveolitis and emphysema [16], whereas wheezes are associated with pulmonary diseases such as asthma and COPD [18].

In literature, there are works which can be categorized into either event based or segment based approaches [3]. Each approach has its own advantages. Segments are generally short duration (100 ms) time domain windows which include either wheeze or crackle sound types and their location is labelled by medical experts. On the other hand, events are long duration (10 s) time domain windows which include adventitious sounds but the exact location of the sound is not known or labelled. The latter is more difficult than the former since transient waveforms (crackles have 20 ms duration) may be located in any part of the long duration signal. Therefore, since crackles or wheezes may be present in the long duration window of the given lung sound, an automatic technique which has finer TF resolution is proposed to preprocess and classify crackle, wheeze, normal and crackle plus wheeze lung sound classes.

In section "Materials and Methods", properties of the dataset and proposed novel time-frequency analysis based method is introduced. In section "Results", experimental setup and results are represented. In section "Conclusions", outcomes are discussed and conclusions are summarized.

## Materials and Methods

### Description of Dataset

The system is trained on a publicly available lung sound database [19]. 4144 labelled respiratory cycles (2072 normal, 1209 crackle, 501 wheeze, 362 crackle plus wheeze) whose duration varies from 10 to 90s are used to form the dataset. Lung sounds are acquired using either microphones or digital stethoscopes from seven different locations on the chest wall. The dataset is very challenging and rich such that, it includes domestic and clinical recordings of children and adult subjects contaminated with various noise components such as coughing, talking, heart and intestinal sounds and motion artefacts [19]. The test set is hidden and has different subjects, so that subject independent testing and training of the algorithm is guaranteed. The locations of the adventitious lung sounds are unknown and their labels are assigned by experts. Moreover, the dataset is collected from patients with different pulmonary disorders.

### General Processing Steps of the Proposed System

In Fig. 1, the complete automatic preprocessing and classification system is represented step by step. As a preprocessing step, the given respiratory sound segment is band pass filtered to eliminate heart sounds and other noise components. Then the segment is separated into three channels to locate wheezes, crackles and background

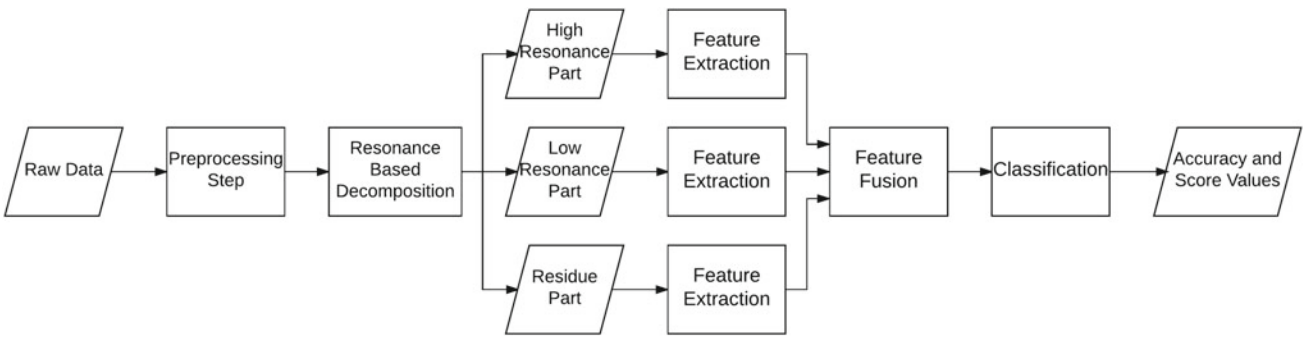



**Fig. 1** Flowchart of the complete proposed preprocessing and classification system

(vesicular) sounds using resonance based decomposition method [20]. Following that step, various time-frequency and time-scale analysis based features are extracted from each individual channel and a fusion of these features are fed into support vector machine classifier to train and test the proposed system. The details of each processing step of the system are presented below.

## Preprocessing Step

At this step, the acquired raw lung sound data are down-sampled to 4000 Hz to set up a coherent feature set. Even on normal lung sound in the dataset which includes various noises such as intestinal and heart sounds and motion artefacts, 12th order Butterworth band pass filter with 120 and 1800 Hz cut-off frequencies is applied to minimize noise effects.

## Proposed Time-Frequency and Time-Scale Analysis Based Method

Since the frequency ranges of lung sound types are highly overlapped and location of the adventitious sounds are unknown, a resonance based decomposition method [21], which is able to represent transient (crackle), oscillatory (wheeze) waveforms and noise components at separate channels, is applied to preprocessed lung sound segments. Adventitious lung sounds may contain both low and high frequency components, therefore instead of using linear frequency based decomposition, non-linear resonance based wavelet decomposition is proposed to decompose given lung sound into three channels. Oscillatory waveforms can be represented with high Q-factor wavelet, while transient waveforms can be better represented using low Q-factor wavelet bases. Once the given lung sound data are decomposed into high resonance (wheeze), low resonance (crackle) and residual (noise) parts, short time Fourier transform (STFT) [22] is applied to each decomposed channel. In the beginning, after TF representation is formed, frequency content of each channel is computed by integrating the TF distribution over time. The motivation behind this step is to capture the discriminative frequency characteristics of lung sound types to better represent different classes because exact location of the adventitious sound is unknown and consequently, time information can not be used. Moreover, to better represent the discriminative characteristics of the dataset, statistical and spectral features derived from the wavelet coefficients are also employed to be used as additional features. Mean, skewness, kurtosis, standard deviation, minimum, maximum of the decomposed wavelet coefficients, linear energy and nonlinear Teager-Kaiser energy of the wavelet coefficients and entropy of wavelet coefficients are employed to increase the performance of the proposed system.

## Feature Fusion and Classification

After the frequency features are extracted using STFT and tunable wavelet transform [20], features are scaled to $[-1, +1]$ range to be normalized. In order to be employed in an online classification system, the dimension of the features is reduced using principal component analysis (PCA) method [23]. Various values of variance (90, 95 and 99%) are employed to measure the success of the proposed system. Support vector machine [24] with radial basis function kernel is employed by exploring the best fit $C(100)$ and $\gamma(0.001)$ parameters after various trials. Extracted STFT or STFT + Wavelet features are fed into the classifier algorithm. STFT is employed using window length of 256

**Table 1** Various performance results of merely STFT and STFT + Wavelet based extracted features of proposed system employing different PCA variances

| Method | Rate | PCA Variances | | |
|---|---|---|---|---|
| | | 0.90 | 0.95 | 0.99 |
| STFT | **Accuracy** | 49.98 | 53.68 | 57.75 |
| | **Sensitivity** | 48.90 | 52.11 | 54.33 |
| | **Specificity** | 77.80 | 82.40 | 87.50 |
| | **Score** | 63.35 | 67.25 | 70.92 |
| STFT + Wavelet | **Accuracy** | 57.88 | 55.29 | 54.15 |
| | **Sensitivity** | 55.29 | 52.60 | 50.61 |
| | **Specificity** | 83.25 | 83.65 | 84.40 |
| | **Score** | 69.27 | 68.13 | 67.51 |

with 80% overlap on Hamming window. For the wavelet part, $Q$ (Q-factor), $r$ (over-sampling rate) and $J$ (number of levels) values of 4, 5 and 45 are employed, respectively.

## Results

In the Table 1, the accuracy, specificity, sensitivity and score (exactly average of specificity and sensitivity) rates of merely STFT and STFT + Wavelet based proposed method is represented employing various PCA variances. As represented in the table, accuracy of the STFT + Wavelet based method is better than STFT even if only 90% of the PCA variance is employed. Moreover, when 90% of the PCA variance is employed the score of the STFT is 63.35%, while the score of the STFT + Wavelet based method is 69.27%. In order to be employed in real time systems, the dimension of the extracted features are reduced using PCA to decrease the computational load. STFT + Wavelet based algorithm represents decreasing performance while the recovered variance of the PCA is increased (except for the specificity rate) as presented in Table 1. This is probably due to the reason that classifier is confused with the redundant information. These results are obtained on the training dataset, since the testing dataset is hidden, the results on the testing dataset could not be provided in detail.

The experimental results showed that proposed wavelet and Fourier based spectral features are not generalizable to the testing set. The reasons for this outcome may be that the kernel of the support vector machine could not be able to learn the bases, Fourier based spectral features could not be able to localize transient sounds (duration of crackles is 10 ms) in 10–90s recordings or normal respiratory sounds are confused with crackle segments due to various intestinal and other noise sources. In the testing experiments 39.97–49.86% accuracies are reached. It is seen that, the proposed STFT + Wavelet based features represent 1.20% better performance than proposed merely STFT based features on the testing dataset.

## Conclusions

A robust and generalizable method is needed to classify respiratory sounds in the presence of various noises. Lower testing scores pointed out that either classifier or trained model show weaker generalizable representation. Redundancy in the extracted features may also decrease the classification performance. Detecting and classifying transient waveforms which last 20 ms in 10 s duration lung segment is very challenging in the presence of heart, intestinal sounds and stethoscope motion. Moreover, normal lung sound segments which include various noise sources may be confused with lung sound segments containing crackle sounds. More advanced spectral features and classifiers will be explored to increase the classification accuracy of the proposed system.

**Acknowledgements** This work is supported by Boğaziçi University Research Fund under grant number 16A02D2. S. Ulukaya is supported by the Ph.D. scholarship (2211) from Turkish Scientific Technological Research Council (TUBITAK).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Bohadana A, Izbicki G, Kraman SS (2014) Fundam Lung Auscult New Engl J Med 370:744–751
2. Gavriely N, Cugell DW (1995) Breath sounds methodology. CRC Press (1995)
3. Pramono RXA, Bowyer S, Rodriguez-Villegas E (2017) Automatic adventitious respiratory sound analysis: a systematic review. PloS One 12:e0177926
4. Bahoura M (2009) Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. Comput Biol Med 39:824–843
5. Emrani S, Gentimis T, Krim H (2014) Persistent homology of delay embeddings and its application to wheeze detection. IEEE Signal Process Lett 21:459–463
6. Wis′niewski M, Zielin′ski TP (2015) Joint application of audio spectral envelope and tonality index in an e-asthma monitoring system. IEEE J Biomed Health Inf 19:1009–1018
7. Serbes G, Sakar CO, Kahya YP, Aydin N (2013) Pulmonary crackle detection using time–frequency and time–scale analysis. Digit Signal Process 23:1012–1021
8. Mendes L, Vogiatzis I M, Perantoni E et al (2016) Detection of crackle events using a multi-feature approach. In: 2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC). IEEE, pp 3679–3683
9. Grønnesby M, Solis JCA, Holsbø E, Melbye H, Bongo LA (2017) Machine learning based crackle detection. In: Lung sounds. arXiv: 1706.00005
10. Naves R, Barbosa BHG, Ferreira DD (2016) Classification of lung sounds using higher-order statistics: a divide-and-conquer approach. Comput Methods Programs Biomed 129:12–20
11. Sengupta N, Sahidullah M, Saha G (2016) Lung sound classification using cepstral-based statistical features. Comput Biol Med 75:118–129
12. Ulukaya S, Serbes G, Kahya YP (2017) Overcomplete discrete wavelet transform based respiratory sound discrimination with feature and decision level fusion. Biomed Signal Process Control 38:322–336
13. Kandaswamy A, Kumar CS, Ramanathan RP, Jayaraman S, Malmurugan N (2004) Neural classification of lung sounds using wavelet coefficients. Comput Biol Med 34:523–537
14. Dokur Z (2009) Respiratory sound classification by using an incremental supervised neural network. Pattern Anal Appl 12:309
15. Postma DS, Rabe KF (2015) The asthma–COPD overlap syndrome. N Engl J Med 373:1241–1249
16. Piirila P, Sovijarvi ARA (1995) Crackles: recording, analysis and clinical significance. Eur Respir J 8:2139–2148
17. Pasterkamp H, Kraman SS, Wodicka GR (1997) Respiratory sounds: advances beyond the stethoscope. Am J Respir Crit Care Med 156:974–987

18. Meslier N, Charbonneau G, Racineux JL (1995) Wheezes. Eur Respir J 8:1942–1948
19. International Conference on Biomedical and Health Informatics (ICBHI) Challenge Database (2017). https://bhichallenge.med.auth.gr/. Accessed 01 Sept 2017
20. Selesnick IW (2011) Wavelet transform with tunable Q-factor. IEEE Trans Signal Process 59:3560–3575
21. Selesnick IW (2011) Sparse signal representations using the tunable Q-factor wavelet transform. In: SPIE Optical Engineering + Applications: 81381U1–81381U15 International Society for Optics and Photonics
22. Cohen L (1995) Time-frequency analysis, vol 778. Prentice Hall PTR, Englewood Cliffs, NJ
23. Jolliffe IT (1986) Principal component analysis and factor analysis. In: Principal component analysis. Springer, pp 115–128
24. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2:1–27

# Detection of Cough and Adventitious Respiratory Sounds in Audio Recordings by Internal Sound Analysis

B. M. Rocha, L. Mendes, I. Chouvarda, P. Carvalho, and R. P. Paiva

**Abstract**

We present a multi-feature approach to the detection of cough and adventitious respiratory sounds. After the removal of near-silent segments, a vector of event boundaries is obtained and a proposed set of 126 features is extracted for each event. Evaluation was performed on a data set comprised of internal audio recordings from 18 patients. The best performance (F-measure = 0.69 ± 0.03; specificity = 0.90 ± 0.01) was achieved when merging wheezes and crackles into a single class of adventitious respiratory sounds.

**Keywords**

Cough • Adventitious respiratory sounds • Automatic classification

## Introduction

Respiratory diseases cause an immense socio-economic impact and are the third leading cause of death worldwide [1] and a burden to public health systems [2]. Therefore, significant research efforts have been dedicated to improving early diagnosis and routine monitoring of patients with respiratory diseases to allow for timely interventions [3].

A great amount of research has been focused on the auscultation and characteristics of cough and respiratory sounds (RS), as they are valuable indicators of respiratory health and respiratory disorders [4].

Cough is a natural respiratory defense mechanism to protect the respiratory tract and one of the most common symptoms of pulmonary disease [5]. It can be characterized by an initial contraction of the expiratory muscles against a closed glottis, followed by a violent expiration as the glottis opens suddenly, producing a characteristic sound [6]. The cough sound is usually divided in three phases: an explosive phase, an intermediate period, whose characteristics are similar to a forced expiration, and a voiced phase. Cough often occurs as an epoch, where an initial inspiration is followed by a series of glottal closures and expiratory efforts, sometimes with interspersed inspirations [7]. In this paper, we consider each cough event, i.e., each glottal closure and expiratory effort, independently.

Respiratory sounds are generally classified as normal or adventitious. Auscultation-based diagnosis and monitoring of respiratory conditions rely heavily on the presence of adventitious sounds and on the altered transmission characteristics of the chest wall. Adventitious sounds are RS superimposed on normal respiratory sounds which can be discontinuous (crackles) or continuous (wheezes). Crackles are discontinuous, explosive, and non-musical adventitious RS that occur frequently in cardiorespiratory diseases [8]. They are usually classified as fine and coarse crackles based on their duration, loudness, pitch, timing in the respiratory cycle, and relationship to coughing and changing body position [9]. Wheezes are musical RS that usually last more than 250 ms. They are a common clinical sign in patients with obstructive airway diseases such as asthma and chronic obstructive pulmonary disease (COPD) [10].

B.M. Rocha (✉) · L. Mendes · P. Carvalho · R.P. Paiva
Centre for Informatics and Systems, University of Coimbra, DEI, Polo 2, Pinhal de Marrocos, Coimbra, Portugal
e-mail: bmrocha@dei.uc.pt

I. Chouvarda
Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_9

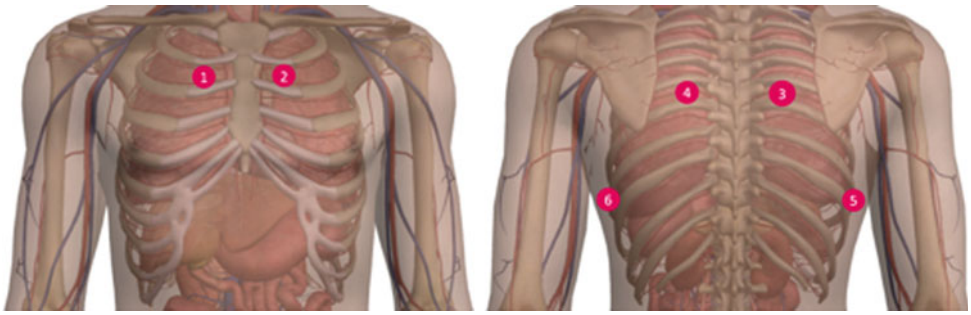



**Fig. 1** Chest locations for the recording of respiratory sounds

The main goal of this work was to design a method for the automatic detection of cough and adventitious RS solely from audio recordings. The automatic detection of cough and adventitious RS has been the subject of many studies in the last decades. Algorithms developed to detect or classify events usually involve two steps; cough and adventitious RS are no exception. The first step is to extract the relevant features that will be used as detection or classification variables. The second step is to use detection or classification techniques on the data, based on the features extracted. The most common features employed in the literature include Mel-frequency cepstral coefficients (MFCCs), spectral features, energy, entropy, and wavelet coefficients. Machine learning algorithms proposed in the literature include empirical rule-based methods, support vector machines (SVMs), artificial neural networks (ANNs), Gaussian mixture models (GMMs), k-nearest neighbors (k-NNs), and logistic regression models [11]. Prior attempts at automated classification of adventitious RS have tried to simplify the problem by focusing on a single type of sound and, to the best of our knowledge, none has tried to classify cough and adventitious RS at the same time.

In Sect. 2 we describe the data collected for this work and the methodology proposed, including the features and classification algorithms used. In Sect. 3 we present the results and discuss their implications. Finally, conclusions of the work are provided in Sect. 4.

## Materials and Methods

### Data Collection

Respiratory sounds were acquired at the Papanikolaou General Hospital, Thessaloniki and at the General Hospital of Imathia (Health Unit of Naousa), Greece. Sounds were collected sequentially from six chest locations, as shown in Fig. 1. The acquisition of RS was performed on adult and elderly patients. All patients had COPD with comorbidities (e.g. heart failure, diabetes, hypertension). Table 1 provides a description of the data set.

These recordings were acquired as part of the European project WELCOME (Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Comorbidities) project and were annotated using Audacity[1] 2.0.6 a free, open source, cross-platform software for recording and editing sounds.

Respiratory sound annotations were performed by three experienced physicians, two specialized pulmonologists and one cardiologist. Annotations discriminated the following sounds: normal (respiratory sound), crackles, wheezes, speech, cough, artifact.

### Pre-processing

In the pre-processing stage, the audio signal is filtered, using an 8th-order infinite impulse response (IIR) high-pass filter at 80 Hz (below the lower bound of the typical adult human voice [12]), and normalized. We then proceed to discard near-silent segments through the following process: given a threshold for length (100 ms) and another for amplitude (5%), segments whose length and amplitude are both below their respective thresholds are classified as near-silent and discarded, i.e., a segment is near-silent if its number of consecutive samples with absolute amplitude below 5% adds up to more than 100 ms. Subsequently, we compute the rms energy in each remaining segment, in 10 ms frames with 80% overlap, to find the onset (threshold: 20%) and ending (threshold: 5%) of each event. These parameters were experimentally obtained and sensitivity analysis proved their robustness. Finally, a vector of event boundaries is fed to the feature extractor.

### Feature Extraction

A total of 42 descriptors were computed in frame windows of 50 ms and 80% overlap. The final number of features computed for each event was 128, corresponding to the

[1] http://audacity.sourceforge.net/.

**Table 1** Description of data set

| | |
|---|---|
| # Patients | 18 |
| Average signal duration (s) | 106 |
| # Cough segments | 574 |
| # Crackle segments | 184 |
| # Wheeze segments | 248 |
| # Speech segments | 440 |
| # Other segments | 602 |

median, the maximum, and the standard deviation of each descriptor.

## Musical features

The MIR Toolbox [13] was used to extract 35 features related to dynamics, timbre, pitch, and harmonic content. Table 2 provides a brief description of the musical features used in this work.

## Other features

Seven other features were extracted in this work. Chirp group delay is a phase-based measure proposed in [14] for highlighting turbulences during glottal production. Harmonic to noise ratio (HNR) was computed for the frequency ranges (0–500) and (0–1500) Hz using the Voice Sauce toolkit [15]. The information entropy is a measure of the disorder of a system; the maximum of the entropy in each

**Table 2** Description of the musical features

| Feature | Description |
|---|---|
| RMS | Root-mean square energy of the frame |
| Spectral centroid | Center of mass of the spectral distribution |
| Spectral brightness | Amount of energy above 1500 Hz |
| Spectral spread | Variance of the spectral distribution |
| Spectral skewness | Skewness of the spectral distribution |
| Spectral kurtosis | Excess kurtosis of the spectral distribution |
| Spectral rolloff 95 | Frequency such that 95% of the total energy is contained below that frequency |
| Spectral rolloff 85 | Frequency such that 85% of the total energy is contained below that frequency |
| Spectral entropy | Complexity of the spectrum |
| Spectral flatness | Noisiness of the spectrum |
| Spectral roughness | Estimation of the sensory dissonance |
| Spectral irregularity | Degree of variation of the successive peaks of the spectrum |
| MFCC | 13 Mel-frequency cepstral coefficients |
| Zero-crossing rate | Waveform sign-change rate |
| Spectral Flux | Distance between the spectrum of successive frames |
| Chromagram centroid | Tonal centroid |
| Chromagram peak | Peak of the tonal centroid |
| Key clarity | Probability of key candidates |
| Mode | Modality estimation |
| Harmonic change detection function | Flux of the tonal centroid |
| Pitch | Pitch estimation |
| Pitch inharmonicity | Ratio of partials that are not multiple of the fundamental frequency, taking into account the amount of energy outside the ideal harmonic series |
| F0 | Fundamental frequency estimation |

**Table 3** Results

| Data set | Specificity | Sensitivity | Precision | F-measure |
|---|---|---|---|---|
| Complete | 0.90 ± 0.01 | 0.67 ± 0.03 | 0.67 ± 0.03 | 0.67 ± 0.03 |
| Merged | 0.90 ± 0.01 | 0.69 ± 0.03 | 0.70 ± 0.03 | 0.69 ± 0.03 |
| Complete 50 | 0.90 ± 0.01 | 0.67 ± 0.03 | 0.67 ± 0.04 | 0.67 ± 0.03 |
| Merged 50 | 0.90 ± 0.01 | 0.69 ± 0.03 | 0.70 ± 0.03 | 0.69 ± 0.03 |
| Complete 20 | 0.90 ± 0.01 | 0.65 ± 0.04 | 0.65 ± 0.04 | 0.65 ± 0.04 |
| Merged 20 | 0.89 ± 0.01 | 0.67 ± 0.03 | 0.68 ± 0.03 | 0.67 ± 0.03 |
| Complete 10 | 0.90 ± 0.01 | 0.62 ± 0.03 | 0.63 ± 0.03 | 0.62 ± 0.03 |
| Merged 10 | 0.88 ± 0.01 | 0.65 ± 0.03 | 0.65 ± 0.03 | 0.65 ± 0.03 |
| Complete 5 | 0.88 ± 0.01 | 0.57 ± 0.03 | 0.58 ± 0.03 | 0.57 ± 0.03 |
| Merged 5 | 0.87 ± 0.01 | 0.61 ± 0.03 | 0.62 ± 0.03 | 0.61 ± 0.03 |

Mean ± standard deviation

frame was used as a feature. Another computed feature was the maximum of the Teager energy in each frame. The maximum of Katz's fractal dimension of the filter WPST-NST, described in [16], was also calculated. Finally, the wheeze signature in the spectrogram space, thoroughly described in [17], was computed.

## Classification

Before classifying the events, the data set is partitioned into 10 stratified folds. Then, the training folds are filtered through the following procedure: (1) a class balancer is applied to reweight the instances in the data so that each class has the same total weight; (2) feature selection is performed and each feature is evaluated according to the information gain it provides; (3) each instance of the training set is classified and misclassified instances are removed. Finally, a random forest classifies each event of the test fold. This algorithm was chosen after validation and comparison with other common machine learning algorithms on a subset of the data. This process is repeated 10 times.

## Evaluation

Six versions of the data set were used for evaluation: *Complete,* with five classes (cough, wheezes, crackles, speech, other) and no feature selection; *Merged*, with four classes, where wheezes and crackles were merged (cough, adventitious sounds, speech, other) and no feature selection was performed; *Complete 50*, i.e., *Complete* with the best 50 features; *Merged 50*, i.e., *Merged* with the best 50 features; *Complete 20*; *Merged 20*; *Complete 10*; *Merged 10*; *Complete 5*; *Merged 5*. Table 3 shows the sensitivity, specificity, precision, and F-measure for all sets.

Regarding specificity, the performance did not change significantly with feature selection or the merging of the adventitious RS classes. Regarding the other metrics, the performance in *Merged* sets is always better than in *Complete* sets and it is especially so in the adventitious RS classes. The removal of features also seems to have less impact in *Merged* sets. Given these results, one can speculate that performing a hierarchical classification where adventitious RS are first merged and then discriminated might improve the performance.

## Conclusions

This paper presents a method for the detection of cough and adventitious RS. A data set comprising a total of 18 patients was used to evaluate the performance. The results indicate that future work should employ a hierarchical classifier.

**Acknowledgements** The authors would like to thank the health professionals and the patients who have agreed to participate in the data collection process. This work was financially supported by the EU project WELCOME[2] (FP7-ICT-2013-10/611223)

**Conflict of Interest** The authors declare that they have no conflict of interest.

## REFERENCES

1. World Health Organization (2015) The top 10 causes of death
2. Gibson GJ, Loddenkemper R, Lundbäck B, Sibille Y (2013) Respiratory health and disease in Europe: the new European lung white book. Eur Respir J 42:559–563

[2] http://www.welcome-project.eu/.

3. Marques A, Oliveira A, Jácome C (2014) Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review. Respir Care 59(5):765–776
4. Earis J, Cheetham B (2000) Current methods used for computerized respiratory sound analysis. Eur Respir Rev 10(77):586–590
5. Korpáš J, Tomori Z (1979) Cough and other respiratory reflexes. S. Karger 81–104
6. Evans JN, Jaeger MJ (1975) Mechanical aspects of coughing. Lung 152(4):253–257
7. Fontana G, Widdicombe J (2007) What is cough and what should be measured? Pulm Pharmacol Ther 20(4):307–312
8. Piirila P, Sovijarvi AR (1995) Crackles: recording, analysis and clinical significance. Eur Respir J 8(12):2139–2148
9. Sarkar M, Madabhavi I, Niranjan N, Dogra M (2015) Auscultation of the respiratory system. Ann Thorac Med 10(3):158
10. Sovijarvi ARA, Malmberg LP, Charbonneau G, Vanderschoot J, Dalmasso F, Sacco C, Rossi M, Earis JE (2000) Characteristics of breath sounds and adventitious respiratory sounds. Eur Respir Rev 10:591–596
11. Pramono RXA, Bowyer S, Rodriguez-Villegas E (2017) Automatic adventitious respiratory sound analysis: A systematic review. PLoS ONE 12(5):e0177926
12. Baken, RJ, Orlikoff, RF (2000) Clinical measurement of speech and voice. Cengage Learning
13. Lartillot O, Toiviainen PA (2007) Matlab toolbox for musical feature extraction from audio. In: international conference on digital audio effects, pp 237–244
14. Drugman T, Dubuisson T, Dutoit T (2011) Phase-based information for voice pathology detection. Int Conf on Acoustics, Speech and Signal Processing
15. Shue YL, Keating P, Vicenik C, Yu K (2011) Voicesauce: a program for voice analysis. In: ICPhS, vol XVII, pp 1846–1849
16. Mendes L, Vogiatzis IM, Perantoni E, Kaimakamis E, Chouvarda I, Maglaveras N, Henriques, J, Carvalho, P, Paiva, RP (2016) Detection of crackle events using a multi-feature approach. In: 38th annual international conference of the IEEE engineering in medicine and biology society pp 3679–83
17. Mendes L, Vogiatzis IM, Perantoni E, Kaimakamis E, Chouvarda I, Maglaveras N, Tsara, V, Teixeira, C, Carvalho, P, Henriques, J, Paiva, RP (2015) Detection of wheezes using their signature in the spectrogram space and musical features. In: 38th annual international conference of the IEEE engineering in medicine and biology society

# Part III
# eHealth Systems, Services and Cloud Computing

# A Content-Aware Analytics Framework for Open Health Data

L. Koumakis, H. Kondylakis, D. G. Katehakis, G. Iatraki, P. Argyropaidas, M. Hatzimina, and K. Marias

**Abstract**

The vision of personalized medicine has led to an unprecedented demand for acquiring, managing and exploiting health related information, which in turn has led to the development of many e-Health systems and applications. However, despite this increasing trend only a limited set of information is currently being exploited for analysis and this has become a major obstacle towards the advancement of personalized medicine. To this direction, this paper presents the design and implementation of a content aware health data-analytics framework. The framework enables first the seamless integration of the available data and their efficient management through big data management systems and staging environments. Then the integrated information is further anonymized at run-time and accessed by the data analysis algorithms in order to provide appropriate statistical information, feature selection correlation and clustering analysis.

**Keywords**

Data analysis • Data mining • Heath data integration • IHE profiles • Semantic interoperability

## Introduction

Healthcare is challenged by large amounts of data that is diverse, unstructured and growing exponentially. The heterogeneity and scale of such data (clinical, environmental, lifestyle, etc.) raises the demand for seamless data access along with the availability of powerful, reliable and efficient data analysis operations, tools and services. Obviously, the amount of information available, the heterogeneity of it and the wide range of terminologies/ontologies available to model this information, dictate the identification of a solution able to handle all this data.

Health data analytics provide mechanisms able to identify patterns or trends in data, screen pre-frailty states and provide different views of data for new management plans. Data mining consists of various methods and algorithms, which have been applied to many research areas, and the healthcare domain is not an exception [1, 2]. Understanding and extracting knowledge from healthcare data, formed the need for advanced analytical methodologies that can effectively transform data into meaningful and actionable information [3].

The major challenge in healthcare analytics is not the data mining algorithms, per se, but rather the framework which leverages legions of disparate, structured, and unstructured data [4]. Analytics can provide insides and draw conclusions for the data only if the data source(s) have been appropriately integrated and populated by reliable content.

To this end, we propose an analytics framework over a well-defined mutli-layer approach, which provides efficient management of big data, seamless integration using semantics and standards, secure interaction and anonymization of data for public open access and a modular

L. Koumakis (✉) · H. Kondylakis · D. G. Katehakis · G. Iatraki · P. Argyropaidas · M. Hatzimina · K. Marias
Institute of Computer Science, Foundation for Research and Technology (FORTH), N. Plastira 100, Heraklion, Greece
e-mail: koumakis@ics.forth.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_10

analytics framework able to incorporate advanced algorithm. The bottom layer can retrieve effectively data from disparate sources and combine those utilizing IHE Profiles [5], the second layer is responsible for the semantic integration and the efficient management of heterogeneous big data providing also data access, and the third layer provides a user-friendly analytics portal. The proposed framework has been adopted by the iManageCancer European project [6] as an open access tool to any researcher for analysing anonymized, health related data. The framework was empirically evaluated by experts using artificial but realistic data that exist in real medical databases such as patients' admission details, demographics, laboratory exams, medications, wellbeing data from smart phones and smart watches, etc. The technology readiness level of the framework is currently at TRL6 and by the end of the iManageCancer project we plan to pass at TRL7.

This paper describes a framework for data analytics over health related big data sources. It presents the system's architecture in section "System Architecture" by means of describing (i) the heterogeneous health sources integration by means of the IHE technical framework, (ii) the staging big data environment and the semantic layer, both feeding (iii) the Data Analysis Layer. Section "Preliminary Evaluation and Discussion" provides preliminary results from the iManageCancer project and section "Conclusions".

## System Architecture

Analytical services try to go much further than traditional statistics by examining the raw data and then attempting to hypothesize relationships within the data. As shown in Fig. 1 on the top, data analysis and data mining is an iterative approach, which combines data from the semantic layer and the big data staging environment, pre-processes the data, performs the analysis and provides the results for visualization based on the data distillation model. The loop closes with the interaction of the end user who can refine the results and continue with a drill-down analysis to extract knowledge from patient cohorts with specific criteria.

The proposed architecture consists of three layers. The data analytics layer, the semantic layer and the heterogeneous health sources integration layer. In the next sections, we analyze in detail each one of the aforementioned layers.

## Heterogeneous Health Sources Integration Using IHE

IHE profiles have the potential to support the sharing of health information in a secure, reliable and incremental manner across the different points of care, through authorized and validated interaction with existing systems and tools. This requires the existence of a commonly agreed interoperability framework to be in place, including (amongst other) detailed conformance statements for each domain under consideration. Implementations of software must be in accordance with the specifications described by IHE Profiles [5], such as those for Cross Document Sharing (XDS), Patient Identifier Cross Referencing (PIXv3), Patient Demographics Query (PDQv3), Cross Community Access (XCA), and Cross Community Patient Discovery (XCPD) for peer-to-peer querying and retrieve with other communities.

Cross-organization health data sharing translates into complex security policies that need to be uniformly man-aged and enforced. New complex requirements include for example the capability of dealing with data-binding concepts such as 'purpose of use' and 'conditions on use' [7].

The IHE Quality, Research and Public Health (QRPH) domain [8] addresses the information exchange and electronic health record content standards that are necessary for the sharing of information relevant to quality improvement in patient care, clinical research and public health monitoring. IHE QRPH addresses the infrastructure and content necessary to share information relevant to quality improvement, improve the liaison between the primary care system and clinical research and provide population base health surveillance, which are all reliant on the secondary use of data gathered in clinical care. Some examples of relevant IHE QRPH profiles indicatively include the Clinical Research Process Content [9] and the Research Matching [10].

Having in our disposal the IHE profiles, we can rely on a health information ecosystem, with well-defined interoperability standards for every health related sources such as Electronic Health Records, Personal Health Records, lifestyle monitoring, laboratory results, etc.

## Semantic Layer and Application Programming Interface

The semantic layer consists of two main sub-layers, i.e. the big data lake and the integrated information layer. On top of these layers APIs use various data services to provide access to the available information exploiting anonymization services according to the security/ethical requirements. Bellow we explain in detail each one of those components.

*Data Lake Layer*

Presented architecture's big data lake layer corresponds to IHE ingestion layer. In the data lake, various information exists in its proprietary form and is further processed and cleaned as a first step of data management. Various databases are staged in this layer such as PostgreSQL storing the

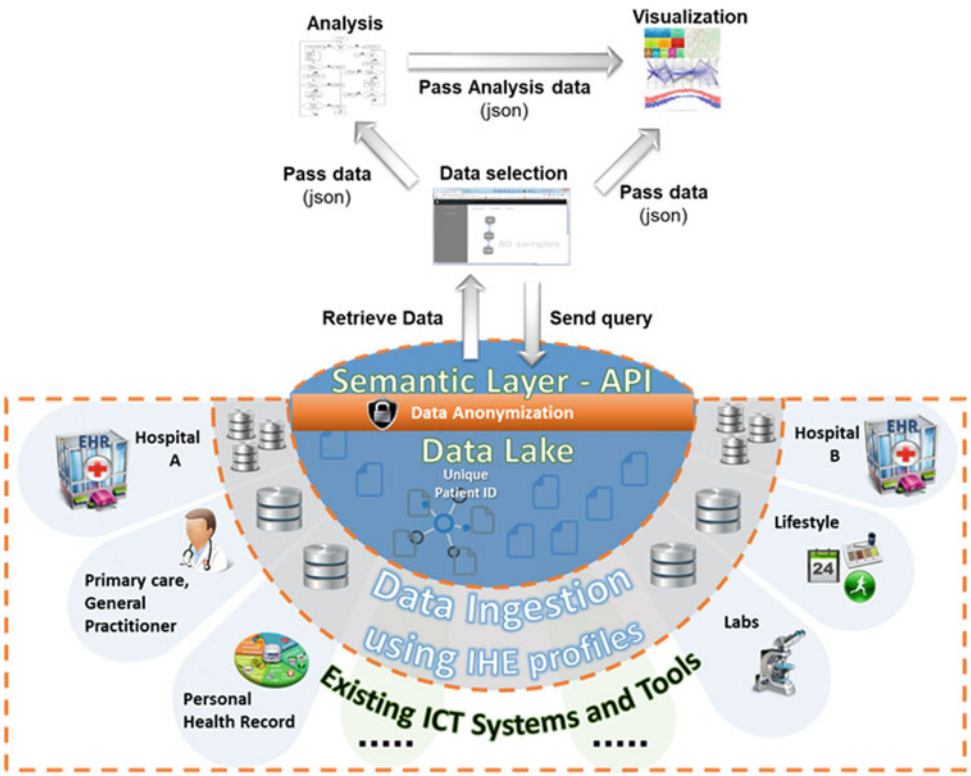



**Fig. 1** The reference architecture of the proposed analytics framework

PHR patient information, MySQL for storing the recommendations to the patients and the corpus to be recommended, Cassandra DBs for staging big data available (e.g. activity monitoring data, sensor data etc.) and other DBs for storing the pushed information from external sources.

*Integrated Information Layer*

Selected information out of the data lake layer is mapped to a modular ontology, the IMC Semantic Core Ontology, developed as part of the iManageCancer project [11]. Then, data are semantically uplifted (through an ETL process), and stored in a Virtuoso Triple Store. A benefit of the approach is that we can recreate from scratch the resulting triples at any time. However, for reasons of efficiency the data integration engine periodically transforms only the newly inserted information by checking the data timestamps. In this process summarization tools [12] allow the quick exploration and understanding of the available information enabling subsequent query formulation.

*Data Access Services (APIs)*

Both the integrated information and the staged information at the data lake can be queried using appropriate Data Access Services through the appropriate APIs. The APIs transform the user request to the appropriate query language (CQL, SQL or SPARQL) and forward the query to the appropriate source. The analytics framework usually queries the integrated information since we would like to analyze the linked information between the sources.

## Data Analysis Layer

The objective of the analytical framework is to extract information from the diverse health data sources and transform it into an understandable structure for better knowledge and further use. The platform is modular enough and allow any data-mining algorithm to be directly embedded in the whole workflow. As we can see from the high-level architecture (Fig. 1 on the top), the components of the analytical framework are the query builder, analysis and visualization.

*Query Builder*

Query builder is the place where the end user poses the research question and pull the anonymized data from the data lake. Since the end users can be also non-IT experts the graphical interface intended to be simple but yet powerful enough for complex queries. The user actually draws an SQL-like query with his/her preconditions and selects the attributes that would like to retrieve data from. The implementation is based on graphs and the user has to create/draw a graph where each node is a feature with specific conditions (e.g. the user wants to view data for all the patients with *age*

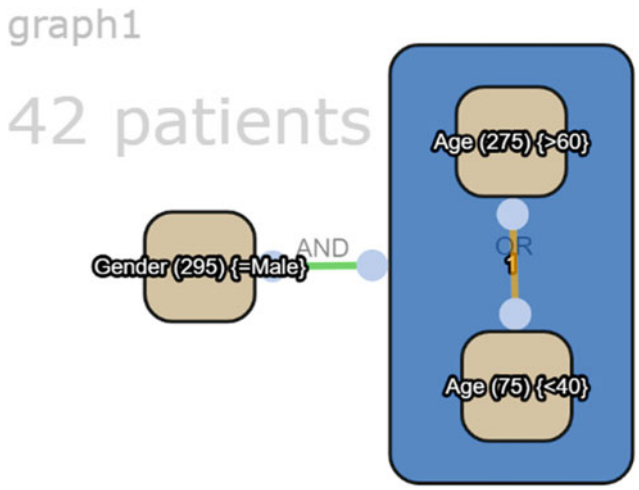



**Fig. 2** Query builder example

*over 60*), while the edges represent the logical condition between the features. The query builder provides the possibility to the user to create more complicated queries using groups e.g. *(age under 40 OR age over 60) AND gender male* as shown in Fig. 2.

*Visualization*

Query builder provides the capability to the user to view more statistics of the generated query at any time and update/modify the query accordingly. The results of each query can be viewed in a graphical way, using various charts, enabling further exploration and enhancement. Each chart can be used as a filter and give instant feedback. The graphical view of the query results from an example query is shown in Fig. 3. Features with numeric values such as Age are visualized as bars charts while the nominal features such as Gender are visualized as pies. The total number of patients is shown to the right of the viewer area. All the charts can be used as single filters or multiple filters (using the logical AND operation for more than one filters).

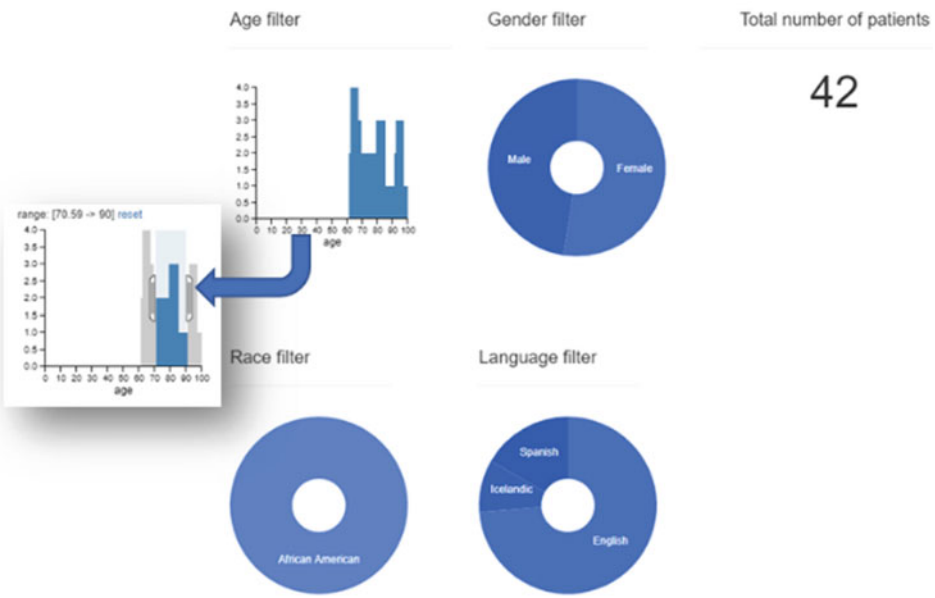
*Data Analysis*

Data mining consists of various methods and algorithms, which have been applied to many research areas and the healthcare domain is not an exception. Main objective of the analytics framework is to hide the complexity of data mining from the end user.

Handling of the diverse and large amount of data is supported by feature elimination algorithms, typically used in the demanding domain of bioinformatics [13, 14] while clustering algorithms provide similarity matching to cases/patients. The reduction of the feature set and the selection of the most relevant features could help to cope with highly dimensional data (e.g. lifestyle data from smart devices), reduce computational cost, and improve classification performance. The framework uses the principal variables [15] in order to select a subset of variables that contain, in some sense, as much information as possible and propose the most informative variables of a cohort to the end user.

One of the most important questions in data analysis is to find the "similar" cases/records in our data. Cluster analysis has been widely used in patient orientated management strategies and identify discrete groups of patients with specific combinations of comorbid conditions [16]. The analytics framework uses the K-Means algorithm [17], one

**Fig. 3** Visualization



of the well-known unsupervised learning algorithms, which clusters data by trying to separate samples in n groups of equal variance, minimizing a criterion known as the inertia or within-cluster sum-of-squares.

## Preliminary Evaluation and Discussion

The proposed data analysis framework has been adopted by the iManageCancer European project. The iManageCancer project aims to support chronic cancer treatment via a cancer disease self-management platform focusing on the wellbeing [18] and patient empowerment [19].

As such the analytics framework has been linked with the iManageCancer data sources using the semantic and interoperability layers. During the implementation of the aforementioned analytics platform, the iManageCancer pilots were in the process of preparation. Therefore, no real data were available. For that reason, we generated artificial but realistic data for testing and development of the framework and the algorithms. The framework, using the artificially generated data, was empirically evaluated by four experts (two physicians, one research nurse and one data analyst) As it was expected the end users focused mainly on the visualization of data and analysis rather that the analytical algorithms. They identified the whole framework as a highly added-value instrument, easy to be understood and used. The query builder was the only part of the analysis workflow that the users were not familiar but after a while, using the trial and error method, all the users managed to create simple and complex queries. The results of a thorough evaluation, with real-datasets will be published in a follow-up after the completion of the pilots.

## Conclusions

Nowadays, medicine combines data collected over time about an individual's genetics, environment, and lifestyle and focuses on the integrated diagnosis, treatment and prevention of disease in individual patients [20]. While the goal is clear, the path to such advances has been fraught with roadblocks mainly in the data management and data integration areas.

We propose a multi-layer framework architecture and we believe that consolidating healthcare data into comprehensive and coherent assets with an analytics frontend on top will aid in the precision medicine area. The architecture is able to combine heterogeneous healthcare sources by exploiting IHE profiles, integrate efficiently big data using semantics and provide anonymized healthcare data over a modular analytics framework. The proposed architecture/framework has been adopted by the iManageCancer project but it could also be used out of the project's context.

**Acknowledgements** This work has been supported by the iManageCancer H2020 EU programme under grant agreement No 643529.

**Conflict of Interest** The authors declare no conflict of interest.

## References

1. Yoo I, Alafaireet P, Marinov M et al (2012) Data mining in healthcare and biomedicine: a survey of the literature. J Med Syst 36:2431–2448. https://doi.org/10.1007/s10916-011-9710-5
2. Potamias G, Koumakis L, Moustakis V (2005) Mining XML clinical data: the healthobs system. Ingénierie des systèmes d'information 10:59–79
3. Reddy CK, Aggarwal CC (2015) Healthcare data analytics. CRC Press
4. Belle A, Thiagarajan R, Soroushmehr SMR, et al (2015) Big data analytics in healthcare. Hindawi Publ Corp, pp 1–16. https://doi.org/10.1155/2015/370194
5. IHE. https://www.ihe.net
6. iManageCancer. http://imanagecancer.eu
7. Chadwick DW, Lievens SF (2008) Enforcing "sticky" security policies throughout a distributed application. In: MidSec '08 proceedings of the 2008 work middleware security, pp 1–6. http://doi.acm.org/10.1145/1463342.1463343
8. QRPH. https://www.ihe.net/Quality_Research_and_Public_Health/
9. Clinical research process content. http://wiki.ihe.net/index.php/Clinical_Research_Process_Content
10. Research Matching. http://wiki.ihe.net/index.php/Research_Matching
11. Kondylakis H, Bucur A, Dong F, et al (2017) iManagecancer: developing a platform for empowering patients and strengthening self-management in cancer diseases. In: 30th IEEE International Symposium on Computer-Based Medical Systems. IEEE CBMS
12. Pappas A, Troullinou G, Roussakis G, et al (2017) exploring importance measures for summarizing RDF/SKBs. In: Blomqvist E, Maynard D, Gangemi A, et al (eds) Proceedings of the 14th international semantic web conference, ESWC 2017, Portorož, Slov. 28 May–1 June 2017, Part I. Springer International Publishing, Cham, pp 387–403
13. Potamias G, Koumakis L, Moustakis V (2004) Gene selection via discretized gene-expression profiles and greedy feature-elimination. In: 2004 Proceedings of the methods and applications of artificial intelligence: third helenic conference on AI, SETN 2004, Samos, Greece, 5–8 May 2004, pp 256–266
14. Koumakis L, Moustakis V, Zervakis M, et al (2012) Coupling regulatory networks and microarays: Revealing molecular regulations of breast cancer treatment responses Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) pp 239–246
15. McCabe GP (1984) Principal variables. Technometrics 26:137–144. https://doi.org/10.2307/1268108
16. Newcomer SR, Steiner JF, Bayliss EA (2011) Identifying subgroups of complex patients with cluster analysis. Am J Manag Care

17. MacQueen JB (1967) Kmeans some methods for classification and analysis of multivariate observations. In: 5th Berkeley Symposium on Mathematical Statistics and Probability 1967, vol 1, pp 281–297. doi:citeulike-article-id:6083430
18. Kondylakis H, Kazantzaki E, Koumakis L et al (2014) Development of interactive empowerment services in support of personalised medicine. Ecancermedicalscience 8
19. Kondylakis H, Koumakis L, Kazantzaki E et al (2015) Patient Empowerment through Personal Medical Recommendations Stud Health Technol Inform p 1117
20. Huang BE, Mulyasasmita W, Rajagopal G (2016) The path from big data to precision medicine. Expert Rev Precis Med Drug Dev 1:129–143. https://doi.org/10.1080/23808993.2016.1157686

# Towards Harmonized Data Processing in SMBG

Sara Zulj, Goran Seketa, and Ratko Magjarevic

**Abstract**

Self-monitoring of blood glucose (SMBG) is the key activity in diabetes management. Patients are required to take measurements and act accordingly, while the physicians use measured data to adjust the therapy. Though the accuracy of individual glucose meters used for SMBG is limited, the main difficulty in interpretation of the recorded data is due to inaccuracy of the records made by patients themselves into the paper diabetic diary. Oftentimes, patients do not record data properly and therefore the data is not reliable for use in determining long-term changes and trends or to use it for further analysis. Therefore, analysis and decision making should rely on the values recorded and stored in glucose memory. The large variety of glucometer models on the market introduce a large problem in using the recorded values since companies which produce and sale glucometers do not necessarily base their data transmission code on accepted standards but they embed custom made code. Data from 37 models of glucometers is transferred into a cloud based platform using previously developed system and available for immediate analysis and for saving into an appropriate health registry in a harmonized structure despite differences in protocols and data structure of different meters. Immediate statistics are given to the physicians upon patient's checkup. However, general statistical metrics usually do not include metrics on glucose variability, which is one of the most important measurements of glycemic control. We added glycemic variability metrics, including other metrics into tool for data analysis using MATLAB. The output of the analysis can be stored in the system and can be combined with the existing healthcare registries to develop multidimensional analysis for new knowledge discovery. This paper describes the system for acquisition of SMBG data, MATLAB analysis software and the notes on the analysis of the previously discussed data set.

**Keywords**

Glucose measurement • Data acquisition • Glucometer • SMBG • MATLAB

## Introduction

Self-monitoring of blood glucose (SMBG) refers to home blood glucose testing for diabetic patients. Guidelines for treatment of diabetics with type 2 diabetes mellitus (T2DM), who are the majority of patients with diabetes, recommend structured self-monitoring of blood glucose (SMBG) for better glucose control through self-efficiency, increased knowledge on their own condition and in changing of life style. Regular blood glucose monitoring provides immediate feedback on how the taken actions are working. SMBG is thus the key activity in diabetes management.

Although clinically approved testing of glycated hemoglobin (HbA1c) gives clinicians good insight into a general

S. Zulj (✉) · G. Seketa · R. Magjarevic
Faculty of Electrical Engineering, University of Zagreb FER, Unska 3, Zagreb, Croatia
e-mail: sara.zulj@fer.hr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_11

measure of blood sugar control over the previous two to three months, if often neglects important daily glycemic excursions and dangerous hypoglycemia events [1]. Without glucose data, the interpretation based on HbA1c only gives little possibility for feedback and for improved therapy and expected outcomes of SMBG associated with it.

Patients are taught to use their immediate SMBG results to try and keep their blood glucose levels within the desired target range. The correction is made by changing the carbohydrate intake, exercising, or using insulin. The frequency with which patients with diabetes should monitor their blood glucose level varies from patient to patient, but most experts agree that insulin treated patients should monitor blood glucose at least four times a day, most commonly fasting, before meals, and before bed [2]. On the other hand, SMBG allows clinicians to adjust patients' therapy and manage their patients' glucose levels more effectively.

SMBG generates massive amounts of data which should be acquired and stored properly. Recorded values of glycaemia (electronic dairies, i.e. data on glycemic values from blood glucose measurement devices, glucose meters) contribute to a physician's better overview of patients' conditions. However, in daily routine, there are several problems to overcome in order to achieve improvement in individual glycemic control. Firstly, the culture of accurate glycemic data acquisition is partially overcome by the introduction of digital memory to the glucometers so that inaccurate entry into paper glucose dairies does not exist in well trained patients [3]. Still, a vast majority of patients does not bring their meters to routine checkups at their GPs.

The availability of SMBG data, allows clinicians and patients to quickly identify glycemic patterns and make more informed decisions about therapeutic adjustments that may be required [4].

Clinicians must analyze and interpret these datasets within a few minutes during a checkup to identify the frequency, severity, and timing of hypoglycemic and hyperglycemic episodes [5]. Complex systems often do not offer a fast and intuitive way to allow easy use for clinicians, where systems which use only basic information do not allow any customization on selected analysis.

There have been a lot of discussions on the effectiveness of SMBG in T2DM patients, which resulted in health technology studies and meta-analysis of literature. These studies showed that the technology implemented for SMBG for structured self-monitoring has a potential which is rarely used: the information from measurements is not well integrated into diabetes management and there is a lack of feedback to patients. Moreover, rarely are glucose measurements stored and kept in electronic health records for trend analysis and pattern management [6], partially because of many glucose measurement systems present at the market and low standardization or harmonization of the data transfer protocols [7]. It is a fact that in conclusions of those reviews there is the statement pointing out that the results of SMBG could improve more advanced use of the acquired data and better training of their analysis. From the point of technology, the feedback loop may be closed even more often that through regular checkups by means of health [8]. Systems for wired remote monitoring have been successfully introduced into clinical settings and lately a number of applications for telemonitoring have been adopted for use by mobile phones [7, 9]. With accessible and accurate data, adopting of individual approach in the treatment of persons with type 2 diabetes mellitus, as recommended by international guidance documents, should not be considered as a problem in any health system.

## Acquisition System for SMBG Data

Our system for acquisition and analysis of SMBG data has previously been developed due to a large number of different glucose meters available on Croatian market. (42 types of glucose meters from 17 manufacturers). The system aims to help clinicians make more informed decisions by providing information on statistics in a chosen time span. Statistics include total number of blood glucose measurements, minimal, maximal and average value of measurements, number of days without measurement and average number of measurements per day, etc. It also includes graphical representations of the data set such as trend, modal day or modal week. The description of the data acquisition system which is used to obtain the data and store it in the cloud is more thoroughly described in [7].

## Matlab-Based Tool for Analysis

The clinician is shown the described statistics and graphical representation of SMBG data for the immediate patient at the checkup. However, the statistics shown to physicians often do not provide complete information on glucose profile that is representative of daily glucose excursions. The MATLAB based tool that we implemented should provide enough customization if needed.

Because glucose values from SMBG do not follow a Gaussian or normal distribution, but the distribution is usually skewed to the right, two common metrics need to be addressed first. The mean and standard deviation (SD) as measures of central tendency (average) and variability, which are very sensitive to outliers, are implemented median and inter-quartile range as metrics which represent dataset more naturally [10].

The tool can be customized to allow users to export relevant metrics for one or more patients. It allows the option

for changing timespan by entering dates. Euglycemic range can be also set for individual patients or for the observed group. But most importantly, measurements of glucose variability such as the mean amplitude of glycemic excursion MAGE, low blood glucose index (LBGI), high blood glucose index (HBGI) have been implemented [11]. The tool also allows to divide SMBG values into time-categories such as nocturnal/daytime measurement and weekend/work day measurement. Several plots are available, such as a histogram of measurements per time of day is added next to modal day, and option for separating bin by hypo-, eu- or hyperglycemia values. The tool utilized the automated script so that selected metrics can be exported to .csv file and plots are generated in the output folder as image files or comma separated values file.

In order to use the tool, users must place data files in comma separated values format into *input* folder. Data files must consist of date of measurement, time of measurement and measurement value. File name corresponds to the patient ID, gender (F/M) and age. Once the data files are in the *input* folder the user can start a script. Depending on the input parameters, user can choose to enter *number of day* or *start date and end date* for limiting data timeline, number of graphical representation (one or more) from the available (trend, modal day, histogram of measurements per hour of day grouped by 2, 4, or 6 h with option to section the ratio of hypo- and/or hyperglycemias), *euglycemia limits* and *per day/total* parameter which determines whether statistics should be calculated by day or in total for each patient. The output is generated as PNG images for each patient for graphical representations, and statistics are given in the form of a CSV file which include patient ID, gender, age, mean, standard deviation, median, interquartile range, total number of measurements, number of days without measurement, number of hypo-, eu- and hyperglycemias nocturnal and daytime, MAGE, LBGI and HBGI.

## Testing the Tool and Results

While testing the tool, we observed the data from the database for patients on intensified therapy. In this case their treatment consists of three or more daily insulin injections or CSII. The main reason for choosing this group of patients was relatively high number of measurements (ideally 3–4 measurements per day) due to the nature of therapy.

The initial data was extracted from database for patients whose average number of performed measurements per day, not including the days that measurements that no measurement was performed, was 3 or more. This yielded data from 603 patients, which included 782.934 SMBG measurements. The last 500 patients were taken into further analysis. We calculated the average number of measurements per day for those patients within the last 90 days prior to the checkup. Time of the checkup was taken from the date of the last data acquisition, which is done in clinicians' office. 103 patients with average number of measurement more than three (33 with more than four) were furtherly extracted with this method.

Even though many glucose meters allow adding flag to the measurement (e.g. fasting, pre-prandial/2-h postprandial at each meal, bedtime, after exercise), the vast majority of patients does not use that option. Therefore, the information on the timing regarding recommended times in measuring structure is neglected. This problem could potentially be overcome by using the mobile app to transfer the data, which would then include other relevant information on meals, therapy and exercise. Also, the time of the glucometers records was corrupted due to the inaccurate setting of the time.

The tool should be modified to allow importing other data, i.e. data from registries, to utilize them and extent the analysis combined with SMBG data and metrics.

## Discussion and Conclusions

Statistical measures often fail to show specific situations. For example, Fig. 1 shows that example of the data from a patient who on average measures more than 3 times a day, but 35 days within 90 days were without any measurement. Also, trend graph showed that most measurements were made in a few days prior to the visit. This can influence the metrics, and clinician tools could use such knowledge to help better understand given metrics and their reliability.

Data collected from patients regularly visiting specialist shown that SMBG data is collected unstructured. Figure 2 shows modal days from the data from two different patients in last 90 days prior to the visit along with the histograms of the measurement on 2-h intervals during the day. The histogram shows the number of hypoglycemic events within all measurements. Figure 2a the patients who measured at consistent times per day, and Fig. 2b shows a patient with no consistency in regard to time of the measurements. A closer

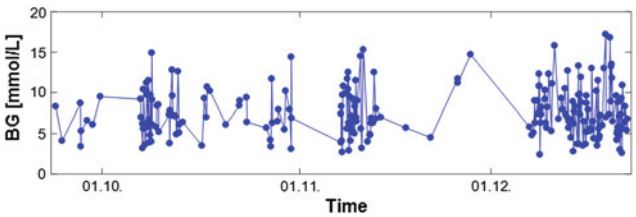



**Fig. 1** Trend graph of patient with inconsistent timing of measurements

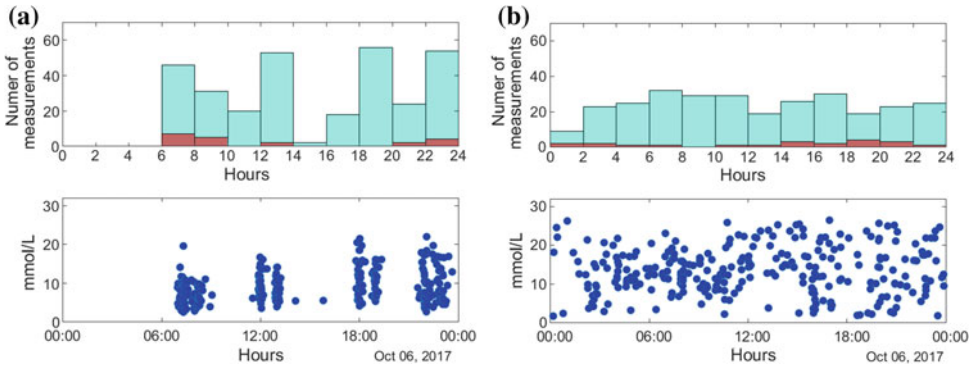



**Fig. 2** Modal day graphs and histograms of measurements per 2-h including hypoglycemia for **a** patients with consistency in timing of measurements, **b** patient with inconsistency in timing of measurements

evaluation showed that even the patient in Fig. 2a does not comply with the rules of structured SMBG measuring.

Data presented in this paper show that medical systems developed for SMBG enable valuable information to physicians and persons with diabetes. However, some of the studies tend to question the efficiency of their use due to unproved short-term results. Due to large prevalence of diabetes, benefits for individuals who were trained in application of and understand the contemporary technology, should be accepted and for increasing the efficiency of the systems for SMBG, more effort should be given to proper training of upstanding use of technology. Also, further development and more sound presence of continuous glucose measurement systems at the market at affordable prices should encourage health policy makers to introduce them since their characteristics already proved its potential in enabling better control of glycaemia [12].

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Dailey G (2007) Assessing glycemic control with self-monitoring of blood glucose and hemoglobin A(1c) measurements. In: Mayo clinic proceedings. Mayo Clinic, vol 82, no 2, pp 229–235; quiz 236. https://doi.org/10.4065/82.2.229
2. Benjamin EM (2002) Self-monitoring of blood glucose: the basics. Clin Diabetes 20(1):45–47. https://doi.org/10.2337/diaclin.20.1.45
3. Prašek M (2011) Self-control diary—challenges of new technological possibilities, pp 972–973. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23508-5_252
4. Parkin CG, Davidson JA, Affiliations A (2009) Value of self-monitoring blood glucose pattern analysis in improving diabetes outcomes. J Diabetes Sci Technol 33(33):500–508
5. Rodbard D (2009) Display of glucose distributions by date, time of day, and day of week: new and improved methods. J Diabetes Sci Technol 3(6):1388–1394. https://doi.org/10.1177/193229680900300619
6. Clar C, Barnard K, Cummins E, Royle P, Waugh N, Aberdeen Health Technology Assessment Group (2010) Self-monitoring of blood glucose in type 2 diabetes: systematic review. Health Technol Assess 14(12):1–140. https://doi.org/10.3310/hta14120
7. Žulj S, Celić L, Grgurević M, Prašek M, Magjarević R (2016) Pilot project: ICT system for management and self- management of diabetes, 29 June
8. Greenwood DA, Blozis SA, Young HM, Nesbitt TS, Quinn CC (2015) Overcoming clinical inertia: a randomized clinical trial of a telehealth remote monitoring intervention using paired glucose testing in adults with type 2 diabetes. J Med Internet Res 17(7): e178. https://doi.org/10.2196/jmir.4112
9. Wojcicki JM, Ladyzynski P, Foltynski P (2013) What we can really expect from telemedicine in intensive diabetes treatment: 10 years later. Diabetes Technol Ther 15(3):260–268. https://doi.org/10.1089/dia.2012.0242
10. Rodbard D (2007) Optimizing display, analysis, interpretation and utility of self-monitoring of blood glucose (SMBG) data for management of patients with diabetes. J Diabetes Sci Technol (Online) 1(1):62–71. https://doi.org/10.1177/193229680700100111
11. Kovatchev B, Cobelli C (2016) Glucose variability: timing, risk analysis, and relationship to hypoglycemia in diabetes. Diabetes Care 39(4):502–510. https://doi.org/10.2337/dc15-2035
12. Kubiak T, Mann CG, Barnard KC, Heinemann L (2016) Psychosocial aspects of continuous glucose monitoring: connecting to the patients' experience. J Diabetes Sci Technol 10(4):859–863. https://doi.org/10.1177/1932296816651450

# Notarization of Knowledge Retrieval from Biomedical Repositories Using Blockchain Technology

P. Mytis-Gkometh, G. Drosatos, P. S. Efraimidis, and E. Kaldoudi

**Abstract**

Biomedical research and clinical decision depend increasingly on a number of authoritative databases, mostly public and continually enriched via peer scientific contributions. Given the dynamic nature of data and their usage in the sensitive domain of biomedical science, it is important to ensure retrieved data integrity and non-repudiation, that is, ensure that retrieved data cannot be modified after retrieval and that the database cannot validly deny that the particular data has been provided as a result of a specific query. In this paper, we propose the use of blockchain technology in combination with digital signatures to create smart digital contracts to seal the query and the respective results each time a third-party requests evidence from a reference biomedical database. The feasibility of the proposed approach is demonstrated using a real blockchain infrastructure and a publicly available medical risk factor reference repository.

**Keywords**

Biomedical repositories • Cryptographic techniques • Blockchain • Integrity Non-repudiation

## Introduction

Biomedical research and clinical practice relies increasingly on authoritative data gathered and curated in reference biomedical databases. Examples include: clinical databases (registries or academic clinical databases) that hold clinical data on patient cohorts [1]; biomedical databases [2] with current data on pharmaceuticals [3], metabolomics [4], inheritance data and other omics (for example, the rich collection available from the European Bioinformatics Institute at http://www.ebi.ac.uk/services); publication repositories and other medical evidence repositories [5], either general purpose (the most prominent example being PubMed service by the National Library of Medicine, USA) or high evidence quality, such as Cochrane Library reports.

Biomedical references databases are continually updated to include new data sets (e.g. PubMed included ~1 M new records in 2016), and are often validated and, if necessary, updated to correct existing data. At any given point in time, these data are heavily accessed by humans (clinicians, patients and researchers alike) and software (via appropriate application programming interfaces) to establish current evidence and inform clinical acts and biomedical research. As such, it is important to ensure that data cannot be manipulated retrospectively and that data 'consumers' can have a proof of what data were retrieved from the database at a given point in time as a result of a specific query.

A reliable knowledge retrieval service has to fulfill at least the following two important requirements; integrity and non-repudiation. Integrity, means that the query and the retrieved data cannot be modified (either by accident or

P. Mytis-Gkometh · P. S. Efraimidis
Department of Electrical and Computer Engineering, Democritus University of Thrace, 67100 Kimmeria, Xanthi, Greece

G. Drosatos (✉) · E. Kaldoudi
School of Medicine, Democritus University of Thrace, University Campus, Dragana, 68100 Alexandroupoli, Greece
e-mail: gdrosato@ee.duth.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_12

deliberately), once the retrieval operation completes. Non-repudiation, in this context means that given any past retrieval operation, the knowledge retrieval service cannot validly deny that the exact data have been provided by the service as a response to the given query at the specific time. An interesting solution that satisfies the above requirements can be found in the emerging field of blockchain infrastructures. Blockchains inherently ensure the integrity of each recorded transaction. Moreover, non-repudiation can also be accomplished if blockchains are for example combined with digital signatures.

In this paper, we propose the use of blockchain technology to create smart digital contracts to seal the query and the respective results each time a third-party requests evidence from a reference biomedical database. The proposed approach is demonstrated on the powerful Ethereum blockchain platform [6] with a retrieval service for the publicly available CARRE risk factor reference repository [7]. The repository has been developed in the context of the European Union funded FP7-ICT project CARRE (Grant no. 611140), which researched and developed novel personalised decision support services for managing comorbidities associated with cardiorenal disease.

## Background

Blockchain is a distributed, incorruptible transaction management technology without one single trusted party. Each new transaction is broadcasted to a distributed network of nodes; once all nodes agree the transaction is valid, the transaction is added to a block. Every block contains a timestamp and the hash (cryptographic seal) of the previous block and the transaction data, thus creating an immutable, append-only chain. Copies of the entire blockchain are maintained by each participating node.

The first blockchain was proposed for and implemented in Bitcoin [8], a distributed infrastructure where users can make financial transactions without the need of a regulator (e.g. a bank). Nowadays, other blockchain infrastructures are emerging, for example the Ethereum [6], where everyone can participate in the blockchain generation, and the Hyperledger Fabric [9], where only approved parties can post to the blockchain. In permissionless blockchains like Bitcoin and Ethereum, all transactions are public, however, no direct links to identities exist. When applied to financial transactions, this privacy preserving features can be enhanced even further [10]. However, in applications that require non-repudiation, identity should be irrevocably maintained; this can be ensured by the appropriate use of public key infrastructure solutions [11, 12].

A recent systematic review on current state, limitations and open research on blockchain technology [13] discusses a number of blockchain applications that extend from cryptocurrency to Internet of things, smart contracts, smart property, digital content distribution, Botnet, and P2P broadcast protocols. Currently, there is considerable optimism that blockchain technology will revolutionize the healthcare industry [14]. Indeed, blockchain technology has been proposed as a solution for privacy-preserving control and sharing of patient personal healthcare data [15–17] and for record management in clinical trials to ensure that data is fully published and not tampered with [18, 19].

## Query Notary Service

In this paper, we propose a lightweight wrapper for conventional databases that uses blockchain technology to offer database query notary services to data consumers (humans and programs alike). The proposed notary service administers contracts that seal a query placed to a database and the returned results. The service offers irrevocable proof of data retrieved by a specific query placed by a specific consumer, thus establishing query transaction integrity and non-repudiation. In this way, the proposed system assures that the consumer is protected against a service that may accidently or intentionally try to repudiate or alter a past query transaction.

The overall architecture is presented in Fig. 1. The main component is the blockchain contract service that acts as a mediator between conventional biomedical databases and data consumers. The structure of the biomedical knowledge could be any database model (i.e. SQL or NoSQL databases), or even semantic repositories (i.e. RDF stores).

The proposed notary service exhibits three computational layers: (a) a data consumer front-end, which can be either an interface for human data consumers or an application programming interface (API) for 3rd party programs that request data from a biomedical database; (b) an interface to communicate with biomedical database interfaces, which is specific to each database API; and (c) the contract generation engine, which collates the query/results data and the consumer, prepares transactions and submits them to a blockchain infrastructure, and stores contract information (contract address and its application binary interface).

The workflow of the notary service is as follows. First, the data consumer front-end undertakes the communication with the party placing the query to the database. In its simplest version, the query is forwarded to the database API via the database API client. As an added-value, the query can also be signed by a public key infrastructure to verify later the identity of the data consumer. The API client places the query via the database API and retrieves the results; both (signed) query and results are forwarded to the blockchain contract service. Subsequently, these data are hashed (e.g.

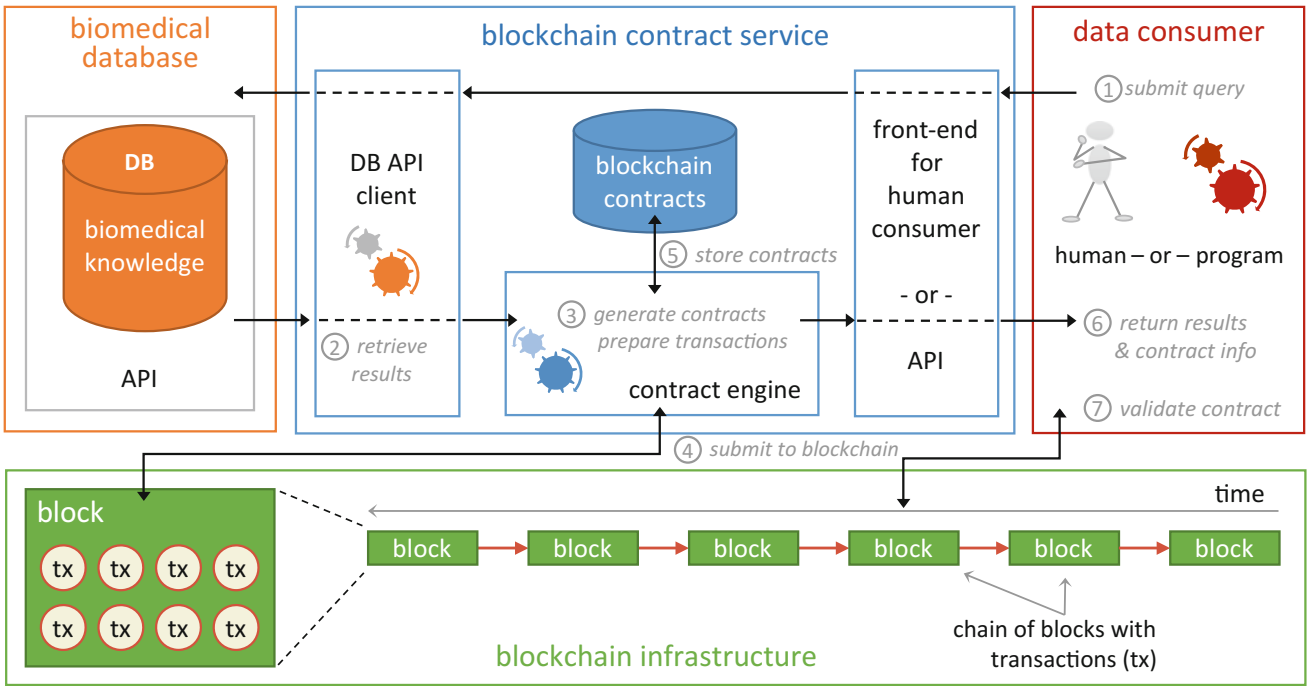



**Fig. 1** The general architecture of our query notary service

using SHA256 [20]) and the hash is included in a smart contract that is deployed to a blockchain infrastructure.

The contract generation engine then returns the query results to the data consumer via the front-end, accompanied by the smart contract's address on the blockchain, the application binary interface (ABI) to interact with the contract, and the (signed) query and its results. A respective entry is also made into the local contract database. The packet returned to the data consumer contains also database certification information to verify the identity of the database and thus ensure query transaction non-repudiation (for example, the database blockchain public key signed by a digital certification authority). The consumer archives the query transaction (query and signed response) in a local database for future reference.

At any later time, the data consumer or any third party can verify the query transaction dataset by retrieving the respective contract from the blockchain infrastructure and comparing the retrieved hash of the original data with a new hash of the claimed (signed) query and respective results.

## Implementation and Experimental Results

The proposed architecture was implemented to provide query notary services for the CARRE risk factor reference repository [7], an open, online database collecting current high-level evidence on risk factors for the cardiorenal syndrome and related comorbidities. In this repository, risk factors are described in a structured way following the CARRE risk factor ontology [21]. Risk evidence descriptions are manually entered by authorized medical experts following a collaborative literature survey process by which appropriate medical publications of high level medical evidence are identified in PubMed and used to extract state-of-the-art medical evidence on risk factors related to cardiorenal disease. The resulting risk factor descriptions are available as Linked Data, following the Resource Description Framework (RDF) format (http://www.w3.org/TR/rdf-syntax), via an open access RDF repository. Currently the CARRE risk factor repository describes more than 100 different risk factors corresponding to 250 risk associations between more than 50 medical conditions related to cardiorenal disease as retrieved from 65 scientific publications.

In this demonstration, blockchains were implemented using the Ethereum infrastructure [6], which in addition to the transaction introduces a programmable logic into the blocks. This functionality allows for smart programmable contracts that have the ability to work autonomously in an if-this-then-that fashion. The use of the Ethereum blockchain infrastructure requires running an Ethereum node using the Geth client (version 1.5.9). Smart contracts are implemented in the Solidity language (https://solidity.readthedocs.io), while a MongoDB database (https://www.mongodb.org) is deployed for the local storage of contracts and respective information.

The front-end was implemented using JavaScript and Ajax asynchronous requests to establish communication

with the CARRE RDF repository via its SPARQL endpoint client (https://devices.duth.carre-project.eu/sparql). The Meteor web framework (https://www.meteor.com) was used to connect the front-end with the backend of the query notary service.

Figure 2 shows a snapshot of the notary service as implemented for CARRE risk factor repository. When a new query is placed via the front end, the notary service communicates with the CARRE repository and creates a smart contract out of the query and the returned results. The contract is compiled locally and then the bytecode is deployed on the Ethereum blockchain. The front end provides information on the status of the procedure and information on past queries and their respective application binary interface on the Ethereum blockchain.

Experimental verification used the Ropsten Ethereum test network (https://testnet.etherscan.io), which simulates the blockchain environment but is provided for free. At the time of testing in July 2017, the transaction confirmation delay was 30–50 s, and the cost of placing a transaction into the Ethereum blockchain was 0.00302 Ether (Ethereum's cryptocurrency). With an exchange rate of 1 Ether = 171.64€ (31 Jul 2017, www.worldcoinindex.com), the cost of one transaction was about 0.52€.

## Discussion

This paper proposes a query notary for biomedical data consumers (humans or programs alike) who need to retrieve accurate and certified data from reference biomedical databases. The proposed approach utilizes blockchain technology and is implemented using a real blockchain infrastructure and a publicly available medical reference repository. The notary is realized following the approach of software-as-a-service, thus bringing the cost to the data consumer on a needs basis. Additionally, a private blockchain network maintained by health regulators, such as healthcare establishments and medical research organizations (similar to that proposed in [19]) could be established to alleviate this cost. Work in progress includes design and development of a mechanism that exploits the blockchain to provide irrefutable version control of database content and give undisputable proof that the results returned via a query correspond to the most recent update. Additional work addresses smart contracts involving dynamic graph data (e.g. Linked Open Data cloud datasets), where the question is to combine certified sub-graphs (for example from different repositories) in order to validate larger, integrated data graphs.

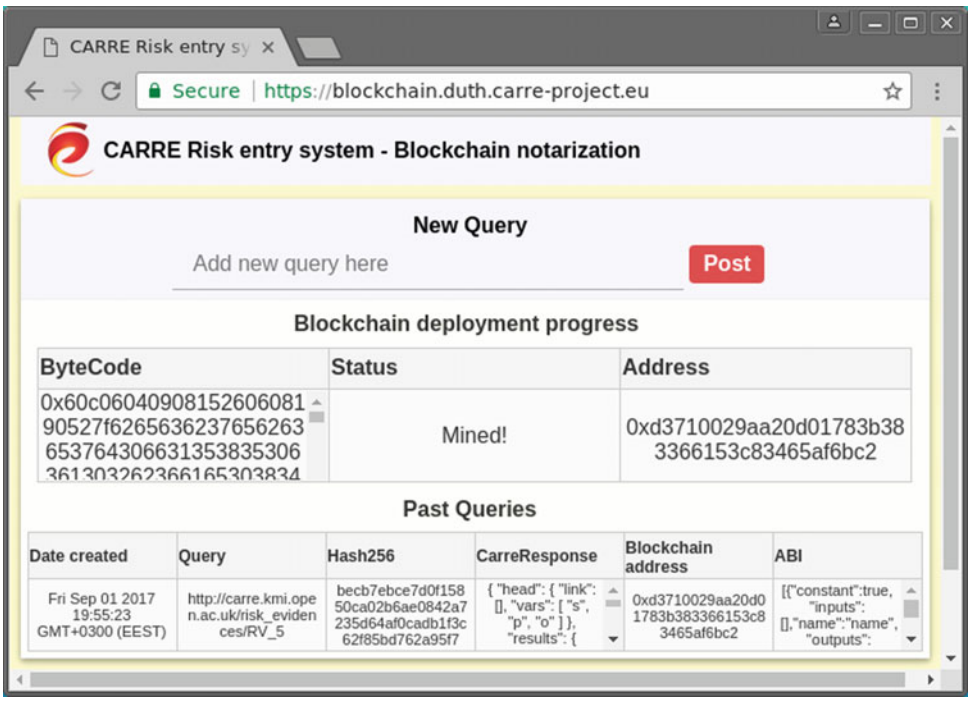



**Fig. 2** Snapshot of contract deployment process in blockchain

**Acknowledgements** This work was partly supported by the European Commission FP7-ICT project CARRE (Grant No. 611140).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Williams WG (2010) Uses and limitations of registry and academic databases. Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu 13:66–70
2. Duck G, Nenadic G, Filannino M et al (2016) A survey of bioinformatics database and software usage through mining the literature. PLoS ONE 11:e0157989
3. Vaughan K, Scolaro KL, Anksorus HN et al (2014) An evaluation of pharmacogenomic information provided by five common drug information resources. J Med Libr Assoc 102:47
4. Go EP (2010) Database resources in metabolomics: an overview. J Neuroimmune Pharmacol 5:18–30
5. Falagas ME, Pitsouni EI, Malietzis GA et al (2008) Comparison of PubMed, scopus, web of science, and google scholar: strengths and weaknesses. FASEB J 22:338–342
6. Buterin V (2014) A next-generation smart contract and decentralized application platform. White Paper. http://www.ethereum.org/pdfs/EthereumWhitePaper.pdf
7. Carre P (2016) CARRE risk factor reference repository. FP7 EU project (FP7-ICT-611140). https://www.carre-project.eu/innovation/carre-risk-factor-entry-system
8. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf
9. Cachin C (2016) Architecture of the hyperledger blockchain fabric. In: Workshop on distributed cryptocurrencies and consensus ledgers (DCCL), Chicago, Illinois, USA
10. Herrera-Joancomartí J (2015) Research and challenges on bitcoin anonymity. In: Data privacy management, autonomous spontaneous security, and security assurance. LNCS, vol 8872. Springer, Cham, pp 3–16
11. Housley R, Ford W, Polk W et al (1998) Internet X. 509 public key infrastructure certificate and CRL profile, pp 2070–1721
12. Matsumoto S, Reischuk RM (2016) IKP: turning a PKI around with blockchains. IACR Cryptol ePrint Arch 2016:1018
13. Yli-Huumo J, Ko D, Choi S et al (2016) Where is current research on blockchain technology?—a systematic review. PLoS ONE 11: e0163477
14. Mettler M (2016) Blockchain technology in healthcare: the revolution starts here. In: 18th international conference on e-health networking, applications and services, Munich, Germany, pp 1–3
15. Yue X, Wang H, Jin D et al (2016) Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst 40:218
16. Roehrs A, da Costa CA, da Rosa Righi R (2017) OmniPHR: a distributed architecture model to integrate personal health records. J Biomed Inform 71:70–81
17. Azaria A, Ekblaw A, Vieira T et al (2016) Medrec: using blockchain for medical data access and permission management. In: International Conference on open and big data (OBD), pp 25–30
18. Carlisle BG (2014) Proof of prespecified endpoints in medical research with the bitcoin blockchain. Web Blog. The Grey Literature. https://www.bgcarlisle.com/blog/2014/08/25/proof-of-prespecified-endpoints-in-medical-research-with-the-bitcoin-blockchain/. Accessed 29 Aug 2017
19. Nugent T, Upton D, Cimpoesu M (2016) Improving data transparency in clinical trials using blockchain smart contracts. F1000Res 5:2541
20. Penard W, van Werkhoven T (2008) On the secure hash algorithm family. National Security Agency, Technical Report
21. Third A, Kaldoudi E, Gkotsis G et al (2015) Capturing scientific knowledge on medical risk factors. In: 1st International workshop on capturing scientific knowledge, collocated with the 8th international conference on knowledge capture (K-CAP), Palisades, NY, USA

# Gap Analysis for Information Security in Interoperable Solutions at a Systemic Level: The KONFIDO Approach

J. Rasmussen, P. Natsiavas, K. Votis, K. Moschou, P. Campegiani, L. Coppolino, I. Cano, D. Marí, G. Faiella, O. Stan, O. Abdelrahman, M. Nalin, I. Baroni, M. Voss-Knude, V. A. Vella, E. Grivas, C. Mesaritakis, J. Dumortier, J. Petersen, D. Tzovaras, L. Romano, I. Komnios, and V. Koutkias

**Abstract**

In this paper, we present a gap analysis study focusing on interoperability of eHealth systems and services coupled with cybersecurity aspects. The study has been conducted in the scope of the KONFIDO EU-funded project, which leverages existing security tools and procedures as well as novel approaches and cutting-edge technology, such as homomorphic encryption and blockchains, in order to create a scalable and holistic paradigm for secure inner and cross-border exchange, storage and overall handling of healthcare data in compliance with legal and ethical norms. The gap analysis relied on desk research, expert opinions and interviews across four thematic areas, namely, eHealth interoperability frameworks, eHealth security software frameworks, end-user perspectives across diverse settings in KONFIDO pilot countries, as well as national cybersecurity strategies and reference reports. A standards-based template has been created as a baseline through which the analysis subjects have been analyzed. The gap analysis identified barriers and constraints as well as open issues and challenges for information security in interoperable

J. Rasmussen (✉) · J. Petersen
MedCom, Odense, Denmark
e-mail: jar@medcom.dk

P. Natsiavas · V. Koutkias
Centre for Research & Technology Hellas, Institute of Applied Biosciences, Thermi, Greece

K. Votis · K. Moschou · D. Tzovaras
Centre for Research & Technology Hellas, Information Technologies Institute, Thermi, Greece

P. Campegiani
Bit4id s.r.l, Naples, Italy

L. Coppolino · L. Romano
Department of Engineering, University of Naples "Parthenope", Naples, Italy

I. Cano
IDIBAPS, Hospital Clinic de Barcelona, Universitat de Barcelona, Barcelona, Spain

D. Marí
eHealth R & D Unit, EURECAT, Barcelona, Spain

G. Faiella
Fondazione Santobono Pausilipon, Naples, Italy

O. Stan
CEA, LIST, Point Courrier 172, 91191 Gif-sur-Yvette Cedex, France

O. Abdelrahman
Depertment of Electrical and Electronic Engineering, Imperial College of Science Technology and Medicine, London, UK

M. Nalin · I. Baroni
Telbios s.r.l, Milan, Italy

M. Voss-Knude
Sundhed.dk, Copenhagen, Denmark

V. A. Vella
Agency for Health Quality and Assessment of Catalonia, Barcelona, Spain

E. Grivas · C. Mesaritakis
Eulambia Advanced Technologies Ltd, Athens, Greece

J. Dumortier
Time.lex, Brussels, Belgium

I. Komnios
Exus Software Ltd, London, UK



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_13

solutions at a systemic level. Recommendations derived from the gap analysis will be brought into the forthcoming phases of KONFIDO to shape its technical solutions accordingly.

**Keywords**
Gap analysis • eHealth • Interoperability • Cross-border health data exchange Cybersecurity

## Introduction

Important systematic efforts have been made recently in the field of digital security and data privacy protection worldwide. Especially in the EU, driven by the need for cross-border health data exchange across Member States, diverse projects have conducted research along this line, by focusing on access policies, as well as on how to enforce security measures in the underlying data transfers and subsequent storage. Equally important, interoperability is an aspect that is explicitly linked with the technical solutions that shall be in place and their user acceptance.

The EU-funded KONFIDO project (http://konfido-project.eu/) aims to leverage proven tools and procedures as well as novel approaches and cutting-edge technology, in order to create a scalable and holistic paradigm for secure inner and cross-border exchange, storage and overall handling of healthcare data in a legal and ethical way at both national and European level. The project aims to: (a) enhance the trust and security of interoperable eHealth services; (b) provide continuous validation and proof of concept demonstrations, and (c) focus on stakeholders, improving user acceptance as well as adherence to standards, legal rules and ethical directives. To achieve these goals, KONFIDO is organized in four distinct though complementary phases, interacting with each other, namely, 'User requirements'; 'Design'; 'Technology development'; and 'Integration, testing and validation'. As part of the 'User requirements' phase, a systematic gap analysis has been conducted focusing on the security and privacy mechanisms developed and/or used in other projects and initiatives in relation to interoperability at a systemic level.

In general, a gap analysis serves the purpose of identifying the difference (gap) between the current and the target state of affairs (product, process, organization, market etc.). Current state refers to the actual state regarding the analysis focus, while the target state refers to the future state where it should be. This entails the comparison between actual performance with potential or desired performance across a range of areas. For the current study, it means that it is possible to see how well a project, initiative, technology, solution etc. meets a set of requirements as identified as relevant for KONFIDO's further work and final output. This analysis is expected to significantly contribute to an improved design of the project's technical solution as a whole and each of its components/tools separately.

The current paper presents the methods employed as well as the consolidated outcomes of the gap analysis study.

## Material and Methods

### Material

As analysis subjects we identified a range of relevant projects, technologies, initiatives, end-user organizations and strategies across four thematic areas, which have been analyzed in detail in the scope of the gap analysis. These were:

- *eHealth Interoperability Frameworks*: Antilope [1], epSOS [2], the Joint Action to Support the eHealth Network (JASeHN) [3] and SemanticHealthNet [4].
- *eHealth Security Software Frameworks*: DECIPHER [5], OpenNCP [6] and STORK 2.0 [7].
- *End-user perspectives across diverse settings in KONFIDO pilot countries*: OUH Odense University Hospital & Svendborg Hospital (Denmark), Santobono Pausilipon Hospital (Italy) and Hospital Clínic Barcelona (Spain).
- *National cybersecurity strategies and reference reports*: the Danish, Italian and Spanish Cybersecurity Strategies as well as relevant ENISA reports [8, 9].

### Methods

For each area, a group of organizations and people with knowledge and experience in the particular area has been identified and a Working Group within the KONFIDO Consortium was formed per thematic area. The analysis subjects have been reviewed by topic experts against a baseline template of security-oriented criteria (or controls), primarily based on the ISO 27 k family of standards concerning information security [10]. The ISO 27 k family

consists of a broad range of standards that addresses information security either to a specific sector/technology, at an overall management level or with specific topic guidelines. In the scope of this study, the following standards have been considered:

(a) ISO/IEC 27002—Information technology—Security techniques—Code of practice for information security controls (https://www.iso.org/standard/54533.html);
(b) ISO/IEC 27010—Information technology—Security techniques—Information security management for inter-sector and inter-organizational communications (https://www.iso.org/standard/68427.html);
(c) ISO/IEC 27040—Information technology—Security techniques—Storage security (https://www.iso.org/standard/44404.html);
(d) ISO 27799—Health informatics—Information security management in health using ISO/IEC 27002 (https://www.iso.org/standard/62777.html);
(e) ISO 22857—Health informatics—Guidelines on data protection to facilitate trans-border flows of personal health information (https://www.iso.org/standard/52955.html), and
(f) ISO/IEC 25010—Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models (https://www.iso.org/standard/35733.html).

Based on the above standards and with input from HIMSS EMR Usability Evaluation Guide for Clinician's Practices (http://www.himss.org/himss-emr-usability-evaluation-guide-clinicians-practices-sample-post-test-questionnaires), a systematic and detailed template for each of the four thematic areas included in the gap analysis was produced. Each Working Group was provided with a baseline template of security-oriented controls organized in various sections and worked in parallel as no group was dependent on the work of the others. In summary, the baseline template relied on the following, upper–level structure: *Security policy; Organizing information security; Asset management; Human resources security; Physical and environmental security; Communications and operations management; Access Control; Information systems acquisition, development and maintenance; Information security incident management; Business continuity management; Compliance,* and *Usability*.

Instructions on how to use the template were offered to the respondents via examples. Furthermore, predefined sets of responses have also been defined where applicable, in order to prevent errors while facilitating the analysis process for the reviewers. Introductory, interim and final meetings were held across the Working Groups to discuss the plan, their progress and results, respectively. Despite that the utilized templates were identical for the four types of subjects considered in the gap analysis, these were distributed with the understanding that for some of the analysis subjects, a part of the security aspects and control questions might not be relevant in all cases, due to the level of details and specifics that were addressed. Respondents were notified of this aspect and, given their knowledge and expertise, they were able to evaluate the relevance and applicability of each security aspect.

The gap analysis data on each analysis subject were obtained via one or a combination of the methods listed below:

- *Desk research*: A review of available material on the analysis subject. This can be project material (reports/deliverables, presentations, videos etc.) as well as articles and other published material.
- *Expert opinion*: The information is from a source with concrete detailed expertise about the analysis subject. It can be someone who was directly involved in the project, or who works in the respective organization.
- *Interview*: The analysis subject was analyzed through an interview with a person with particular knowledge on the analysis subject and security aspects.

As a result, a comprehensive dataset was compiled for our analysis, but the process also flagged some weaknesses and risks, e.g. relevant projects and initiatives are still evolving. Therefore, mitigations were put in place during the process when possible, making the gap analysis an iterative process.

## Results

### eHealth Interoperability Frameworks

The gap analysis of existing eHealth interoperability frameworks showed that while the SemanticHealthNet project refers to information security and references ISO standards as guidelines, the project per se does not provide details concerning information security. The Joint Action to support the eHealth Network (JAseHN) focuses on semantic interoperability of exchanging information in a cross-border fashion and Antilope on quality management and testing processes. In both cases, while some security aspects identified in the baseline template are addressed in the projects, most topics can be considered out of these projects' scope. Concerning epSOS, a distinction must be made between the project itself and the open-source software that came out of the project, namely, OpenNCP (included in the gap analysis of eHealth Software Security Frameworks).

Some indicative common gap patterns were identified through the reviewed interoperability projects/initiatives:

- Cases of inadequate information on aspects concerning 'Information security policy' and 'Management commitment'. For epSOS, the scope of the security policy might have been too narrow concerning also the National Contact Points (NCP) and internal organizational matters.
- 'Assets management' in terms of responsibility, classification and exchange protection may be enhanced.
- 'Compliance' in terms of cryptographic control and information system audit could have been added.

## eHealth Security Software Frameworks

The gap analysis of STORK 2.0 disclosed an overall adherence to the information security aspects according to the standards. Most gaps were related to what can be considered as local operations rather than the technology per se. However, the reviewed STORK 2.0 documentation, for instance, does not reference any controls against 'Protection against malicious and mobile code', 'Back-up' and to some extent 'Network security' and 'Media handling'. The STORK 2.0 project results are carried over to eIDAS, which is the Regulation 2014/910 on electronic identification and trust services for electronic transactions in the internal market. Furthermore, the review of DECIPHER showed several gaps. However, DECIPHER is a pre-commercial procurement (PCP) project, which to a large extent impacts its applicability and relevance in relation to a gap analysis of existing eHealth security software. In DECIPHER, the security aspects are not dealt with as the standards define, since the main focus of the project is on the PCP process itself. Finally, according to the OpenNCP project's gap analysis, some shortcomings in relation to 'Information security' have been identified: an audit trail that is potentially forgeable; no protection against malicious cloud provider or administrator; and only basic encryption technologies are used.

## Local End-User Organisations

A gap analysis of the local end-users was conducted through desk research, interviews with relevant people and expert opinions (e.g. IT staff in hospitals and healthcare professionals) and reveals that the local end-user hospitals in Spain, Denmark and Italy, respectively, all adhere to a high-level compliance with the background standards on information and health security. The gaps identified related mostly to whether or not there was full compliance to the ideal procedures according to the standards.

## National and European Cybersecurity Strategies and Reference Reports

The national cybersecurity strategies and ENISA reports address security at a general level given their strategic perspective, so many operational gaps were expectedly identified. However, for the most part, these gaps were not considered as significant. Nevertheless, it has been identified that there are aspects of the baseline template, which are not at all addressed. Furthermore, it should be noted that further elaboration regarding local operational application should be considered.

## Consolidation of Gap Analysis Outcomes

The gap analysis uncovered barriers and constraints as well as open issues, challenges and recommendations for information security in interoperable solutions at a systemic level. The consolidated outcomes are summarized below:

(a) *Barriers and constraints:*
- Adherence to the security targets and controls set by international standards are met to various degrees but rarely regardless of the analysis subjects.
- Technological advances happen at a pace which can quickly render security mechanisms applied outdated or not state-of-the-art.
- The activities of the projects and frameworks are developed in parallel but not integrated.

(b) *Open issues and challenges:*
- The level of adherence to standards on information security management varies between end-users (very high) to the frameworks, strategies etc.
- Analyzing frameworks does not address the details of the local operations and execution, hence some aspects are perhaps not captured.
- The dependency on adaption and implementation of specific technologies deriving from the frameworks impacts the KONFIDO solution.

(c) *Recommendations:*

Based on the full outcome of the gap analysis, five recommendations for addressing information and cybersecurity in an eHealth setting at a systemic and holistic level were

formulated, which are applicable to KONFIDO and its future activities as well as in a broader setting:

- Strive for high adherence to standards across all domains and subjects as it ensures trust and is in line with the end-users' approach.
- Take into account the users of a centralised technology/framework in terms of details of information security.
- Implement state-of-the-art security technologies and measures.
- Ensure information security sustainability in technical solutions.
- Explore further the implementation issues of the relevant security software frameworks.

## Conclusions

The conducted analysis has shown a range of gaps in other projects and initiatives, which does give indication of gaps at a systemic level and raises specific issues to consider. Recommendations derived from the gap analysis will be brought into the forthcoming phases of our project, but also have a wider European added value and, hence, will be disseminated appropriately. An iterative process for the gap analysis will strengthen the knowledge pool for KONFIDO and other relevant projects in the domain.

**Acknowledgements** The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727528 (KONFIDO—Secure and Trusted Paradigm for Interoperable eHealth Services). This paper reflects only the authors' views and the Commission is not liable for any use that may be made of the information contained therein.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. The Antilope project: https://www.antilope-project.eu/. Accessed 6 Oct 2017
2. The epSOS project: http://www.epsos.eu/. Accessed 6 Oct 2017
3. The JASeHN project: http://jasehn.eu/. Accessed 6 Oct 2017
4. The SemanticHealthNet project: http://www.semantichealthnet.eu/. Accessed 6 Oct 2017
5. The DECIPHER project: http://www.decipherpcp.eu/. Accessed 6 Oct 2017
6. The OpenNCP project: https://openncp.atlassian.net/wiki/. Accessed 6 Oct 2017
7. The STORK 2.0 project: https://www.eid-stork2.eu/. Accessed 6 Oct 2017
8. European union agency for network and information security, security and resilience in eHealth: security challenges and risks, 18 Dec 2015
9. European union agency for network and information security, cyber security and resilience for smart hospitals, 24 Nov 2016
10. The ISO 27 k family of standards: http://www.iso27001security.com/. Accessed 6 Oct 2017

# Identification of Barriers and Facilitators for eHealth Acceptance: The KONFIDO Study

P. Natsiavas, C. Kakalou, K. Votis, D. Tzovaras, N. Maglaveras, I. Komnios, and V. Koutkias

**Abstract**

In this paper, we present one of the key KONFIDO project's activities, the identification of key barriers and facilitators regarding eHealth solutions acceptance, focusing on security and interoperability. The methodology presented includes an end-user survey and an end-user workshop, engaging various stakeholders from Europe, in order to gain value out of their experience and insight in real-world healthcare settings. The analysis of the results provides a list of explicitly identified barriers and facilitators of adopting eHealth solutions in a Europe-wide scale, useful in the context of KONFIDO and beyond.

**Keywords**

eHealth acceptance • Barriers • Facilitators • Cross-border health data exchange • Security of eHealth systems

## Introduction

Recent advances in health IT intend to transform the healthcare delivery, especially via the increase of use of tele-monitoring solutions, mHealth applications and genomic data. However, the constantly increasing digitalization and use of sensitive data come along with the cost of proliferation of cyber-crime. For example, 2015 has been an all-time record year for security breaches in healthcare with over 100 million health records accessed by hackers globally [1]. Despite the benefits of technological advances, security is considered as one of the most important barriers for the large-scale adoption of new eHealth services. Lack of security results to patients and healthcare personnel unwillingness to share health data and adopt eHealth solutions, as well as of investors (both private and public) to fund such activities. In addition, during the last decade, we witness a considerable increase of citizen's mobility in Europe for education, training, work and tourism. Nevertheless, people suffering from chronic diseases are facing obstacles in travelling either within or outside their country of residence, due to the lack of an established, systematic and secure framework for data exchange among healthcare organizations across EU.

The KONFIDO project (http://konfido-project.eu/) aims to leverage proven tools and procedures as well as novel approaches and cutting-edge technology, such as homomorphic encryption and blockchains, to create a holistic paradigm for secure cross-border exchange, storage and overall handling of healthcare data. KONFIDO aspires to fulfil the prerequisites for cross-border patient mobility, in the interest of EU citizens, allowing secure cross-border exchange of personal health data. KONFIDO is organised in four complementary phases, namely, 'User requirements';

P. Natsiavas (✉) · C. Kakalou · V. Koutkias
Centre for Research & Technology Hellas, Institute of Applied Biosciences, 6th Km. Charilaou-Thermi Rd, 6036157001 Thermi, Greece
e-mail: pnatsiavas@certh.gr

K. Votis · D. Tzovaras
Centre for Research & Technology Hellas, Information Technologies Institute, Thermi, Greece

N. Maglaveras
Department of Electrical Engineering & Computer Science, McCormick School of Engineering & Applied Sciences, Northwestern University, Evanston, IL, USA

I. Komnios
Exus Software Ltd, London, UK



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_14

'Design'; 'Technology development'; and 'Integration, testing and validation'. As part of the 'User requirements' phase, KONFIDO reviews and maps applicable legal frameworks, ethical and social norms at EU level and in the project's pilot-site countries (i.e. Denmark, Italy and Spain), defining operational constraints and requirements. This process includes surveying all relevant stakeholders in participating countries (and other European countries too) to identify key factors and weak-signals that may considerably affect (at present and in the future) user acceptance, go-to-market strategies and overall operational sustainability of eHealth.

The current paper presents the methods employed to identify barriers and facilitators of eHealth acceptance linked with security and interoperability and the obtained results.

## Methods

The two pillars employed to identify key barriers and facilitators to adopt eHealth solutions linked with security and interoperability were (a) an End-user survey, and (b) an End-user Workshop. The scope of these pillars as well as their organization details are presented in the following.

### End-User Survey

The End-user survey focused on identifying the facilitators and barriers of applying security practices in real-world healthcare delivery settings. Thus, its main goal was to identify the currently applied practices regarding security and interoperability on existing e-health infrastructures, for organizations of varying size and nature (e.g. private and public), also focusing on specific stakeholder categories, namely managers, healthcare professionals (HCPs) and health IT stuff working in hospitals.

The survey has been implemented as an online survey, enabling sophisticated features like: (a) conditional workflow of questions based on answers submitted on earlier questions, so that only relevant questions appear for the user; (b) validation of input to avoid erroneous or malicious input; (c) export of the collected responses in a format convenient for further analysis, and (d) creation of personalized invitations and automatic reminders for the involved participants. The design of the survey is based on the guidelines presented in [2].

While making the survey public and using widely accepted forums, email lists, social media etc. would certainly increase the obtained responses, it would inevitably increase the risk of receiving answers of questionable value. Therefore, we decided to avoid a totally open invitation policy for survey submissions and invited specific individuals (experts) expecting high-value responses. The target audience has been carefully selected among stakeholders working on hospitals or health administrative regional units across Europe, able to provide the anticipated insights. A time window of one week has been given to the participants to submit their responses, while further extensions of this deadline have also been given. While the responses of the participants have been treated as anonymous, each invitation has been related with an automatically produced token to allow trace-back of the submitted responses for quality control reasons.

The survey questions have been structured in 6 sections:

- *Organization profile section*: refers to the organization's size and structure (e.g. employees number, activities' domain etc.) to provide a context for the responses.
- *Security facts section*: focuses on security incidents happened in the organization. This section targets technical staff (engineers and IT security staff) and managers.
- *Security policy section*: refers to policies applied in the organization (e.g. existence of security and risk management policies, use of encryption etc.).
- *Security incident management section*: targets on the handling of security breaches in a technical level. This section targets mostly technical staff (engineers and IT security staff) and managers as medical staff could not practically provide details on such issues.
- *Barriers and facilitators section*: aims on identifying key issues that facilitate or discourage the adoption of security oriented best practices.
- *Personal view section*: focuses on awareness (e.g. use of publicly available cloud storage services, importance of security in everyday work etc.) and satisfaction regarding current security status.

### End-User Workshop

The End-user Workshop has attracted more than 30 key stakeholders from the eHealth and healthcare across Europe. It has been organized to encourage open discussion, exploring the open issues in the domain of cross-border health data exchange through eHealth solutions. Personal invitations were sent to candidate participants from diverse organizations (healthcare, standards developing organizations, health IT associations, regional healthcare authorities, privacy authorities, research/academia, etc.), to obtain input from the widest possible spectrum of stakeholders composing the eHealth ecosystem. In each of the Workshop sessions, short presentations concerning key aspects of the project were provided, while sufficient time was assigned for discussion among participants. Discussions were recorded,

to allow transcription and elaboration upon the discussed issues.

## Results

### End-User Survey

#### Organization profile section

The end-user survey has been completed by 39 selected stakeholders across Europe. More than 50% of the participants refer to organizations with more than 1,000 employees and 80% of the submissions refer to organizations having more than one facility location. This is important as large organizations might tend to handle infrastructure issues in a more systematic way than small ones and this would probably lead in a more systematic approach on IT security issues, policies, etc., given that smaller organizations (e.g. peripheral hospitals) probably lack resources, expertise, etc. Furthermore, it should be also noted that participants' distribution among occupations was rather well distributed, keeping a balance among participants with a technical background (engineers and IT security staff), HCPs, and management stuff. A high number of participants is also leaning towards research (about 34%). Finally, the collected responses refer to a wide distribution of IT system types, with Electronic Health Records (EHR), Digital Prescription Records (DPR) and Laboratory Information Systems (LIS) being the most frequent.

#### Security Policy

The submitted responses clearly identify that security policies are widely applied. However, the lack of an overall security mentality is also identified. A clear majority of the survey participants (over 80%) answered that a specific IT security policy in their organization exists, while more than 40% identify standards or legislation on which their organization's policy refers to. 14.29% answered that there is no responsible person for IT security and more than 75% declared that encryption is used, at least to some extent. Personal data (clinical, demographic and personnel data) were recognized as more important than other operational data (ERP, CRM, email). However, only 14.29% of the survey participants declared that there is an incentive to discover and report security breaches, 27.59% knew of a specific information classification scheme used in their organization and 40% declared that there is no budget regarding IT security or it is lower than 1%, while 20% declared ignorance. Furthermore, none of the participants declared that there is breach insurance available (60% declared that there is no breach insurance and 40% declared ignorance). These findings clearly depict a lack of an everyday security-oriented mentality.

Regarding inter-organizations' data exchange, 68.57% explicitly declared that they regularly exchange data with other organizations and 40% exchange data with foreign organizations. Almost 80% act upon agreements with third-party organizations (e.g. other hospitals), to securely exchange sensitive data, while 35.71% declared that their state cross-border data exchange agreements are GDRP [3] compliant.

Technically, only 51.43% declared that there is a central antivirus management in their organizations, while 37.4% declared that there is a central IT resources access mechanism (Active Directory or LDAP). It should be noted that engineers, managers and HCPs have unlimited access to highly sensitive data, once they log in, in percentages reaching 40%.

These findings depict that KONFIDO should focus on raising awareness on real-world security policy issues.

#### Security Incident Management

Regarding the main cause of security breaches, "External attacks" is identified as leading cause and "Employee negligence" follows, while "Employee negligence" was characterized as the "most undetected" by 70% of the participants. Furthermore, organizations are ready to conduct risk assessment, but do not take actual measures to enforce security like breach monitoring and mitigation. The most frequently used security tool identified is VPN (70%), while there is little usage of advanced tools like Intrusion Detection Systems or Intrusion Prevention Systems (10% and 15%, respectively).

The above findings clearly depict that KONFIDO should also focus on raising awareness regarding more advanced security tools that are currently available. KONFIDO is expected to provide a Security Event Information Management (SIEM) system, that would also improve the respective organizations' incident management capabilities.

#### Barriers and Facilitators

Many survey participants referred to security measures as an obstacle for usability (34.29%). 75% declared that sometimes they skip a security rule and that their colleagues also try to skip security rules, either regularly or occasionally, implying a direct link between usability and security measures. Furthermore, almost 70% identified that lack of budget is a clear obstacle towards a more secure IT infrastructure, while almost 60% identifies the shortage of IT staff as barrier as well. Regarding the evolution of security measures and their efficiency, 57.14% feels that the overall security has improved, mostly due to the commitment of the management towards applying security practices. Only 5.71% declared that the situation has worsened, mostly due to the lack of management commitment. Conclusively, management commitment, budget increase and definition of

a specific policy are the main issues expected to facilitate the adoption of security-oriented best practices.

### Personal View

Network intrusions, malicious insider attacks, phishing and loosing storage devices were identified as the most important threats, according to the participants. 42.86% consider public cloud storage services unsafe to use and another 20% is not allowed to use them. Loss of productivity, fear for assets and organization's image are considered the most affected factors by a security incidence. More than 50% considers IT security as top priority, while only 30% appears to be satisfied with the level of IT security of their organization.

## End-User Workshop

The major outcomes of the End-user Workshop concerning barriers and facilitators include:

(a1) *Information flow barriers:*
- Cross-border health data exchange is typically manual and document-driven. The capability to "translate" the content is crucial and could be a major burden.
- Terminology issues can cause ambiguities.
- The diversity of the internal workflows applied in healthcare organizations. Applying IT-oriented approaches could lead to new information workflows that HCPs are reluctant to accept.
- Lack of trust among organizations (partially due to different information handling workflows), discourages (if not prohibits) data sharing.

(a2) *Information flow facilitators:*
- Using widely accepted, standardized encodings, thesauri, ontologies, etc. could facilitate interoperability and reduce ambiguities among different organizations due to the use of different languages and terminologies.
- Workflow heterogeneity among the healthcare organizations could be overcome through the definition of a simple common/baseline workflow (it could include only medical data sharing with no administrative data), acting as a proxy among the various local workflows.

(b1) *Legislation barriers:*
- National and European legislation formulates a complex grid of unaligned laws, hard to be interpreted and applied.
- Liability issues are not yet sufficiently clarified regarding data sharing scenarios.
- Ambiguities regarding data ownership.
- As technology evolves, new scenarios of data usage and transfer are emerging, causing legal gaps, as legislation cannot keep pace with fast evolving IT. Giving consent for health data handling is very important and could be proven as a major problem.

(b2) *Legislation facilitators:*
- GDRP and the 95/46/EC directive [4] provide a robust basis on which the EU Member States build agreements. Having a clear legal point of reference can significantly facilitate data process agreements among organizations.
- Several EU initiatives work on the alignment of legislation among EU Member States. Working groups, forums, etc. are constantly being formulated to facilitate interstate agreements and the application of EU directives in a uniform way across.
- A legal process that could facilitate data exchange and minimize legal problems is patient's explicit consent. Patient's consent could overcome the need of cross-institutional or international agreements. As patients tend to care most for their treatment than for their privacy, the process of consent must be carefully designed and consider proper information providing, opt-out or regret capabilities and delegation processes, e.g. for cases of patients being unconscious.

(c1) *Technical barriers:*
- Lack of a clearly established technical infrastructure for health data sharing.
- Usability shall be a top priority.
- Network availability has risen as a consideration in cases where high mobility is required.
- While standards exist for almost any procedure of data sharing, their application is a barrier itself.

(c2) *Technical facilitators:*
- Advanced technologies could be used to extract structure out of unstructured free-text data (e.g. Natural Language Processing) and encode the outcomes using widely-accepted, standardized coding schemes.
- Provenance information could be proven as a key facilitator for data validity. Through provenance, errors could be identified and possibly corrected.
- To overcome usability issues, the paradigm of a summary report document could be useful as a

paradigm that they the various stakeholders are familiar with.

- Several technologies and standards in the context of secure and interoperable cross-border health data exchange already exist, providing valuable tools and experience for the construction of real-world infrastructure.

## Conclusions

The main outcome of the End-user survey is that currently applied security practices are far from ideal due to several reasons, including lack of management commitment, lack of security culture in every-day activities but also shortage of available funding. Furthermore, in real-world settings, interoperability (cross-border or in-border) practices are not widely applied through standardized procedures, while the use of central, security-oriented tools (Active Directory, LDAP, Intrusion Detection Systems, etc.) is severely lacking. Moreover, management commitment to ensure the adoption of efficient security best practices must go further than just defining a policy and must practically enhance procedures through budget and specialized IT personnel. In the scope of the End-user Workshop, we identified many technical, organizational and legal barriers targeting cross-border health data exchange scenarios. Despite all the available technological achievements and the already evolving legal initiatives within EU, a lot of effort must be invested in aligning the legislations and the workflow of actions among EU Member States to provide the context of an IT solution facilitating cross-border secure data exchange and processing.

**Acknowledgements** The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727528 (KONFIDO—Secure and Trusted Paradigm for Interoperable eHealth Services). This paper reflects only the authors' views and the Commission is not liable for any use that may be made of the information contained therein. The authors would like to thank the KONFIDO partners for promoting the survey and for participating at the End-user Workshop.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. HIT Consultants: http://hitconsultant.net/2016/01/05/healthcare-cyber-attacks-in-2015-infographic. Accessed 6 Oct 2017
2. Shaughnessy J et al (2011) Research methods in psychology, 9th ed. McGraw Hill, New York, NY, pp 161–175
3. General data protection regulation portal. http://www.eugdpr.org/. Accessed 6 Oct 2017
4. Directive 95/46/EC of the European Parliament and of the Council of 24 Oct 1995

# A Personalized Cloud-Based Platform for AAL Support to Cognitively Impaired Elderly People

Stefanos Stavrotheodoros, Nikolaos Kaklanis, and Dimitrios Tzovaras

**Abstract**

Population ageing due to declining fertility rates and/or rising life expectancy is poised to significantly transform our societies in the upcoming years. Although a lot of work has been done in the field of AAL for the creation of ICT solutions that will prolong and support the autonomous living of elderly individuals with cognitive impairments, these solutions are unable to meet all user needs and/or to provide an easy to use mechanism for further extension with new services. The present paper presents a cloud-based solution that provides easy, transparent, personalized and contextualized access to all the supported AAL services to the cognitively impaired elderly end-users and their caregivers by also offering a mechanism for easy registration and integration of new AAL services into the platform.

**Keywords**

AAL services • Ontology • Matchmaking • Cloud based architecture

## Introduction

According to the World Report on Ageing and Health [1] released by the World Health Organization (WHO), today the vast majority of people can expect to live into their 60 s and beyond. Cross-sectional comparisons have consistently revealed that increased age is associated with lower levels of cognitive performance [2]. Cognitive loss has an important impact on the capacity to conduct activities of daily living (ADL) in older people, resulting in dependency, distress, and reduced quality of life [3].

The basic infrastructure that supports people with cognitive impairments, such as homes for the elderly, nursing homes, and other care facilities, are becoming insufficient to deal with this increase, thus there is a need to address their unmet needs through the use of ICT [4]. According to Lauriks et al. [5], the elderly have various needs that can be summarized in the following categories: (a) need for general and personalized information, (b) need for support with regard to cognitive decline symptoms, (c) need for social contact and company, and (d) need for health monitoring and perceived safety.

Although, several ICT applications and services have been developed to support people with cognitive impairments, a major challenge is to provide a holistic solution that addresses all the aforementioned needs by also supporting an easy to integrate mechanism of new Ambient Assisted Living (AAL) services.

This paper presents the IN LIFE cloud-based platform developed in the context of the IN LIFE H2020 EU project that aims to lengthen and support the independent living of elderly individuals with cognitive impairments, through interoperable, open, personalized and seamless ICT solutions. The main goal of the platform is (1) to provide personalized and easy access to all the supported AAL services to the elderly cognitively impaired users and their caregivers, and (2) to ensure that external service providers and device manufacturers will be able to register their assets in a user-transparent, open and standards-abiding way.

S. Stavrotheodoros (✉) · N. Kaklanis · D. Tzovaras
Information Technologies Institute, Centre for Research and Technology, Hellas 6th Klm. Charilaou - Thermi Road, P.O. BOX 60361 GR - 570 01 Thessaloniki, Greece
e-mail: stavrotheodoros@iti.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_15

## Related Work

Existing solutions in the AAL domain can be classified into two main categories: (1) on-premises solutions and (2) cloud-based solutions. Some of the on-premises solutions focus on the establishment of reference architectures for AAL systems, such as the PERSONA project [6], which aimed at the development of a scalable, open-standard technological platform for building a range of AAL services. universAAL [7] reuses many components of PERSONA and it supports non-cloud features by also providing runtime support for software components and services for different types of hardware device. OASIS [8] introduced an ontology-driven, open reference architecture and platform that facilitates interoperability, seamless connectivity, and sharing of content between different services, while SOPRANO [9] introduced an ontology-centered platform for offering AAL solutions through an extensible service-oriented OSGi modular architecture.

All the aforementioned projects are considered to be the most consolidated and well-known AAL platforms, but all of them are based on non-cloud based architectures. However, there are some other AAL solutions that utilize cloud computing features, like the DOMEO project [8], which uses integrated cloud-services for personalized homecare services for tele-presence. The iWalkActive project [10] provides to users with walking disabilities cloud services that make use of indoor and outdoor navigation. The MyLifeMyWay [11] project provides through a cloud platform a Personal Virtual Assistant that can integrate existing and proven technology and home automation appliances. The iCarer project [12] provides a personalized and adaptive cloud-based platform to offer informal carers support by means of monitoring activities of daily care, while the SOCIALIZE project [13] introduced a service-oriented software architecture to supply network services with cloud computing modalities in order to promote elderly social interaction.

Although the aforementioned solutions utilize some cloud computing features, they are not entirely based on a cloud-based architecture for the provision of various types of AAL service. Additionally, some of them try to meet specific user needs like mobility or communication without following a more holistic approach. Moreover, they do not utilize semantic technologies.

The solution presented in this paper goes one step beyond the state of the art by presenting a unified cloud-based framework, enhanced with semantic technologies, based on existing reference architectures, further extended in order to provide advanced functionalities for transparent, personalized and contextualized access to all the supported services as well as easy service registration/integration of new services. The adoption of a cloud-based approach ensures the accessibility of the registered tools/services on demand from everywhere, while the ontology-driven architecture enhances the semantic interoperability between the architectural elements of the platform and takes advantage of well-established health describing ontological frameworks, like ICF [14].

## In Life Platform

The main functionalities supported by the IN LIFE platform are the following:

(a) Monitor user activities and preferences in an unobtrusive way.
(b) Support elderly people with cognitive impairments in a variety of indoor and outdoor activities by providing easy and personalized access to the IN LIFE services and applications.
(c) Provide help and instructions to care givers.
(d) Enable service/application providers to easily incorporate their products in the IN LIFE framework.

In order to support them a cloud infrastructure was set, where the heart of the system is an OSGi/universAAL-based platform [7]. A knowledge base consisting of three ontologies (i.e. User Ontology, Service Ontology and Sensor Ontology) described in detail by Konstantinidou et al. [15] is used for storing information regarding the registered users, the supported AAL services, as well as data coming from monitoring in a semantic manner. Figure 1 illustrates an overview of the platform's architecture, while its main functionalities are presented in detail in the following paragraphs.

### User Clustering and Modelling

Different kind of users have different needs, should be the target of different interventions, and possibly of more granular monitoring indicators. Using the user's capacity to function in terms of daily life activities or of disability-free status, and his/her Socio-Economic Status (SES), the personal information of the cognitively impaired individuals is used by the platform's matchmaker in order to cluster the user according to a user taxonomy. The 4 supported taxonomies (i.e. dependent, assisted, at risk and active) represent 4 archetypes of users with different needs [15].

For the described functionality, the information stored in User Ontology is used. The User Ontology defines in a

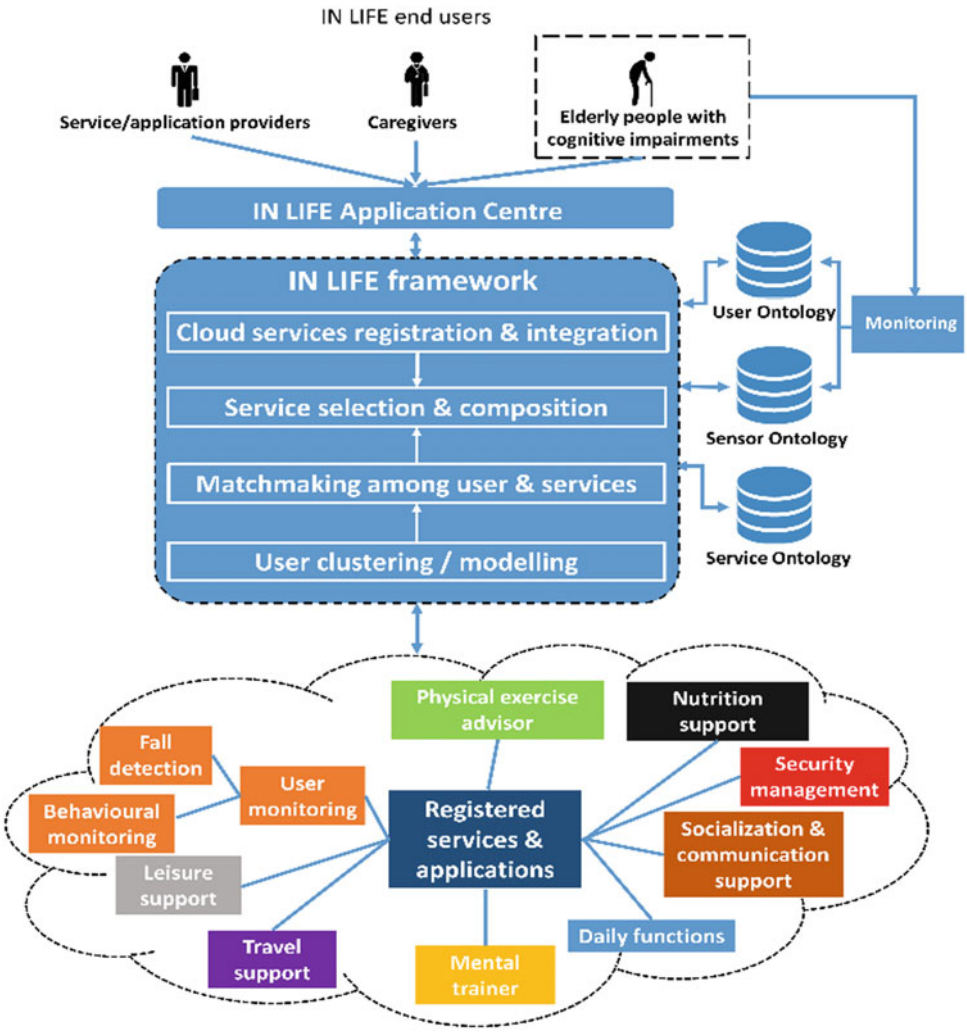



**Fig. 1** INLIFE architecture

semantic way all users of the IN LIFE Platform and more specifically assisted people, their informal and formal carers and service providers. It is based on the ontology of the ACCESSIBLE project [16] that was based on ICF [14], thus making it suitable for describing health and health-related states, and the Profiling Ontology of universAAL [17], which models AAL user profiles and their characteristics.

## Matchmaking - Service Selection and Composition

Towards providing personalized services to the elderly end-users and resolving emergency cases (e.g. the user has fallen), a hybrid matchmaking mechanism has been developed that consolidates the results of a rule-based matchmaker and a statistical matchmaker.

The rule-based matchmaker, in case there is an alarm notification that indicates an emergency for the user, through a set of rules selects the most proper set of services to be called. By applying the *Exact* and *Plug in* OWLS-MX filters [18] in the selected services, it identifies which of them can be combined and called in a sequential way. Moreover, another set of rules selects the most suitable tool categories and tools for the user to be displayed by using his/her current activity and status stored in the User Ontology and the semantic description of the tools stored in the Service Ontology.

The statistical matchmaker tries to improve the recommendation accuracy of the rule-based and recommends tools for the user by creating clusters of likeminded user. Two different clustering approaches are used: a collaboration filtering method that utilizes the tool usages of the users and a demographic filtering method that uses certain personal

attributes like socio-economic status, country, education level, etc.

The tool recommendations of both matchmakers are combined into a hybrid matchmaker that uses user-specific weights in order to combine the results of all techniques into a single recommendation value. The weights are dynamically adjusted in order to optimize the predictions of the system.

## Application Centre

The Application Centre is the main interface of the IN LIFE platform and through it cognitively impaired individuals and caregivers can access the AAL services registered in the IN LIFE ecosystem or browse/search/connect with other users and service providers can register new services/applications.

A personalized dashboard with different functionalities for each role is supported. In addition, according to the current activity and status of the user, and by using the matchmaking capabilities of the system, the most suitable set of services is proposed and displayed.

An Alarm Management mechanism has been implemented in order to manage effectively alarm notifications created by AAL tools and services that are integrated into the IN LIFE ecosystem. Each received alarm notification is stored in the Sensor Ontology and it is instantly displayed through the Application Centre to all caregivers connected with the elderly.

An accessible and user-friendly interface, appropriate for cognitive impaired elderly, is provided, based on the guidelines and techniques described mainly in the WCAG 2.0. All pages of the Application Center website have checked by two online web accessibility evaluation tools: the A Checker [19] that resulted a WAI "AAA" classification and the checker of the European Internet Inclusive Initiative project [20] that resulted a 94.015% average compliance.

Finally, due to the multilingual environment of Europe, the following 9 languages are supported: English, Greek, Spanish, Dutch, Slovenian, Swedish, German, French and Italian. The Application Centre, in the context of the IN LIFE project, was tested, evaluated and further refined by large-scale, Europe-wide pilots conducted in Greece, the Netherlands, Slovenia, Spain, Sweden and UK.

## Cloud Services Registration and Integration

Application/ service providers can easily register their AAL solutions in a semantic way and also integrate them on the cloud platform. The registration is made through a set of web forms provided by the Application Centre. The semantic description of the services is stored in the Service Ontology, a core component of the IN LIFE platform, that is based on the universAAL Profiling Ontology [17] and supports also the OWL-S standard [21] for the proper technical description of a service.

This ontology was further extended in order to include the technical description of services not only from a syntactic point of view, but also in terms of semantic capabilities in order to indicate different service features and the way to interact with them. Semantic capabilities have been taken into account by specifying a set of terms/classes, to which each service is mapped, creating this way a common dictionary that ensures the use of the same terms when describing same things, thus enabling the use of generic rules by the matchmaker for mapping services and users according to their status by also identifying which of these services can be used in a combined matter.

## Conclusions

In the present paper, the IN LIFE platform, a holistic cloud-based ICT solution for supporting the independent living of elderly individuals, was presented. Through its innovative hybrid matchmaking approach it provides personalized solutions to the elderly end-users and resolves emergency cases by selecting and combining the most proper services. Through Application Centre it provides accessible, personalized dashboards that elderlies can use to access the most suitable tools, caregivers can effectively monitor their connected cognitively impaired individuals and service providers can register their assets, building this way a distributed ecosystem of services targeting elderly.

**Acknowledgements** This work is supported by the EU funded project IN LIFE (H2020–643442).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. World Health Organization (2015) World report on ageing and health. World Health Organization, Geneva, Switzerland
2. Salthouse TA (2009) When does age-related cognitive decline begin? Neurobiol Aging 30:507–514
3. Mograbi DC, de Assis Faria DC, Fichman HC, Paradela EM, Lourenco RA (2014) Relationship between activities of daily living and cognitive ability in a sample of older adults with heterogeneous educational level. Ann Indian Acad Neurol 17 (1):71–76
4. Genet N, Boerma W, Kroneman M, Hutchinson A, Saltman RB (2012) Home care across europe. current structure and future challenges. observatory studies series 27. In: The European observatory on health systems and policies

5. Lauriks S, Reinersmann A, Van der Roest HG, Meiland FJ, Davies RJ, Moelaert F et al (2007) Review of ICT-based services for identified unmet needs in people with dementia. Ageing Res Rev 6(3):223–246
6. Perceptive spaces promoting independent aging-PERSONA (FP6-IST-045459)
7. Hanke S, Mayer C, Hoeftberger O, Boos H, Wichert R, Tazari MR, Wolf P, Furfari F (2011) UniversAAL—An open and consolidated AAL platform. In: Ambient assisted living, pp 127–140
8. Fazekas G, Zsiga K, Pilissy T, Tóth A, Rumeau P, Dupourque V, Dénes Z (2012) Field test of a home-care robot for elderly assistance: first experiences. In: 9th congress of the mediterranean rehabilitation forum
9. Muller S., Sixsmith AJ. (2008) User requirements for ambient assisted living: results of the SOPRANO project, Gerotechnology, 7(2), 168.ev, 6 (October (3)) (2007):223–246
10. An active rollator for active people—iWalkActive (AAL). www.iwalkactive.eu/. Accessed 10 Apr 2017
11. MyLifeMyWay project (AAL). http://www.mylifemyway-aal.eu/. Accessed 10 Apr 2017
12. Intelligent care guidance and learning services platform—iCarer (ALL). http://icarer-project.eu/. Accessed 10 Apr 2017
13. SOCIALIZE project (AAL) at www.aal-europe.eu/projects/socialize/. Accessed 10 Apr 2017
14. World Health Organization (2001) International classification of functioning, disability and health: ICF. World Health Organization, Geneva, Switzerland
15. Konstantinidou A, Kaklanis N, Tzovaras D (2016) A unified cloud-based framework for AAL services provision to elderly with cognitive impairments. In: 7th international conference on cognitive infocommunications, wroclaw, Poland
16. Lopes R, Votis K, Carriço L, Likothanassis S (2009) A service oriented ontological framework for the semantic validation of web accessibility. In: Cruz-Cunha M, Oliveira E, Tavares A, Ferreira L (eds) Handbook of research on social dimensions of semantic technologies and web services. IGI Global, Hershey, pp 49–67
17. UniversAAL Profiling Ontology at https://github.com/universAAL/ontology/wiki/Profile. Accessed 10 Apr 2017
18. Klusch M, Fries B, Khalid M, Sycara K (2009) OWLS-MX: hybrid OWLS service matchmaking. Web Semant Sci Serv Ag World Wide Web 7:121–133
19. Gay G, Li QC (2010) A Checker: open, interactive, customizable, web accessibility checking. In: Proceedings of the international cross-disciplinary conference on web accessibility (W4A2010), Raleigh, USA, April 2010. http://dx.doi.org/10.1145/1805986.1806019
20. European internet inclusion initiative at http://checkers.eiii.eu/. Accessed 10 Apr 2017
21. OWL-S: semantic markup for web services at https://www.w3.org/Submission/OWL-S/

# Enhanced Healthcare System Based on Mobile Communication

Cheng-Huei Yang, Tsung-Che Wu, and Hsiu-Chen Huang

**Abstract**

This paper aims to develop an enhanced healthcare system that can not only increase the rate of medication adherence through multimedia technology, but also improve the efficacy of diagnosis and treatment. The system comprises five parts including a medicine bag, a home server, a hospital server, a prescription terminal, and a pharmacy sever. The wireless communication technology performs data communication among different parts in the system, which achieves remote monitoring and rapid efficacy assessment. The presented system is intended for patients, physicians, and caregivers, to enable not only medication prescribing and treatment efficacy reporting during illness, but also complete health record collection before illness. The experimental results show that the proposed system can effectively achieve a better healthcare and fast recovery from illness.

**Keywords**

Healthcare • Medication adherence • Treatment efficacy • Mobile communication

## Introduction

Health is priceless. It influences not just physical abilities, but mental and social aspects as well. Better healthcare including completed health examination records before illness, strict adherence to medication, efficient treatment efficacy reports during illness, and sufficient nutrition supplement after illness are the best ways to maintain health condition for long time. Although illness is unavoidable, fast recovery is determined by appropriate medical care.

Previous research regarding medication care includes many varieties such as drug dispensers, pill wallets, pill organizers, medication, and so on. Some studies tried to integrate different features to achieve a better healthcare goal. Gors [1] used a mobile device to enhance the pill dispensing function including patient identification and medicine in-take recognition. These products provided a lot of help and convenience while sick. Since medication non-adherence results in various degrees of impairment, Hayes [2] developed electronic devices with different features for monitoring medication adherence so that patients can take medicine as prescribed. Regarding health record information technology, Pires [3] introduced an integrated e-healthcare system that can monitor and record the biomedical parameters of a person in real time via wireless communication without interfering in his/her daily life to achieve a better diagnosis.

The proposed system mainly consists of five parts: a medicine bag, a home server, a hospital server, a prescription terminal and a pharmacy sever. Several radio frequency (RF) communication modules are embedded to support wireless communication among the system components. The Internet is also linked to perform data transmission to support treatment evaluation and short message service when help is needed. For the most flexibility, the system provides

C.-H. Yang (✉) · T.-C. Wu
Department of Electronic Communication Engineering, National Kaohsiung Marine University, Kaohsiung, Taiwan
e-mail: tosos.chyang@msa.hinet.net

H.-C. Huang
Department of Physical Medicine and Rehabilitation, Chiayi Christian Hospital, Chiayi, Taiwan



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_16

two operation modes including stand alone and online mode. In the stand alone modes, the patient can independently arrange medication events such that they can enter the name, quantity, time, and picture of medicine. In the on-line mode, all medication events start with sending the patient's health record to the doctor's prescription server and then receiving prescription information from the doctor after completing diagnosis. Later, having prescription medicine, the system initializes medication reminders and records the medication time. Furthermore, the system automatically starts a medication efficacy investigation procedure which is used to inquire about the patient's response to the medicine after taking the medication for a period of time. The efficacy report is then sent to the medical doctors or caregivers for further processing.

## Methods

Complete health data collection including illness before and after the treatment period is the most important step for fast recovery of heath. The health data before illness not only are good for individuals (users) to know their health condition, but also help medical doctors' diagnose precisely. Similarly, the health data collected during the treatment period aim to record adherence to medication treatment and, more importantly, report treatment efficacy, which is also used as an index for whether current medication is appropriate or not. To complete health data collection, we developed a health care system that can not only collect health data completely in different periods but also report patient experience of medication efficacy rapidly. Each home server may communicate with several family members' medicine bags, and every prescription server may also communicate with many patients' medicine bags, as shown in Fig. 1.

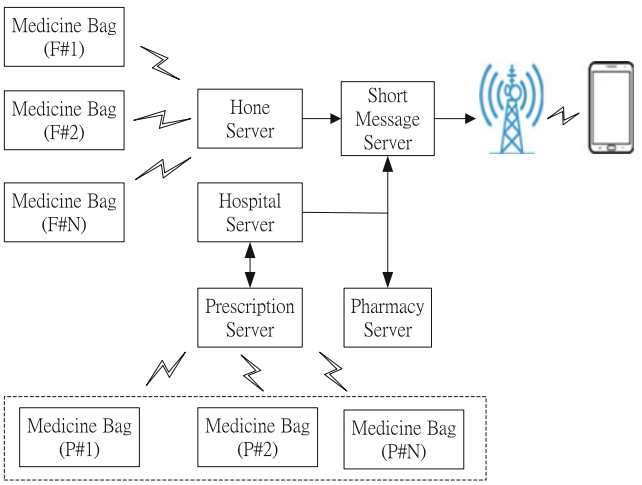



**Fig. 1** System architecture

### Family Health Data Managing

Many diseases have been reported to be related to family inheritance. Gathering family members' illness history from the treatment record of their medicine bags and then performing statistical analysis will be helpful in finding family genetic diseases. The home server is a personal-based computer used to store all family member health information. A relational database management system is built to manage family members' health data, and an analytical mechanism is designed to perform various analyses to conclude genetic disease findings. A disease warning and emergency rescue mechanism is also constructed to allow family members to receive warning messages and help in time when necessary. The short message server handles emergency message delivery issued by the home server through the base station and the cellular phone to obtain immediate help. Family members use this bag to report their health examination records while they are away from home. Once they return home, these health records are sent to the home server; thus, the complete health information of the family is then constructed.

### Diagnosis and Prescription

The correct diagnosis depends on the detailed disease symptom description, and the treatment efficacy depends on the precise prescription. During outpatient clinic, the doctor uses the prescription server to receive electronic health data from the patient's medicine bag via wireless communication. With these data, the doctor can have much more information to know patients' illness situation and, therefore, reach a better diagnosis and then prescribe the most appropriate treatment. The electronic prescription is sent back to the patient's medicine bag through wireless communication with other medication information such as drug picture, indications, and precautions. In addition, this information is also sent to the hospital server for later inquiry and to the hospital's pharmacy for medicine preparation. The hospital server is used to save all information including diagnostic, medication, genetic disease, and other related treatment information.

### Prescription Verification and Medicine Preparation

Having prescription medicine as early as possible will speed up treatment and recovery. The proposed system provides two alternatives for getting prescription drugs. Firstly, the pharmacy server inside the hospital receives prescription information from the prescription server via the hospital

server. Secondly, if there is no pharmacy server inside the hospital, the patient has to find a drug store somewhere outside the hospital. The patient either gives his/her prescription to the pharmacist or sends his/her electronic prescription via wireless communication to the pharmacy server of the drug store, and then the pharmacist performs prescription information verification by delivering the prescription data to the hospital server through the Internet. Once verified, the pharmacist starts to prepare and dispense medication.

## Medication Adherence Monitoring and Treatment Efficacy Reporting

The medicine bag is used to store medicine and initialize medication reminders and then record medication. Later, it performs a treatment efficacy evaluation to collect medication efficacy information immediately, thus, ineffective medical care can be reduced. The medicine bag is a microprocessor-based bag connected with control kernel and other peripherals. The control kernel is a data management center, which is used to receive the prescription information directly from the doctor's prescription server and sense (close and open) signals of the medicine storage compartment via a micro control unit (MCU). The control kernel also performs medication reminders including switching on the indicator light and generating an appropriate voice message, and then it records the date and time automatically when the user completes medication.

## Experimental Results and Discussions

Healthcare is the maintenance of health including disease prevention, diagnosis, and treatment. To know one's body situation through health examination is the best way to prevent disease. Complete and detailed health checkup information is very helpful in assessing and predicting the dangers of disease both for physicians and individuals. To achieve multi-function, a portable medicine bag was developed. This bag includes four layers for medicine and two extra layers for storage and four indication lights used to guide correct medicine use. Moreover, four different color drug boxes are used for different medicine periods. The medicine bag is designed not only to store daily physical checkup information but also prescription medicine. With its multi-layer design, the bag can be used to store either different medicines received from different hospitals or different medicines based on different medication periods from the same hospital. Moreover, it performs medication monitoring


**Fig. 2** Outlook of the medicine bag

starting from the medication time setting, which can be received either from the prescription server or input through self-management. The medicine bag will inform the user to open the layer to take medicine at the medication-taking time. The prescription medication pictures as shown in Fig. 2 and other information are automatically displayed on the screen. Furthermore, the indication light of the corresponding layer of the medicine bag also flashes on to avoid taking the wrong medicine.

Ineffective treatment caused by inappropriate medication not only certainly delays recovery, but also might hurt patients' health. Thus, immediately reporting treatment efficacy is very important. The proposed system designed an interactive reporting mechanism initiated within a period after taking medication. The form page for reporting is automatically displayed. Three efficacy states including better, no improvement, and worse can be chosen. The patient can simply press the appropriate button through the touch-screen to report their feeling about treatment efficacy. The reporting data are stored and compared with the control parameter programmed previously to determine whether to start up a remedial action such as informing caregivers, medical doctors, and so on. With the display, people including medical doctors, caregivers, and patients themselves can easily understand the medication adherence situation. This information is also sent to the medical doctor during the next outpatient clinic through wireless communication.

## Conclusions

Complete health care should generally include care for illness before, during, and after treatment. The proposed system presented is a multi-function system that can achieve three periods of care using the same device. Furthermore, the treatment efficacy reporting mechanism allows medical staffs

or caregivers to know the patient's situation immediately, thus, give an appropriate treatment and prescription. Undoubtedly, better treatment relies on better diagnosis, and better diagnosis is determined by disease symptom information provided by the patient. In other words, complete symptom information will influence disease treatment and, later, patient's health recovery.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Gors M, Albert M, Schwedhelm K, Herrmann C, Schilling K (2016) Design of an advanced telemedicine system for remote supervision. IEEE Syst J 10(3):1089–1097
2. Hayes TL, Hunt JM, Adami A, Kaye JA (2006) An electronic pillbox for continuous monitoring of medication adherence. In: Conference Proceedings on IEEE Engineering in Medicine and Biology Society
3. Pires P, Mendes L, Mendes J, Rodrigues R, Pereira A (2016) Integrated e-healthcare system for elderly support. Cogn Comput 8 (2):368–384

# Experience of Using the WELCOME Remote Monitoring System on Patients with COPD and Comorbidities

E. Kaimakamis, E. Perantoni, E. Serasli, V. Kilintzis, I. Chouvarda, R. Kayyali, S. Nabhani-Gebara, J. Chang, R. Siva, R. Hibbert, N. Philips, D. Karamitros, A. Raptopoulos, I. Frerichs, J. Wacker, and N. Maglaveras

**Abstract**

The WELCOME system is an innovative telemonitoring system designed to provide constant monitoring of COPD patients also suffering from other major comorbidities. It consists of a sensors vest capable of recording various vital parameters in real time and transmitting them to the cloud via a tablet PC and wireless connection. In addition, peripheral devices record extra physiological data which is coupled with responses to validated health questionnaires that the patient responds to via the tablet. A dedicated medical decision support system (DSS) and a medical professional user interface support the system in providing automated detection of abnormal conditions. The obtained signals included Heart and Respiratory Rate, Body Posture, $SpO_2$, multi-lead ECG, Auscultation and Electric Impedance Tomography. The system is tested with pilot studies in two European countries, Greece and UK. The preliminary results from the Greek pilot study are presented in this paper, highlighting the main findings from the first operational use of the WELCOME infrastructure.

**Keywords**

Comorbidities • COPD • Decision support system • Pilot study • Telemonitoring system

E. Kaimakamis (✉) · E. Perantoni · V. Kilintzis · I. Chouvarda · N. Maglaveras
Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Ag. Dimitriou street, Thessaloniki, Greece
e-mail: vkaimak@med.auth.gr

E. Serasli
Pulmonary Department, General Hospital "G. Papanikolaou", Thessaloniki, Greece

R. Kayyali · S. Nabhani-Gebara · R. Hibbert · N. Philips
Faculty of Science, Engineering and Computing, Kingston University, Kingston-upon-Thames, UK

J. Chang · R. Siva
Chest Clinic and Research and Development, Croydon University Hospital, Croydon, UK

D. Karamitros · A. Raptopoulos
EXUS Company, Athens, Greece

I. Frerichs
Department of Anesthesiology and Intensive Care Medicine, University Medical Centre Schleswig-Holstein, Kiel, Germany

J. Wacker
CSEM Company, Neuchatel, Switzerland

## Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a preventable, progressive respiratory condition characterized by airflow reduction and persistent respiratory symptoms [1]. It is characterized by varying degrees of airway and alveolar damage and represents a major health problem, generating considerable premature mortality and economic loss [2]. It is often associated with other comorbid conditions, like anxiety, chronic heart failure, depression, and diabetes mellitus [3]. The patients present variable rates of disease exacerbations, potentially life threatening and often leading to hospital admissions. The complexity of the disease its complications and its combination with comorbidities necessitate a holistic approach to its monitoring and management [4].

Many efforts have been made to create novel integrated care schemas in order to better address the multimorbid condition of COPD and other diseases [5, 6], as it is evident that such interventions may have the ability to reduce the utilization of



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_17

healthcare provisions and have an impact on reduction of at least short term associated costs [7]. In addition, when these interventions utilize technology in the integrated care process, there is evidence of stabilization of important clinical variables, even two years after the initial intervention [8].

The European project WELCOME (Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Comorbidities, Grant Agreement FP7-611223) is an innovative European funded project aiming to provide integrated care for the management of COPD with associated comorbidities. The project has produced a novel telehealth solution, the WELCOME system, able to provide continuous monitoring and detection of the disease variability on a daily basis.

The monitoring function is coupled with a medical Decision Support System (DSS) that takes data from a series of wearable and portable sensors, providing information on the patient's health status and potential deterioration of the disease state.

The system is being tested with pilot studies in two European countries, namely Greece and UK. A preliminary trial with healthy volunteers preceded the pilot trials and took place in Germany. The first results from the Greek pilot study are presented in this paper, highlighting the main findings from the first operational use of the WELCOME infrastructure.

## Methods

### System Description

The array of sensors deployed to gather the biophysiological signals and present them to the DSS includes a sensor-vest that records body posture, heart and respiratory rate, $SpO_2$, multi-lead ECG, Lung Auscultation and Electric Impedance Tomography (EIT) imaging [9]. Figure 1 shows the approximate position of the embedded sensors of the vest on the patient's skin. While the vest is worn the patient may move freely within the Bluetooth connectivity range from the tablet. A number of peripheral medical devices also perform measurements and transmit them to the tablet PC via Bluetooth. The measured parameters are blood pressure, body weight, body temperature, blood glucose and spirometry. The WELCOME solution is designed to identify parameters that indicate when unscheduled clinical interventions might be required. The patients periodically respond to validated health questionnaires (addressing the full range of symptoms and signs of the underlying diseases) via the tablet. A dedicated medical decision support system (DSS) and a medical professional user interface support the system in providing automated detection of abnormal conditions. The DSS takes into consideration the obtained biosignals from the WELCOME system and follows a set of complex rules which analyze the possibility of a disease progression, based on the combination of findings and the existing comorbidities in each patient. Personalized average values are constantly compared to the obtained values from the continuous measurements to ensure close monitoring of patient condition and refinement of the rules. The expert set of rules has the form: *IF* patient has COPD *AND* Heart Failure *AND* lower $SpO_2$ than personal baseline *AND* crackles in auscultation *AND* increased dyspnea *AND* Body weight increase from personal baseline *AND* swollen ankles *AND* bilateral decreased ventilation in EIT signal, *THEN*: Possible pulmonary edema alert generated.

The system was designed based on a series of interviews from patients, healthcare professionals and formal or informal carers [10]. The main technologies used were developed in a multinational level and various parallel studies helped to evolve and validate associated algorithms and solutions for automated signal analysis and annotation [11].

### Data Collection Process

Patients were recruited at the outpatient clinic of the Pulmonary Department at "G. Papanikolaou" General Hospital of Thessaloniki and were informed about the study design and objectives. All of them signed an informed consent form and the study protocol was approved by the ethics committee of the Hospital. The recruited patients were examined at the outpatient setting and then properly trained to use the full set of equipment for the collection and transmission of data at their home environment. The number of recruited patients reached the level of 17, with 14 male and 3 female subjects, with an average age of 68.4 years and COPD GOLD stages II to IV. Regarding the comorbidities, there were four patients with comorbid Diabetes Mellitus, three with Congestive Heart Failure and two with Anxiety/Depression.

Two patients dropped out of the study after its initiation, either due to perceived inability to operate complex technology or due to delays because of technical problems encountered during the trial. Another two patients were recruited during hospitalization due to COPD exacerbation and were recorded during their stay at the hospital.

### Experimental Protocol

The study protocol dictated that the patients had to log into the WELCOME system using the tablet PC with their personal credentials and then perform the vital signs measurements with the peripheral devices, sending the recorded data to the cloud via the tablet. The next step was to wear the vest for 30 min twice a day and connect the vest to the tablet wirelessly. The system transmitted the recorded data to the

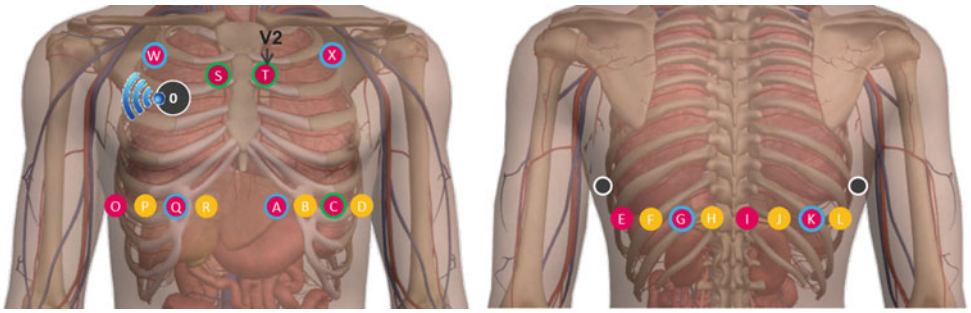



**Fig. 1** Chest locations of the embedded sensors of the WELCOME vest

cloud every 5 min. After the termination of the recording, the vest should be put on charge and start the automatic transmission of the largest signals to the tablet via Wi-Fi connection. The final step was for the patients to answer disease-specific validated questionnaires through their tablet regarding their daily symptoms, the control over their disease and the quality of life.

After the completion of the study, the subjects filled the TAM3 and SUTAC questionnaires that are designed to capture the experience they had regarding the use of technology and how this intervention affected their health status and healthcare utilization.

The patient hub interface used by the subjects was designed for the tablet PC and allowed the users to perform the transmission of data and communicate with the doctors or view educational content assigned to them by their doctors. This interface was simplified and intuitive for use by elderly technology-naïve patients (Fig. 2).

## Results from the Greek Pilot Study

### Preliminary Results

In total, more than 1250 chunks of 5-min recordings were successfully transmitted to the cloud and included all of the aforementioned bioparameters collected from every patient. The big data volume requires an analytic approach for the evaluation of the obtained signals and their contribution to a better understanding of the disease progression concerning the daily fluctuations in real life and potential progression towards exacerbation events. The specially designed medical decision support system fired a set of alerts in numerous cases when the biosignals exceeded predefined thresholds or thresholds set by the individual variability and was focused on the combined management of COPD and its major comorbidities, thus patient-centered and personalized. There were 241 simple rules that were fired together with 93

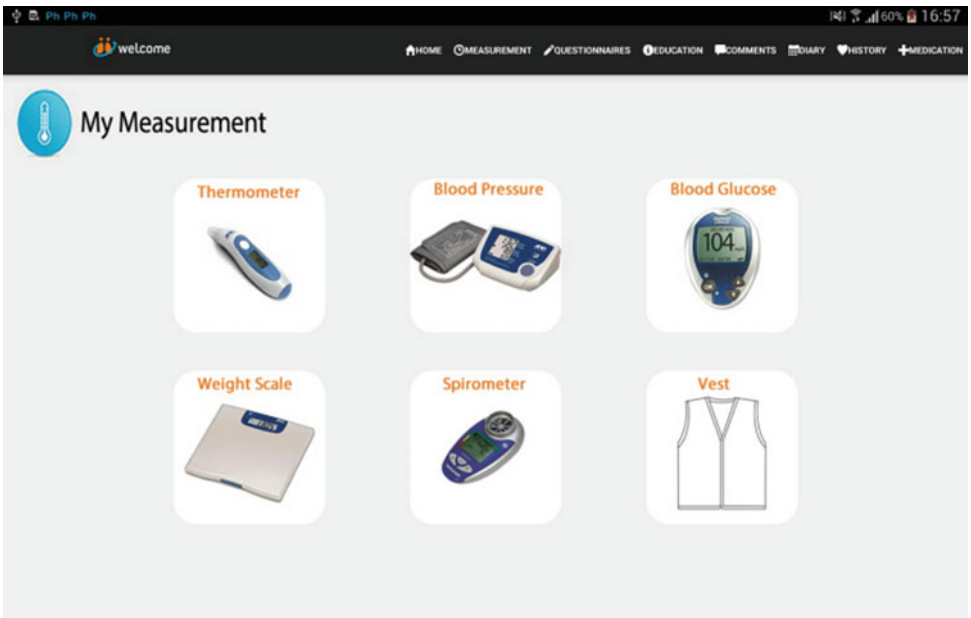



**Fig. 2** Screenshot of the user interface of the patient hub software on the tablet screen

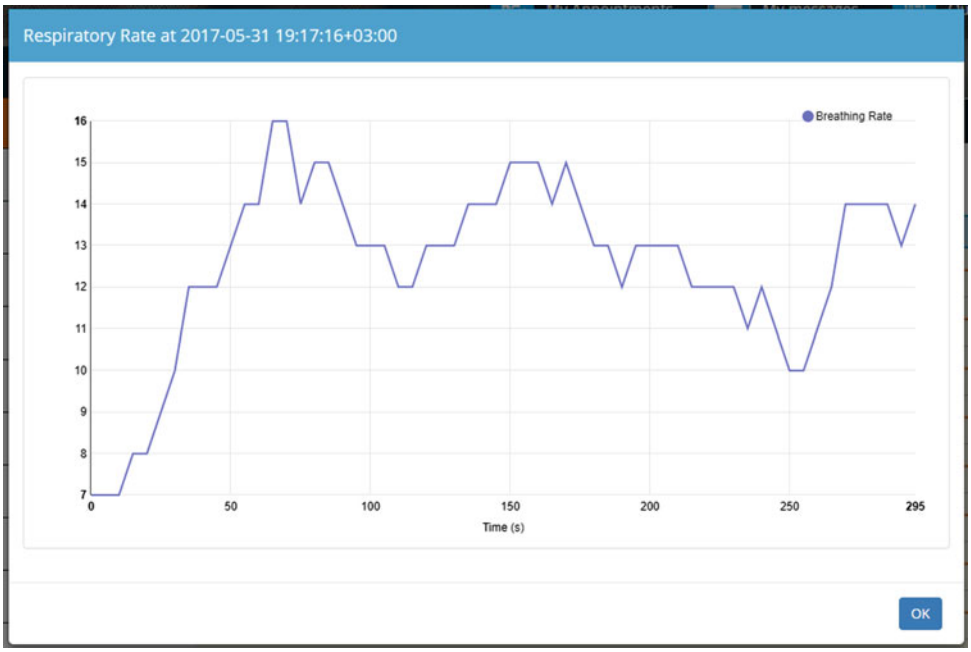



**Fig. 3** A sample of the respiratory rate graph during a 5-min chunk, as shown on the screen of the healthcare application

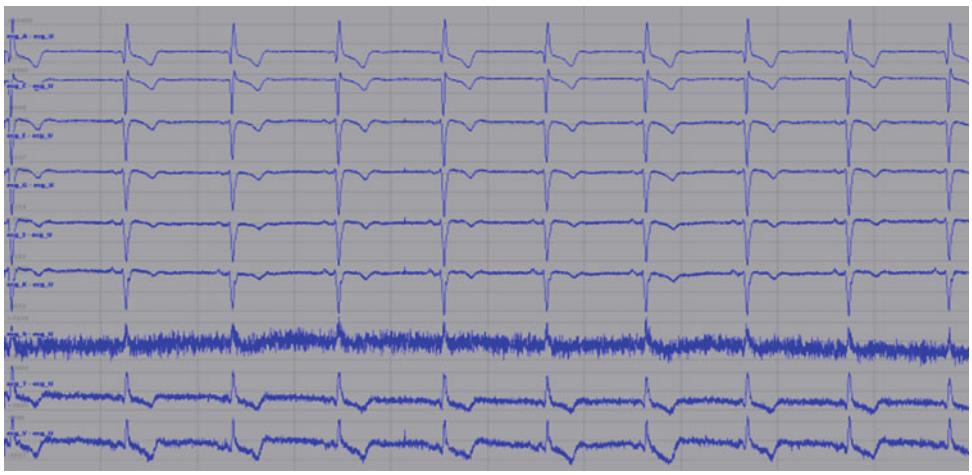


**Fig. 4** A segment from a 12-lead ECG file

clinical alerts from the DSS for the total number of patients during the Greek pilot. Currently, the association of the fired alerts to events that took place during the participation in the study is being analyzed in order to evaluate the validity of the design.

## Challenges During the Pilot

There were certain technical problems that emerged at the initial stages of the Greek pilot study. These mainly concerned the connectivity between the vest and the tablet when uploading the recorded data to the cloud and between the peripheral devices and the tablet. The connectivity issues were more prudent in cases of congested environment from

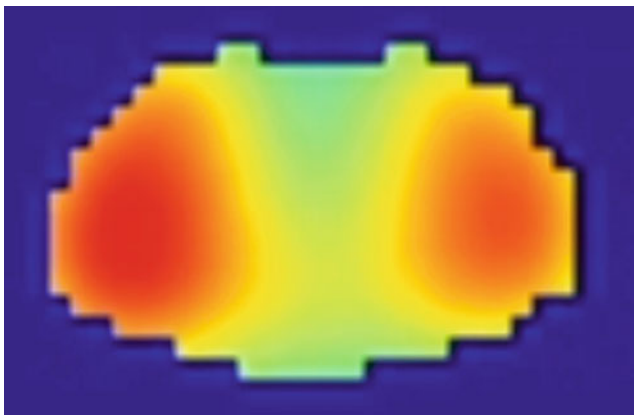


**Fig. 5** A functional EIT image taken from a WELCOME patient. The dark red color depicts higher airflow into the lungs

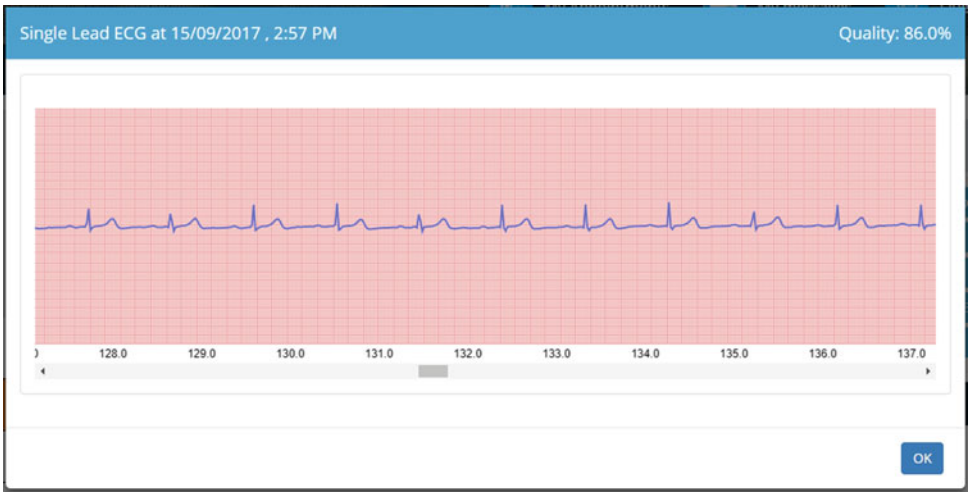



**Fig. 6** A single lead ECG strip, as shown on the screen of the healthcare application

multiple Wi-Fi sources. Moreover, some elderly patients found it difficult to operate the equipment, resulting in relatively low compliance in some cases. They were mainly people who had never used computers, tablets or smartphones in the past. The sensors vest operated better in specific body types. When the waist circumference was too large compared to the chest dimension, there was suboptimal contact of the sensors to the skin and low quality of many different recordings. The auscultation files were unreliable in many cases, due to the body type and the presence of hair on the surface of the chest in some patients.

Under normal conditions, the obtained biosignals were captured and transmitted to the cloud database successfully and were available to the attending physicians in almost real time for most of them.

## Examples of Collected Signals

The images provided in this paper are some examples of the biosignals recorded and transmitted via the WELCOME system. As can be seen at the screenshots (Figs. 3, 4, 5 and 6), the medical doctors are able to view the recorded biosignals, plot the trends in specific measured values and have some form of imaging of the chest while the patients are at home with the EIT modality (Fig. 5). The importance of the EIT imaging is highly emphasized, since it is the first effort to receive some form of lung imaging at the home setting of the patients and in a continuous manner, capable of highlighting changes in the distribution of air into the lungs (in various pathological conditions, like pneumonia, obstruction, pleural large effusions, etc.).

## Discussion

The WELCOME system may provide the medical doctors unprecedented insight on the disease progression and variability on a daily basis and the DSS module is currently under review for its efficacy and precision. The wide range and the volume of the recorded bioparameters allows for a clear view of a patient's health status and this is possible in a comprehensive way from a remote location for the first time.

Despite the problems encountered, most of the subjects reported excitement from the participation in the study and the majority felt that the system could help them better manage their chronic condition and promote the concept of personalized and integrated care.

**Acknowledgements** The authors would like to thank the WELCOME project coordinator and consortium (http://www.welcome-project.eu/) for providing some of the images and the associated information.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Rabe K, Hurd S, Anzueto A, Barnes PJ et al (2007) Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 176:55–532
2. Gibson GJ, Loddenkemper R, Lundbäck B, Sibille Y (2013) Respiratory health and disease in Europe: the new European Lung White Book. Eur Respir J 42:63–559
3. Smith MC1, Wrobel JP (2014) Epidemiology and clinical impact of major comorbidities in patients with COPD. Int J Chron

Obstruct Pulmon Dis 9:871–88. https://doi.org/10.2147/COPD.S49621
4. Hillas G, Perlikos F, Tsiligianni I, Tzanakis N (2015) Managing comorbidities in COPD. Int J Chron Obstruct Pulmon Dis 7 (10):95–109. https://doi.org/10.2147/COPD.S54473
5. Bernocchi P, Scalvini S, Galli T, Paneroni M et al (2016) A multidisciplinary telehealth program in patients with combined chronic obstructive pulmonary disease and chronic heart failure: study protocol for a randomized controlled trial. Trials 17(1):462
6. Phanareth K, Vingtoft S, Christensen AS, Nielsen JS et al (2017) The epital care model: a new person-centered model of technology-enabled integrated care for people with long term conditions. JMIR Res Protoc. 6(1):e6. https://doi.org/10.2196/resprot.6506
7. Guertin JR, Bowen JM, Gosse C, Blackhouse G et al (2017) Preliminary results of the adoption and application of the integrated comprehensive care bundle care program when treating patients with chronic obstructive pulmonary disease. Can Respir J 7049483. https://doi.org/10.1155/2017/7049483
8. Esteban C, Moraza J, Iriberri M, Aguirre U et al (2016) Outcomes of a telemonitoring-based program (telEPOC) in frequently hospitalized COPD patients. Int J Chron Obstruct Pulmon Dis. 11:2919–2930
9. Frerichs I, Amato MB, van Kaam AH, Tingay DG, Zhao Z et al (2016) Chest electrical impedance tomography examination, data analysis, terminology, critical use and recommendations: consensus statement of the translational EIT development study group. Thorax 72:83–93
10. Kayyali R, Odeh B, Frerichs I, Davies N, Perantoni E et al (2016) COPD care delivery pathways in five European union countries: mapping and health care professionals' perceptions. Int J Chron Obstruct Pulmon Dis 11:8–2831
11. Mendes L, Vogiatzis IM, Perantoni E, Kaimakamis E, Chouvarda I, Maglaveras N et al (2016) Detection of crackle events using a multi-feature approach. In: 2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC), pp 3679–3683

# Part IV
# Biosignals and Biomarkers

# Adipose Tissue as a Biomarker in Data Mining Predictive Models of Metabolic Pathophysiologies

O. Tsave, I. Kavakiotis, I. Vlahavas, and A. Salifoglou

**Abstract**

It is well known that the metabolic syndrome emerges as one of the major public health issues worldwide. In diabetes and other metabolism related diseases, further complexity is added in diagnosis and prognosis due to the presence of metabolic syndrome, including obesity. Obesity, which is defined as an excess of body fat, can be described as an underlying risk factor of almost any of the aforementioned metabolic related pathologies. Moreover, a very likely potential link between such pathologies and obesity is the adipose tissue, which functions as an endocrine organ. Since obesity serves as general key to metabolism related disorders and complications, the adipose tissue can be a useful tool in predicting such pathologies. In the present mini review work, several representative studies are discussed with respect to the effectiveness of adipose tissue as a valuable biomarker along with other factors taken into consideration with data mining approaches. Taken together, adipose tissue can be used in data mining as a predictive tool in diabetes, mortality, cardiometabolic risk and other metabolism related pathologies.

**Keywords**

Adipose tissue • Metabolic abnormalities • Clinical biomarkers • Data mining • Prediction

## Introduction

A biomarker is defined as an indicator of a certain condition, such as a biological process or entity, in (patho)physiological conditions. The rising need for new, novel and efficient biomarkers is crucial in numerous diseases relating to the onset, diagnosis, prognosis, therapeutic strategy or causal events. The metabolic syndrome emerges as one of the major public health issues worldwide. By definition, the metabolic syndrome, along with its related pathologies, is one of the most controversial research/medical issues [1, 2]. In general, it refers to a constellation of metabolic risk factors involving (a) atherogenic dyslipidemia, (b) elevated blood pressure, (c) elevated glucose levels strongly related to insulin resistance, (d) prothrombotic state, and (e) proinflammatory state [2]. The aforementioned factors are not clinically determined despite the fact that they can be easily identified through specific assays. It should be noticed that the onset of metabolic syndrome is strongly related to several pathophysiological conditions that include cardiovascular risk, obesity, diabetes, and other metabolic complications [2]. Although the two-way association of these risk factors-pathologies has been known for more than 80 years, clustering had received poor attention until 1988, when Reaven defined syndrome X with respect to insulin resistance, hyperglycemia, hypertension, low HDL-cholesterol, and raised VLDL-triglycerides. It is worth mentioning that obesity, which is defined as an excess of body fat, can be described as an underlying risk

O. Tsave (✉) · A. Salifoglou
Department of Chemical Engineering, Laboratory of Inorganic Chemistry and Advanced Materials, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
e-mail: tsaveolga@gmail.com

I. Kavakiotis · I. Vlahavas
School of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_18

factor of almost any of the aforementioned metabolic related pathologies. The prevalence of obesity and diabetes has increased dramatically over the past decades and there is a strong causal relationship between diabetes and insulin resistance [3–6]. A very likely link between diabetes and obesity is the adipose tissue. The adipose tissue functions as an endocrine organ, secreting hormones and cytokines able to regulate and maintain metabolic homeostasis in other tissues. Normally, excess fat is stored in adipocytes and only low amounts of triglycerides exist in non-adipose cells [7, 8]. In obesity, the capacity of adipocytes to accommodate excess lipids can be exceeded, resulting in an imbalanced accumulation of lipids in other tissues such as in the epicardial region. Numerous studies (both experimental and clinical) support the fact that obesity acts on several metabolic pathways, resulting in the onset of multiple potential risk factors. This complexity provides a great challenge in when investigating potential biomarkers in an effort to deal with new targets of therapy, early diagnosis, prognosis and biotechnological applications [9]. Taken together, since obesity serves as a general key to metabolism-related disorders and complications, the adipose tissue (characteristics, physiology, structure, volume etc.) can be employed as a tool for predicting such pathologies. The aim of the present work is to (a) assess the capacity of adipose tissue to serve as a predictive biomarker through an overview of recent representative studies in humans, and (b) investigate/suggest the most efficient adipose tissue characteristic-parameter taken into consideration.

## Adipose Tissue as a Predictive Tool

In an attempt to use adipose tissue as a biomarker, several parameters and specific characteristics are employed that can be measured.

This section describes representative studies dealing with adipose tissue as a predictive marker of mortality, diabetes and cardiometabolic risk, employing volume, structure, tissue–specific components, accumulation values/indicators etc., in combination with other measurable features (such as anthropometric factors).

In [10], researchers employed the value of epicardial fat thickness (EFT) as a prognostic biomarker of increased inflammatory profile, in patients diagnosed with Diabetes Mellitus type II (DM type II) and acute myocardial infarction (AMI). It has been shown that epicardial adipose tissue is linked with the onset/progression of coronary artery disease. This is due to the secretion of several pro-inflammatory cytokines. In this study, efforts were made to assess the correlation between three distinct factors (EFT, prolonged presence of elevated circulating levels of hs-CRP after an AMI, and left ventricular (LV) remodeling) in patients with DM type II. The significance of such an attempt is driven by the fact that a great amount of death incidents in diabetes are strongly related to coronary artery disease, mainly through AMI, which in turn is tightly correlated with the degree of LV remodeling and levels of inflammatory status following AMI. In [11], adipose tissue density was used as a biomarker in the prediction of the risk of mortality in older adults. As previously mentioned, the composition of adipose tissue is strongly related to obesity and its complications. Several characteristics of adipose tissue, including the amount or ectopic locations and distribution, can be used in evaluating the potential risk for mortality caused by obesity. In this study, the density of VAT (visceral adipose tissue) and SAT (subcutaneous adipose tissue) was employed to investigate the potential correlation between mortality and density of adipose tissue. The measurements included clinical and subclinical body composition in both humans and nonhuman primates. Interestingly, in the undertaken study, an association of adipose tissue density at baseline in quintiles with adipocytokines was also performed, thereby enabling further correlations among tissue-specific molecules, such as adiponectin and leptin, which serve as valuable clinical biomarkers. The herein attempted evaluation supports the fact that adipose tissue density can serve as a predictive factor of survival in older adults, even when an adipose depot itself is not associated with death risk. Overall obesity and Body Fat Distribution were also used in [12], predicting mortality. Researchers took into consideration factors, such as ethnicity, education, history of diagnosed respiratory disease, body weight, and hypertension. Age appears to have a strong impact on the association of adiposity and mortality. The collective results suggest that ratio measurements of BFD can be informative in identifying increased risk of mortality in middle-age adults. By the same token, in [13], several adiposity indices were used to extract potential associations with cardiometabolic risk(s). The aim of the specific cohort study was to compare the general adiposity index, referred to as BMI (body mass index), with abdominal obesity indices, such as waist circumference, waist-to-hip ratio, and waist-to-height ratio, in order to unravel the most efficient predictive marker. It is well-stated that adipose tissue excess is strongly associated with a number of cardiometabolic factors involving hypertension, dyslipidemia and DM. The measurements employed were based on anthropometric data (hip circumference, body weight, etc.), blood pressure (diastolic, systolic), and analysis of fasting blood samples. In the study, cardiometabolic risk factors increase directly with rising obesity, whereas indices of body fat centralization emerge as a better predictor of cardiometabolic risk factors compared to BMI. Similar to that, in [14], proton density fat fraction (PDFF) was measured by magnetic resonance imaging (MRI). In this study, supraclavicular and gluteal adipose tissue with subcutaneous

and visceral adipose tissue (SAT and VAT), and liver fat fraction in combination with several anthropometric obesity markers (e.g. waist circumference) were employed. The results project a significant difference between PDFF in the supraclavicular and gluteal fat depots and thus differences in cell-related characteristics. In addition, there is a positive correlation among PDFF, imaging and anthropometric obesity markers, with adipose tissue PDFF emerging as a valuable biomarker toward obese phenotype. Several studies support the fact that Visceral Adiposity Index (VAI) has been proven to be a direct marker of adipose distribution. That, indirectly, can be used in the prediction of cardiometabolic risk onset [15, 16]. Moreover, VAI can be used as an indicator of developing metabolic syndrome and hence DM type II. In [17], VAI (sex-specific) was used in a cohort study as a tool, in an effort to identify potential risks for DM type II. More specifically, the study aimed at evaluating the predictive capacity of VAI, Body Fat Indices, and BMI, whereas the authors also used biomedical measurements (fasting blood glucose, HLDL). The overall results suggested that VAI can be associated with the risk of DM and stands as the best indicator compared to the aforementioned ones. Similarly, in [18], VAI, in combination with the hypertriglyceridemic waist (HTGW) phenotype, was used to predict the development of hypertension in a high-risk population that does not appear to have either DM or hypertension. Data concerning age, sex, BMI, hemoglobin A1c (HbA1c), total cholesterol (TC), low density lipoprotein cholesterol (LDL-C), HDL-C, TG, BP, and family/personal medical history were also taken into consideration. In conclusion, VAI, in addition to the HTGW phenotype, proved to be a weak predictor of hypertension. Interestingly, in [19], employment of VAI in predicting DM type II, failed to improve prediction. VAI was used along with several (easily) measurable anthropometric parameters. Although VAI appears to be a strong predictor of DM, its predictive capacity does not surpass the efficiency of other indicators, such as BMI, WC, WHtR, and WHR in the current cohort study.

## Datasets, Data Mining and Statistical Analysis

In this section, the datasets and methods used in the performed studies will be presented briefly. Opincariu et al. [10] tried to assess the prognostic value of epicardial fat thickness as a biomarker of increased inflammatory status in patients with DM Type 2 and Acute Myocardial Infarction. Analysis was performed using a dataset of 98 patients (45 diabetic and 43 non-diabetic). Authors used Fisher's exact test and the Student's t-test to compare the baseline characteristics of patients, linear regression to evaluate the correlation between the left ventricular ejection fraction and the remodeling index, and the hs-CRP values, and finally logistic regression to identify independent predictors of left ventricular remodeling. Murphy et al. [11] proposed the adipose tissue density as a biomarker to predict mortality risk in the elderly. In their study, the authors used a subset of data from two different cohort studies, the Health, Aging, and Body Composition (Health ABC) study, on one hand, and the Age, Gene and/or Environment Susceptibility-Reykjavik (AGES-Reykjavik) study on the other, following application of some strict inclusion criteria (Health ABC n = 2,735 and AGES-Reykjavik n = 5,131) as well as 24 non-human primates. In order to examine associations between adipose density and mortality authors, they used Cox proportional hazard models, reflecting a class of survival models. Frantz et al. [14] tried to investigate the association of proton density fat fraction in adipose tissue with imaging-based and anthropometric obesity markers, analyzing a dataset of 61 adults through Student's t-test and Pearson correlation analysis. Palacio et al. [13] investigated the association between adiposity indices and cardiometabolic risk factors through (a) Pearson's product moment correlation coefficients, (b) linear regression, and (c) the area under the receiver-operating characteristic (ROC) curves. It was a regional study including 858 adults. Reis et al. [12] compared the overall obesity with body fat distribution, in an attempt to predict the risk of mortality. The analysis performed on a dataset comprised of 5,799 men and 6,429 women (total 12,228 adults) was included in the third National Health and Nutrition Examination Survey. Multivariable Cox proportional hazard regression models were again used to examine the relation of adiposity with mortality, after examining various potential confounders. Chen et al. [17] studied the visceral adiposity index as a predictor for DM type 2 and compared its predicting performance with other body fatness indices. Authors utilized Cox proportional hazard regression models, receiver operating characteristic (ROC) curves and areas under curve (AUC) to analyze a dataset of 3,461 participants. Janghorbani et al. [18] examined the use of the Visceral Adiposity Index and Hypertriglyceridemic Waist Phenotype for the prediction of Incident Hypertension. The authors analyzed a 1,375 individual dataset using a series of statistical analysis methods: the Student t test, Mann-Whitney U test; ANOVA; the Kruskal-Wallis test; the chi-square test, Pearson correlation; Spearman rank correlation; and binary logistic regression. Finally, Janghorbani and Amini [19] concluded in their study that the visceral adiposity index did not improve the prediction of diabetes compared to measurable anthropometric markers, such as the hypertriglyceridemic waist (HTGW) phenotype, body mass index (BMI), waist circumference (WC), waist-to-height ratio (WHtR) and waist-to-hip ratio (WHR). In their study, the authors

analyzed a dataset of 1720 individuals, using logistic regression and receiver operating characteristic (ROC) curves.

## Conclusions

In the present work, adipose tissue is discussed (mini review) as a potential biomarker in the prediction of related metabolic disorders (diabetes, mortality, cardiometabolic risk) using data mining approaches. With respect to adipose tissue, the features examined include the (a) site of the body (epicardial, gluteal, etc.), (b) density, (c) composition (e.g. adipocytokines), and (d) structure. The specific values were investigated through several assays at the clinical level and were evaluated in combination with other, not directly related parameters, such as ethnicity or sex. The results suggest that although adipose tissue could be used as a marker of prediction, a number of well-defined parameters (topology, anatomy, structure, composition etc.) should always be taken into account.

When it comes to the analysis process, the most frequently used computational analysis methods included regression, both logistic and linear, and Cox proportional hazard methods. Collectively, adipose tissue could serve as a source of factors parameterizing clinical data into predictive biomarkers in a variety of metabolic pathologies.

**Acknowledgements** O.T. was financially supported as a postdoctoral fellow by Greek State Scholarships Foundation, through the Siemens Program: "IKY Fellowships of Excellence for Postgraduate Studies in Greece—Siemens Program (2015–2017)".

**Conflict of Interest** "The authors declare that they have no conflict of interest".

## References

1. Kaur J (2014) A comprehensive review on metabolic syndrome. Cardiol Res Pract 2014:943162. https://doi.org/10.1155/2014/943162
2. Huang PL (2009) A comprehensive definition for metabolic syndrome. Dis Model Mech 2(5–6):231–237. https://doi.org/10.1242/dmm.001180
3. Reynisdottir S, Ellerfeld K, Wahrenberg H et al (1994) J Clin Investig 93(6):2590–2599
4. Reaven GM (1988) Demonstration of the central role of insulin resistance in type 2 diabetes and cardiovascular disease. Diabetes 37:1595–1607
5. Warram JH, Martin BC et al (1990) Slow glucose removal rate and hyperinsulinemia precede the development of type II diabetes in the offspring of diabetic parents. Ann Intern Med 113:909–915
6. Lillioja S, Mott DM et al (1988) Impaired glucose tolerance as a disorder of insulin action. Longitudinal and cross-sectional studies in Pima Indians. N Engl J Med 318:1217–1225
7. Guilherme A, Virbasius JV et al (2008) Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. Nat Rev Mol Cell Biol 9(5):367–377
8. Frayn KN, Tan GD, Karpe F (2007) Adipose tissue: a key target for diabetes pathophysiology and treatment? adipose tissue and insulin resistance. Horm Metab Res 39:739–742
9. Caveney EJ, Cohen OJ (2011) Diabetes and biomarkers. J Diabetes Sci Technol 5(1):192–197. https://doi.org/10.1177/193229681100500127
10. Opincariu1 O, Mester1 A, Dobra M, et al (2016) Prognostic value of epicardial fat thickness as a biomarker of increased inflammatory status in patients with type 2 diabetes mellitus and acute myocardial infarction. J Cardiovasc Emerg 2(1):11–18
11. Murphy RA, Register TC, Shively CA et al (2014) Adipose tissue density, a novel biomarker predicting mortality risk in older adults. J Gerontol A Biol Sci Med Sci 69(1):109–117. https://doi.org/10.1093/gerona/glt070
12. Reis JP, Macera CA, Araneta MR et al (2009) Comparison of overall obesity and body fat distribution in predicting risk of mortality. Obesity (Silver Spring) 17(6):1232–1239. https://doi.org/10.1038/oby.2008.664
13. Palacios C, Pérez CM, Guzmán M et al (2011) Association between adiposity indices and cardiometabolic risk factors among adults living in Puerto Rico. Public Health Nutr. 14(10):1714–1723. https://doi.org/10.1017/S1368980011000796
14. Franz D, Weidlich D, Freitag F et al (2017) Association of proton density fat fraction in adipose tissue with imaging-based and anthropometric obesity markers in adults. Int J Obes (Lond). https://doi.org/10.1038/ijo.2017.194
15. Amato MC, Giordano C (2014) Visceral adiposity index: an indicator of adipose tissue dysfunction. 2014 Int J Endocrinol 2014, 7 pp. https://doi.org/10.1155/2014/730827
16. Amato MC, Giordano C, Galia M et al (2010) Visceral adiposity index a reliable indicator of visceral fat function associated with cardiometabolic risk. Diabetes Care 33(4):920–922. https://doi.org/10.2337/dc09-1825
17. Chen C, Xu Y, Guo ZR et al (2014) The application of visceral adiposity index in identifying type 2 diabetes risks based on a prospective cohort in China. Lipids Health Dis 13:108. https://doi.org/10.1186/1476-511X-13-108
18. Janghorbani M, Salamat MR, Aminorroaya A, Amini M (2017) Utility of the visceral adiposity index and hypertriglyceridemic waist phenotype for predicting incident hypertension. Endocrinol Metab (Seoul) 32(2):221–229. https://doi.org/10.3803/EnM.2017.32.2.221
19. Janghorbani M, Amini M (2016) The visceral adiposity index in comparison with easily measurable anthropometric markers did not improve prediction of diabetes. Can J Diabetes 40(5):393–398. https://doi.org/10.1016/j.jcjd.2016.02.008

# Portable Near-Infrared Spectroscopy for Detecting Peripheral Arterial Occlusion

W.-C. Lu, S.-H. Lu, M.-F. Chen, T.-C. Fu, K.-P. Lin, and C.-L. Tsai

**Abstract**

The prevalence rate of peripheral arterial disease has been increasing in Taiwan in recent years. The occlusion of artery will reflect on the oxygen saturation of muscle, especially after exercise. A wireless tissue oximetry based on near-infrared spectroscopy is developed in this study to help to detect arterial occlusion. The measurement of change in hemoglobin oxygen saturation was measured on gastrocnemius muscle using air cuff with different pressure to simulate the blockage of blood flow.

**Keywords**

Near-infrared spectroscopy • Arterial occlusion • Peripheral arterial disease (PAD)

## Introduction

The occlusion of peripheral arteries has become a serious health problem in aging societies. In Taiwan, the elderly population has accounted for 13.2% of the total population in 2016. According to a record from the Ministry of Interior of Taiwan, population aging together with greasy diet result in a constant increase of patients with peripheral arterial disease.

The main symptom of peripheral arterial disease is atheroma of low limbs. Fat plaque developed and accumulated on arterial walls causes the loss of elasticity. Intimal hyperplasia starts to narrow arteries and reduce blood flow, which in turn decreases transmission of nutrition and oxygen through capillaries to tissues. If the disease were not diagnosed in time and treated, severe limb ischemia may occur and require amputation to save the life of patient. According to a statistic, Patients who suffered from the disease for one year receiving amputation accounts for 30% and the death rate accounts for 25% [1].

The prevailing rate of diabetes in Taiwan has risen 25% and diabetics are frequently suffered from limb ischemia. They not only have poor microcirculation but also with PAD, which was worsened to require an amputation and their death rate in a short time after the surgery is much higher than non-diabetics. Therefore, it is extremely important to early diagnose PAD.

Ankle-Brachial index (ABI) is the measurement currently applied in clinic to diagnose PAD. However, it is also indicated that the difference of ABI between PAD patients with no symptoms and those in the early stages is less than 0.01–0.02 [2, 3]. This means it is hard to diagnose PAD in advance by using ABI examination. In addition, ABI is an indirect measurement and it can't identify which artery is occluded. Therefore, patients have to take vessel ultrasound exam, contrast x-ray imaging, or nuclear magnetic resonance imaging as a further exam. These exams can only be operated in a well-equipped hospital with experienced doctors to

W.-C. Lu · S.-H. Lu · C.-L. Tsai (✉)
Biomedical Engineering Department, Chung Yuan Christian University, Taoyuan, Taiwan
e-mail: clt@cycu.edu.tw

M.-F. Chen · K.-P. Lin
Electrical Engineering Department, Chung Yuan Christian University, Taoyuan, Taiwan

T.-C. Fu
Department of Internal Medicine, Heart Failure Center, Chang Gung Memorial Hospital, Keelung, Keelung, Taiwan

K.-P. Lin · C.-L. Tsai
Technology Translation Center for Medical Device, Chung Yuan Christian University, 200, Chung-Pei Rd., Taoyuan, Taiwan



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_19

indicate the exact location of arterial occlusion. The study is to develop a PAD early detection system using non-invasive optic measurement to diagnose arterial occlusion. The price is relatively low and it is expected to be conduct in small or medium size outpatient clinics.

## Materials and Methods

### A. *Near-Infrared Spectroscopy*

The red color of muscle tissue is mainly caused by myoglobin in muscle fiber rather than hemoglobin in blood. Myoglobin is also a protein that stores oxygen in muscle tissue. Since myoglobin has a stronger oxygen binding capacity than hemoglobin, it can easily convert to oxidative myoglobin when get in contact with oxygen molecule. The oxygen is released by oxy-hemoglobin in capillary due to the oxygen concentration gradient and diffused into muscle fiber. During exercise oxygen is consumed to produce carbonate acid. If there were not enough blood supply to muscle, more and more oxy-myoglobin would be reduced to myoglobin. For normal people oxygen saturation of myoglobin will decrease during exercise, and quickly recovered after exercise. For patient with occluded artery, the saturation value will drop lower and take a much longer time to recover to normal value. It is possible to measure such change during the recovery period to detect the severity of arteria occlusion.

Both hemoglobin and myoglobin change color when they are bound with oxygen. The color of blood and muscle is determined by the amount of hemoglobin in blood and myoglobin in muscle tissue, as well as their oxygen saturation. Therefore, their oxygen content can be measured by non-invasive optical methods. Since muscle tissue is generally cover by a thick layer of skin and adipose tissue, only near-infrared at the range between 600 and 1300 nm is suitable for detecting the oxygen saturation of deep muscle tissue. Within this spectral range, oxidative hemoglobin (HbO), hemoglobin (Hb), oxidative myoglobin (MbO), myoglobin (Mb), and cytochrome (Cyt) are the major light absorption chromophores that absorb the light energy. Scattering of light in biological tissue is another factor that affects the reflectance of light back to skin surface. The great variability of tissue structure between different subjects makes it almost impossible to carry out any useful quantification measurement. Only the self-reference technique of measurement as used in pulse oximeter can automatically cancel out the individual difference.

Three near-infrared wavelengths: 740, 808 and 850 nm are chosen in this study for detecting the oxygen saturation of myoglobin. These wavelengths are all close to an isobestic point of the absorption spectra of myoglobin and oxidative myoglobin at around 805 nm. With similar wavelength, they also have very similar scattering property that makes their pass lengths in tissue roughly the same.

To eliminate the influence of skin color, thicknesses and structure of skin and adipose tissue, two LEDs with the same wavelength are placed at two slightly different distances from the photo diode, as drawn in Fig. 1. The distance of near LED is 25 mm and the distance of far LED is 30 mm. The light emitted by both LEDs are able to travel through muscle tissue if the source-detector distance is longer than about twice the thickness of skin and adipose tissue. The light from near and far LEDs travel through similar path in tissue. The major path difference is the portion in muscle tissue. This means the intensity difference of measurements from two LEDs is caused by two factors, the absorption of myoglobin and oxy-myoglobin.

To calculate the absorption by myoglobin, the logarithm of the light intensity ratio of incident to reflection is treated as the absorbance of light by tissue. The absorbance is mostly contributed by the absorption of hemoglobin and myoglobin, as well as the unpredictable scattering factor of skin and adipose tissues. Fortunately, the contribution of tissue scattering in two measurements will cancel out when the difference of absorbance is taken. The rest part of absorbance is expressed in Eq. (1) as the absorption by myoglobin and oxy-myoglobin. The last term in equation is for compensating the dc drift of light signal.

$$\Delta A = a_{MbO}\Delta l_{MbO} + a_{Mb}\Delta l_{Mb} + S_M\Delta l_s \qquad (1)$$

where $\Delta A$ is the absorbance difference of measurement taken from light sources at different distance, $a_{Mb}$ and $a_{MbO}$ are the absorption coefficients of myoglobin and oxy-myoglobin. $\Delta l_{Mb}$ and $\Delta l_{MbO}$ are the mean optical path differences of myoglobin and oxy-myoglobin, respectively. $S_M$ is the

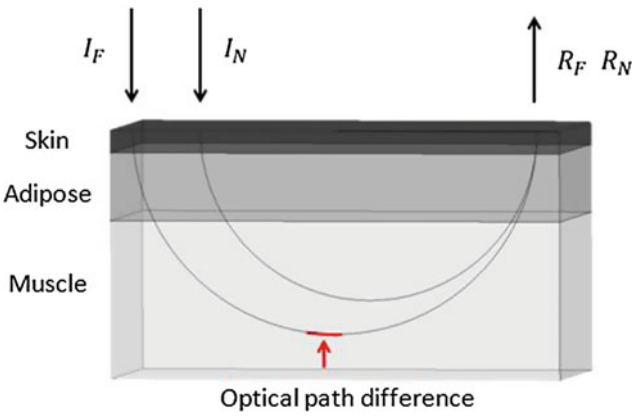



**Fig. 1** Illustration of different reflective light paths through tissue layers

scattering coefficient of muscle and $\Delta l_s$ is the mean optical path difference due to tissue scattering.

The difference of absorbance in Eq. (1) is expressed and a linear combination of three different terms. Each term includes an absorption of scattering coefficient multiplied by a variable, the difference of mean path length. These variables of represent the concentration of myoglobin and oxy-myoglobin in muscle tissue. Because there are three variables in the equation, it requires measurement at three wavelengths to have three equations to solve the three unknown variables.

### B. *Measurement System*

Most PAD examining devices are bulky or wired to the patients. Patients cannot walk around freely to show intermittent claudication during the test. In order to do the measurement in normal walk or exercise condition, this measuring system is designed as a portable device. The block diagram of measurement system can be divided into two parts, as shown in Fig. 2. One is the near-infrared spectroscopy device which is to attach to the measurement site. The optical probe includes 6 LED light sources, two for each wavelength. A photodiode light sensor detects and converts light signal into electrical current. The signal current flows into a microprocessor (HY16F198, Hycon Tech. Inc.) and is digitized by a 24 bits analog-to-digital convertor. Then, the digital signal is sent to a computer through Bluetooth wireless connection for further data analysis. The other part of the system is a computer that receives the transmitted signals and calculates the change in oxygen content.

To simulate the arterial occlusion condition on normal subjects, blood flow through upper leg was blocked by air cuff. The testing procedure refer to previous studies [4, 5]. During the experiment, the subject lie down on an experimental bed with an optical probe placed on the gastrocnemius muscle, as shown in Fig. 3.

An air cuff for blood pressure measurement is tied on the upper leg to block blood flow to lower leg. At the beginning of test, the subject was asked to lie down and rest for 3 min to achieve a steady state. Then, cuff pressure was pumped up to reach the target pressure in 15 s. The pressure is maintained for 3 min. After cuff pressure was released, subject stays in rest for 10 min to record the recovery status.

The typical physiological reaction of blood supply to leg during vascular occlusion experiment has several features. The concentration of myoglobin and oxy-hemoglobin remain constant before occlusion. When blood vessel is occluded, the concentration of myoglobin starts to increase and oxy-hemoglobin decreases, respectively. Because there is a compensatory effect that continue to increase tissue blood perfusion, oxygen saturation of myoglobin is expected to rebound to a peak value higher than its initial value. After that, concentration of myoglobin and oxidative myoglobin will gradually drift back to their normal value. For patient with arterial blockage, it will take much more time for them to return to original state.

## Results

The self-made optical probe for arterial vessel obstruction is shown in Fig. 4. The grid paper at background is 5 mm by 5 mm for each grid. LEDs are put on two concentric circles with radiuses of 25 and 30 mm while the photodiode is placed at the center. The optical probe is fixed on leg using elastic bandage.

Figure 5 shows the results of vascular occlusion test with air cuff pressure to 220 mmHg and the measurement was taken at gastrocnemius muscle. The calculated change in concentration of oxidative myoglobin (red) and myoglobin (blue) are quite similar to expectation.

At the beginning of arterial occlusion, the concentration of oxidative myoglobin gradually decreased, and the concentration of myoglobin increased, respectively. This is caused by the consumption of oxygen by tissue in the lower leg, and the lack of oxygen supply. Right after the release of cuff pressure, these concentrations quickly rebounded toward their initial values. The $T_{50}$ value represents the time for oxidative myoglobin to recover 50% from the lowest value right before cuff pressure was released. In this case, it took 12 s to reach the middle value.

Figure 6 shows the similar result with the air cuff pressure pumped up to 250 mmHg in the arterial occlusion test. With such a high pressure, the blood flow to leg was probably

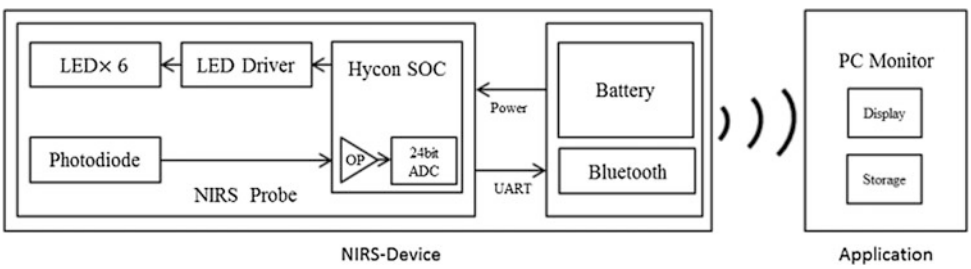



**Fig. 2** Block diagram of measurement system

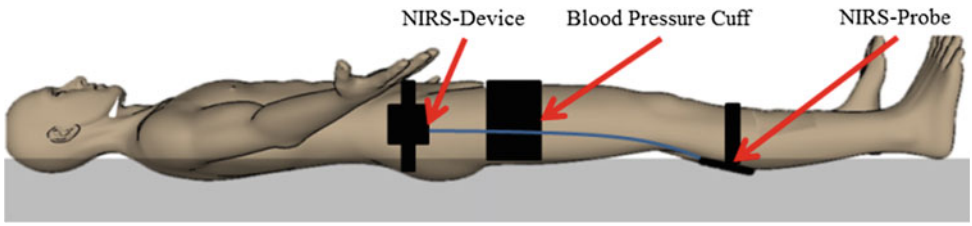



**Fig. 3** Arterial occlusion experiment in supine position

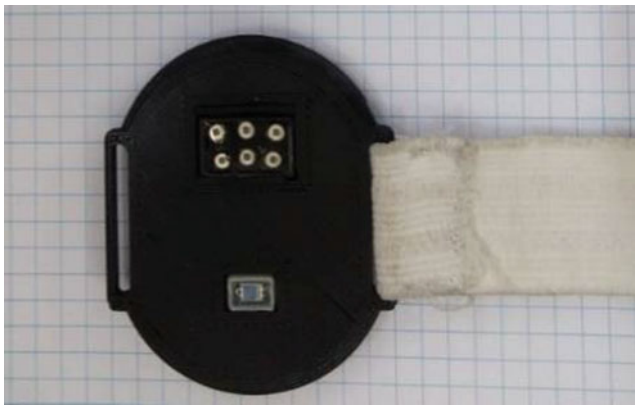


**Fig. 4** Optical probe of near-infrared spectroscopy

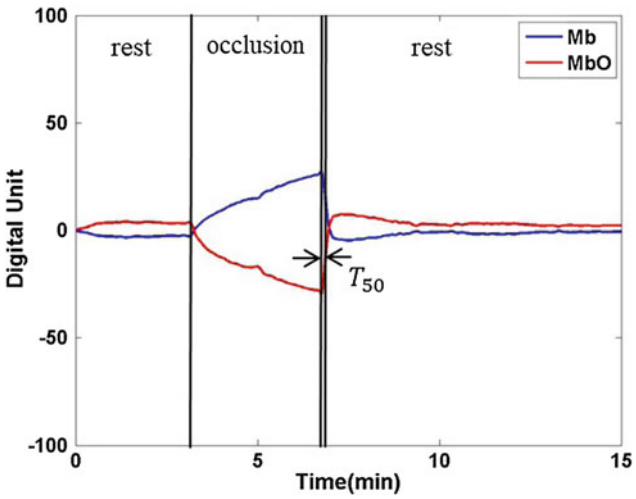



**Fig. 5** Measurement of gastrocnemius muscle in arterial occlusion experiment with cuff pressure at 220 mmHg

totally blocked. The concentration of oxidative myoglobin greatly dropped and approached a steady state value, and so is the concentration of myoglobin. The acute change of concentrations before pressure release is most likely by a probe motion artefact during measurement. The $T_{50}$ value for 250 mmHg was 15 s which is longer than that of

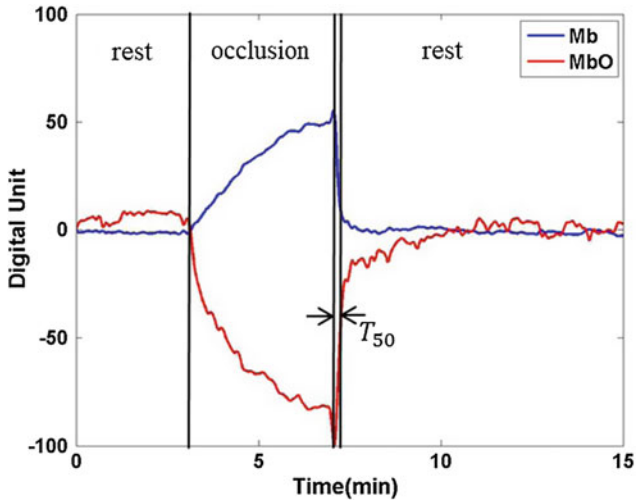



**Fig. 6** Measurement of gastrocnemius muscle in arterial occlusion experiment with cuff pressure at 250 mmHg

220 mmHg. The most significant feature for 250 mmHg is the unstable variation during the recovery time after $T_{50}$.

## Discussion

The measurement and calculation of change in the concentration of myoglobin and oxidative myoglobin during the occlusion experiments have very reasonable results. Oxidative hemoglobin decreased with the blockage of blood supply, and the concentration recovered to normal value after the cuff pressure was released. The concentration of myoglobin changed in the opposite direction, respectively. When the cuff pressure is higher, the concentration of myoglobin changed accordingly. These features show the portable near-infrared spectrograph is a promising device for detecting the blood vessel occlusion of peripheral arterial disease.

The device is currently only working as a qualitative evaluation machine. In order to accurately measure the change in myoglobin concentration or myoglobin oxygen saturation, the measurement has to be carefully calibrated by

some sophisticated systems or phantoms that can provide known values of myoglobin concentration or oxygen saturation.

The vessel occlusion experiment is a simulation of peripheral arterial disease test on normal subjects. Measurements on real PAD patients are still needed to further improve the performance of this device as well as to carry out calibration on human subject.

The main components used to build this device, such as light emitting diode and photodiodes are relatively cheaper than lasers and photomultiplier tube that are used in some commercial NIRS system. Therefore, it is possible to greatly deduce the cost of manufacturing this device.

## Conclusion

The performance of this portable near-infrared spectrograph is very close to some expensive NIRS systems that are currently used in some hospitals for detecting peripheral arterial disease. The light weight, small in size and low in price of this device make it much easier to promote it to small clinics as a possible screening device for PAD.

**Acknowledgements** This work is supported by the grants: MOST 105-2314-B-182A-130- and 106-2221-E-033-014- from the Ministry of Science and Technology, R.O.C. (Taiwan).

**Conflict of Interest** All contributing authors declare no conflict of interest.

## References

1. Cui W, Kumar C, Chance B (1991) Experimental study of migration depths for the photons measured at sample surface. SPIE 1431:180–191
2. Hiatt WR, Hoag S, Hamman RF (1991) Effect of diagnostic criteria on the prevalence of peripheral arterial disease. The San Luis valley diabetes study. Circulation 91:1472–1479
3. Allison MA, Ho E, Denenberg JO, Langer RD, Newman AB, Fabsitz RR, Criqui MH (2007) Ethnic-specific prevalence of peripheral arterial disease in the United States. Am J Prev Med 32:328–333
4. Willingham Thomas B, Southern W, McCully KK (2016) Measuring reactive hyperemia in the lower limb using near-infrared spectroscopy. J Biomed Opt 21(9):091
5. Wolf U, Wolf M, Choi JH, Levi M, Choudhury D, Hull S, Daniel Coussirat L, Paunescu A, Safonova LP, Michalos A, Mantulin WW, Gratton E (2003) Localized irregularities in hemoglobin flow and oxygenation in calf muscle in patients with peripheral vascular disease detected with near-infrared spectrophotometry. J Vasc Surg 37:1017–1026

# Physiological Monitoring of Cold-Air Stimulated Rhinitis

M.-S. Jhuang, C.-M. Chen, S.-H. Lu, M.-F. Chen, K.-P. Lin, and C.-L. Tsai

**Abstract**

Patients with vasomotor rhinitis accounts for most of the non-allergic rhinitis ones. A wearable measurement system was developed to record multiple physiological signals to record the onset of non-allergic rhinitis. Vasomotor rhinitis symptoms are generally stimulated by a sudden change in ambient temperature. The temperature of two rooms were set at 20 and 30 °C, respectively. Subjects were asked to stay in the warm room then enter the cold room to trigger the nasal congestion. A non-allergic rhinitis subject show very different response from that of a Normal subject. The measurement system is to quantitatively record the stimulation and evaluate the treatment result.

**Keywords**

Vasomotor rhinitis • Optical rhinometry • Nasal swelling

## Introduction

The number of rhinitis patients has greatly increased in recent years. There were more than 60,000 people in Taiwan received medical treatments because of allergy between year 2000 and 2007, and rhinitis patients accounted for 49.8% of them [1]. The morbidity rate of rhinitis also has been continuously rising yearly.

Rhinitis is considered as a result of sever air pollution from overdeveloped industrialization. Although rhinitis is generally stimulated by allergens, the correlation between allergens and rhinitis and could not be found on some patients after allergy test. Such kind of rhinitis is called non-allergic rhinitis (NAR), which accounts for 25% of rhinitis patients [2]. In 2001, a report showed that 52% of patients with allergies were diagnosed to also have NAR [3]. And, vasomotor rhinitis (VMR) accounted for 71% of patients with NAR [4].

VMR symptoms are frequently triggered by the change in environmental temperature. It might happen after going into a cool air-conditioned room during a hot day, and getting down from warm bed in cold morning. When VMR patients are affected by temperature changing, the symptoms include nasal swelling, running nose and sneezing. These symptoms are generally ignored and without doing anything to treat them, because they are not fatal or life threatening. For some patients, the swelling and running nose is quite annoying and greatly affect their daily life. They could lower the patients' work efficiency or distract their attention from study, which decreases their life quality. In addition, it is difficult to know exactly when and how the symptoms would appear making it hard to prevent [5].

Vascular congestion stimulated by low temperature is an important factor that causes nasal swelling [1]. In this study,

M.-S. Jhuang · S.-H. Lu · C.-L. Tsai (✉)
Biomedical Engineering Department, Chung Yuan Christian University, Taoyuan, Taiwan
e-mail: clt@cycu.edu.tw

C.-M. Chen
Otolaryngological Department, Ten-Chen General Hospital, Taoyuan, Taiwan

M.-F. Chen · K.-P. Lin
Electrical Engineering Department, Chung Yuan Christian University, Taoyuan, Taiwan

K.-P. Lin · C.-L. Tsai
Technology Translation Center for Medical Device, Chung Yuan Christian University, Taoyuan, Taiwan



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_20

a wearable multi-physiological signal monitor was designed as a noninvasive measuring system to record nasal swelling due to the change in ambient temperature [6]. This is to help to have better understanding of VMR and provide physicians more information about how to improve rhinitis patients' condition [7].

## Materials and Methods

A. *Optical rhinometry for portable continuous monitoring*

The rhinometry built for long term continuous monitoring of nasal swelling in this study is shown as Fig. 1. The system includes the measurement of light transmission through the nasal bridge and the airflow through both nostrils.

A near infrared light emitting diode (LED) with wavelength of 940 nm is used as the light source of optical rhinometry. Near-infrared has the strongest penetration ability in biological tissue. It can enter from one side of nose, transmit through nasal bridge, and reach the other side of nose to be detected. The transmittance varies mainly with the change in blood perfusion due to the congestion of nasal lining. A quick nasal discharge is also possible when blood pressure increases suddenly. The plethysmograph signal also provides the information of heart rate and its variation. The light transmission changes with the narrowing of airway, and it will have a sudden change in intensity when two walls of airway get in contact during the obstruction. However, it is still difficult to tell whether the airway is really blocked just by the light transmission signal.

A direct measurement of airflow is needed to confirm the obstruction of airway. To access the obstruction of nasal breathing with congestion, three thermistors are placed in front of the two nostrils and the mouth, respectively. The one in front of mouth is to detect whether the subject is breathing through the mouth. The other thermistors were fixed at the centers of 3D-printed barrels placed at nostrils. These tow thermistors are for detecting the airflow through nostrils. If nasal airways are narrowed by the swelling close to the nostrils, airflow will increase according to Bernoulli's principle. The heat dissipation from thermistor will also be increased and lower the sensor temperature. This is much obvious during the breath in period. An accelerometer was attached to the forehead to record the change in subject postures. All these signals are wirelessly transmitted to a personal computer to record for further analysis.

B. *Environmental temperature variations experiment*

During the hot day in summer, it is easy for NAR patients to have nasal obstruction when entering a relatively cold inside place from the hot outside. To have a stable and reproducible ambient temperature simulating such a condition, two neighboring rooms were set at 20 and 30 °C separately. Before the experiment, a subject was asked to sit and rest in the warm room for 10 min to stabilize the heart rate and breathing. At the beginning, the subject kept on staying in the warm room for another 20 min with signals being recorded. Then, the subject was asked to move to the cool room to sit for 30 min, as shown in Fig. 2.

## Results and Discussion

The measurement system records multiple physiological signals. The nasal bridge transmission is a photo- plethysmograph signal varying with the change in blood volume of nose. It also provides the pulse rate signal for calculating heart rate. The change in heart rate is also a sign of the influence of autonomic nervous system on the nasal vasomotor activity which affects blood pressure. Vessel dilation and pressure increase are all possible causes of congestion. However, a continuous pressure monitoring is much difficult.

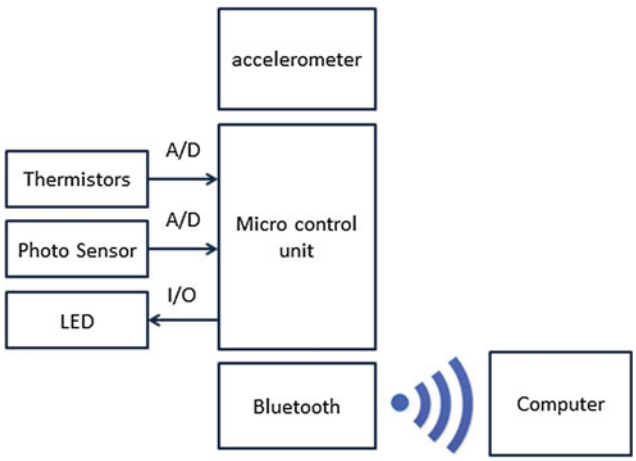



**Fig. 1** Block diagram of physiological signals measurement system

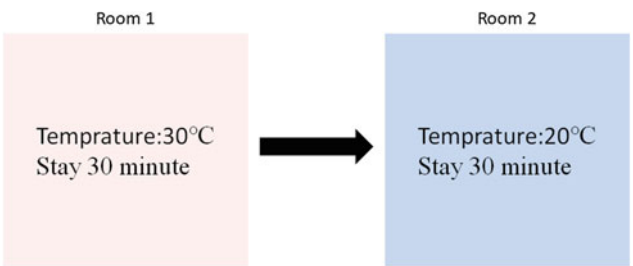



**Fig. 2** Block diagram of experiment rooms

Figure 3 shows the trends of physiological signals obtained from a normal subject. The photo-plethysmograph signal has a stable and low baseline only with small drifting. The temperature at nostrils are also very stable and small in amplitude because the room air temperature is close to body temperature. The air temperature flow through nostril varies between 30 and 37 °C. Moreover, the response time of thermistor limits the variation range down to between 33° and 35°. Since most people do not have symmetry nasal airway, the variation of temperature with breathing on two nostrils are slightly different in amplitude.

Once the subject entered the cold room, the amplitude of temperature variation became much larger. And, the amplitude of right nostril is still larger than that of left nostril. This probably shows the right nasal airway is larger than the left one. On the whole, the variation of temperature is still very stable in the cold room. The increase of heart rate while changing room is caused by the activity of subject. After that, the heart rate become similar to that in the hot room and remain stable. The most obvious change is the up shift of photo-plethysmograph signal which means blood perfusion to nose must be larger in the cold room. This is to compensate for the increasing heat lost by breathing air with lower temperature. The baseline of photo-plethysmograph also shows a lightly upward drifting over the 30 min period of time.

Figure 4 shows the measurement of a non-allergic rhinitis subject. The trends of signal in the warm room are similar to those of the normal subject. The heart rate shows a slightly larger variation during the first 30 min in warm room. When the NAR subject entered the cold room, the signals showed very different physiological response patterns. Not only the temperature at nostrils show an irregular change in amplitude, his heart rate also shows a much larger variation. During the 30 min of siting in cold room, the subject

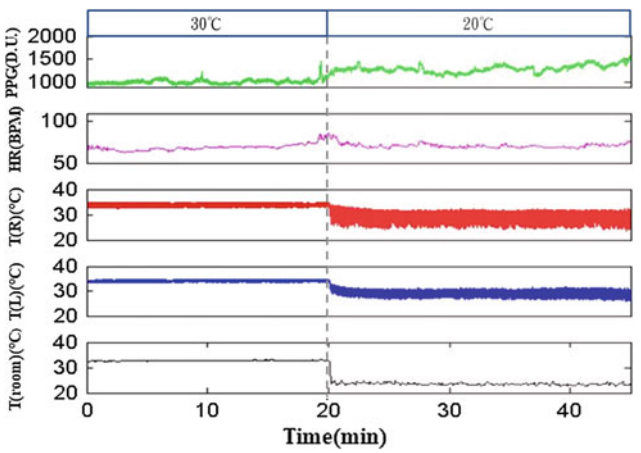



**Fig. 3** Measuring results of a normal subject during the ambient temperature changing experiment

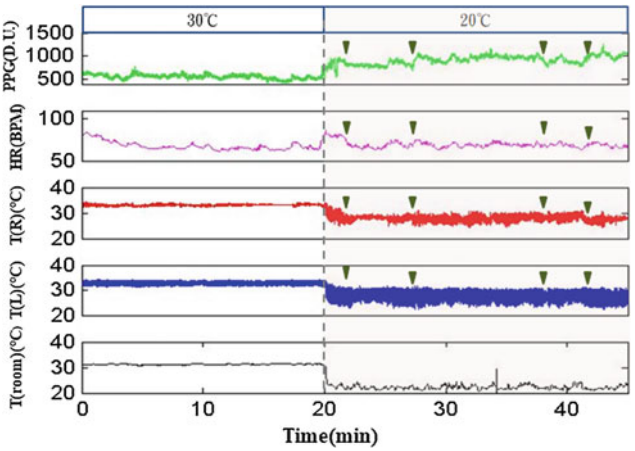



**Fig. 4** Measuring results of a subject with non-allergic rhinitis during the ambient temperature changing experiment

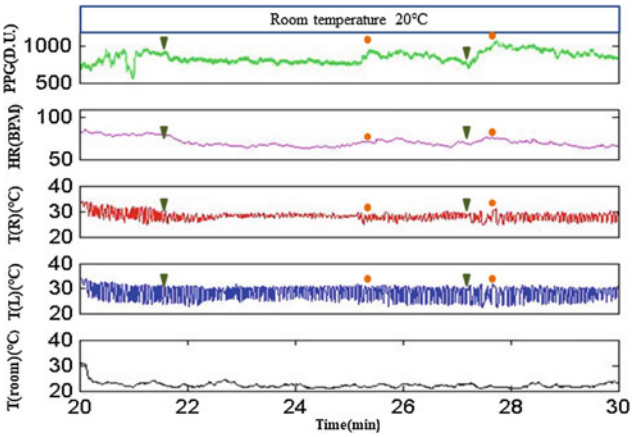



**Fig. 5** The 10 min measurement results of the non-allergic rhinitis subject right after entering the cool room

reported five times of repeated partial nasal airway obstruction and reopen. The time of these occurrences are marked on the trends using black triangles.

The first 10 min period of time after entering the cold room is expanded in Fig. 5 to show more details of signals at the first occurrence. The subject has a narrower right nasal airway, and it is much easier to be blocked by nasal congestion. The nasal airway started to have a first obstruction just about one and half minutes after changing the ambient temperature, as marked by the triangle. The baseline of photo-plethysmograph signal rose before the obstruction, then gradually decreased until the reopen of airway. The figure shows a simultaneous change in heart rate and right nostril temperature. The temperature at left nostril was less affected by the obstruction. The right nasal airway was total blocked, and subject only depended on the left nostril to breath. The small variation of left temperature was caused by

air turbulent near the nostril. The red dot marked the opening of airway where photo-plethysmograph signal showed sudden increase in intensity. Heart rate showed simultaneous increasing with the reopening. After the first reopening, the right nasal airway remains slightly opened. Small amount of airflow can be seen in the figure as the change in temperature amplitude. At the 27th min, the right airway was totally blocked again for a very short period of time. But, it reopened in just a few seconds.

## Conclusion

A multi-parameter continuously monitoring system of nasal congestion and airway obstruction was developed in this study. The measurement system can provide many detail information of the onset of rhinitis. The correlation among different physiological signals will help the physicians to acquire more time course of fast reaction of non-allergic rhinitis in future study.

**Acknowledgements** This work is supported by the grants: MOST 103-2221-E-033-028- from the Ministry of Science and Technology, R. O.C. (Taiwan).

**Conflict of Interest** All contributing authors declare no conflict of interest.

## References

1. Hwang CY, Chen YJ, Lin MW (2010) Prevalence of atopic dermatitis, allergic rhinitis and asthma in Taiwan: a national study 2000–2007. Acta DV 90:589–594
2. Medscape at http://book-med.info/rhinitis/64411
3. Settipane RA, Lieberman P (2001) Update on nonallergic rhinitis. Annals AAI 86:494–508
4. Settipane RA (2009) Epidemiology of vasomotor rhinitis. WAO J 2:115–118
5. Schroer B, Pien LC (2012) Nonallergic rhinitis: common problem, chronic symptoms. Cleve Clin J Med 79:285–293
6. Hampel U, Schleicher E, Wustenberg EG, Huttenbrink KB (2004) Optical measurement of nasal swellings. IEEE Trans Biomed Eng 51:1673–1679
7. Lu SH, Tsai CL, Lin KP (2016) Optical rhinometry for long term monitoring. In: XIV Mediterranean conference on medical and biological engineering and computing, pp 762–765. https://doi.org/10.1007/9783319327037260

# Association Between $SpO_2$ Signal Characteristics and Sleep Architecture with Insulin Resistance in Patients with Obstructive Sleep Apnea Syndrome

E. Perantoni, P. Steiropoulos, D. Filos, N. Maglaveras, K. Nikolaou, and I. Chouvarda

**Abstract**

Obstructive Sleep Apnea Syndrome (OSAS) may contribute to the increasing frequency of metabolic disorders. Intermittent hypoxia (IH) is a major characteristic of the syndrome. However, the existing indices of hypoxia in sleep cannot express accurately the effect of the mild desaturations. In this study, a total of 51 patients without other comorbidities were examined by polysomnography (PSG). Hypoxia parameters were analyzed, in the intervals with low values of $SpO_2$ signal. More specifically, the thresholds were set at 94 and 92% and the average value (M) of the $SpO_2$ signal, in areas below thresholds, were calculated. Moreover, the desaturations were analyzed, together with their duration within the recording in terms of $SpO_2$ signal parameters. The patients' blood sample was analyzed for metabolic parameters. In total, 28 individuals were diagnosed with severe OSAS, (Apnea Hypopnea index (AHI) 59.11 $\pm$ 26.10/h, aver$SpO_2$ 91.64 $\pm$ 4.50%, min$SpO_2$ 78.18 $\pm$ 10.26%, t < 90 21.42 $\pm$ 28.64 and ODI 35.48 $\pm$ 33.79/h). A statistically significant correlation between the average M92 value with insulin levels ($r = 0.401$, $p < 0.03$) and homeostasis model assessment (HOMA) ($r = 0.431$, $p < 0.022$) was displayed. Likewise, a correlation between the amount of desaturations and fasting glucose ($r = 0.400$, $p < 0.035$) was observed. Moreover, a statistically significant correlation between the desaturations' average value with insulin ($r = 0.378$, $p < 0.047$) and CRP ($r = 0.400$, $p < 0.035$) levels was also revealed. A strong correlation also emerged from the cumulative desaturations' duration as recorded by the $SpO_2$ signal with fasting glucose levels ($r = 0.964$ $p < 0.001$), glycosylated hemoglobin ($r = 0.860$, $p < 0.000$) and HOMA index ($r = 0.580$, $p < 0.001$). The results suggest that the Hypoxia factors derived from $SpO_2$ signal analysis, are strongly correlated with the insulin resistance and with fasting glucose levels. The correlations of the proposed hypoxia parameters were found to be stronger than the already known hypoxia indices, deriving from the PSG, however a more extended analysis is necessary in order to consolidate the findings of this study.

**Keywords**

Hypoxia • Obstructive sleep apnea syndrome • Metabolic syndrome • Insulin resistance Signal analysis

E. Perantoni (✉) · D. Filos · N. Maglaveras · I. Chouvarda
Lab of Computing, Medical Informatics & Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Aristotle University Campus, 54124 Thessaloniki, Greece
e-mail: elperantoni@gmail.com

P. Steiropoulos
MSc Program in Sleep Medicine, Medical School, Democritus University of Thrace, Alexandroupolis, Greece

K. Nikolaou
Pulmonary Clinic, G.H. G. Papanikolaou, Thessaloniki, Greece

## Introduction

It has already been reported that Obstructive Sleep Apnea Syndrome (OSAS) is characterized by recurrent episodes completely of partial obstruction of the upper airways. It is characterized by hypoxic episodes alternating with normal



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_21

oxygenation and by arousals aiming at restoring airflow which result in fragmentation of sleep. Obesity is considered as a major risk factor for the development and progression of OSAS. The incidence of the syndrome in obese or severely obese patients is almost twice than adults with normal weight [1]. The deposition of fat around the chest reduces chest compliance and functional residual capacity, conditions that can lead to increased oxygen demand [2]. The visceral obesity is considered a common risk factor for the onset of diabetes and OSAS.

There is a rapid increase in recent years of both obesity and diabetes, which has been attributed to environmental, socioeconomic, demographic factors as well as lifestyle changes [3]. Many clinical studies suggest that up to 50% of patients with type 2 diabetes have moderate to severe OSAS. Most patients with diabetes report insufficient sleep, both in duration and quality. The prevalence of sleep apnea in patients with type 2 diabetes is estimated at 2–70%. A remarkable finding of the study by Foster et al. was the high prevalence of undiagnosed OSAS (86.6%) among obese patients with type 2 diabetes [4]. This report requires certainly further study on the link between diabetes and sleep apnea.

Intermittent hypoxia (IH), characterized by small repetitive desaturation cycles followed by rapid re-oxygenation, is a feature of the OSAS. Many researchers have developed experimental models in animals and at cellular level in order to study the role of IH. Studies in normal and obese mice, has shown that chronic exposure to IH contributes to decreased insulin sensitivity, a result that is partially reversible after stimulus discontinuation [5].

Nocturnal oxyhemoglobin desaturation and OSAS severity has been associated with metabolic parameters including higher fasting insulin, glucose levels and impaired insulin sensitivity [6].

The present study aims to explore whether parameters derived from the $SpO_2$ signal during sleep can be related to insulin resistance, with the perspective to generate hypoxia markers that can predict diabetes mellitus in the future.

## Materials and Methods

### Data Description

The data for this work were collected in one year period, following all ethical procedures. Totally, 51 patients without known comorbidities, who had been referred to the sleep lab of the G.H. "G. Papanikolaou", Thessaloniki, Greece with symptoms suggesting sleep-related breathing disorders were included in the study. All patients had undergone clinical and laboratory testing. A full polysomnography (PSG) was performed for the assessment of their sleep. Finally, 28 patients were diagnosed with OSAS and 23 were found to have normal breathing function in sleep. A blood sample was collected after at least 8 h of overnight fasting. The following parameters were examined: high sensitivity C-reactive protein, fibrinogen, homocysteine, complete lipid profile, blood glucose as well as glycosylated hemoglobin, insulin levels and HOMA (Homeostasis model assessment) which is an insulin resistance marker. The characteristics of the subjects are displayed in Table 1.

**Table 1** Patients' demographics and standard $SpO_2$ parameters (mean ± std)

| | OSAS | Controls |
|---|---|---|
| N | 28 | 23 |
| Age | 47.14 ± 11.543 | 46.00 ± 3.69 |
| BMI | 33.79 ± 6.42 | 28.61 ± 3.29 |
| AHI | 59.11 ± 26.10 | 3.47 ± 4.6 |
| ESS | 11.32 ± 5.93 | 5.5 ± 4.64 |
| $AverSpO_2$ | 89.87 ± 5.20 | 93.80 ± 2.002 |
| $minSpO_2$ | 71.36 ± 9.03 | 86.47 ± 2.82 |
| t < 90 | 34.44 ± 27.96 | 5.56 ± 20.65 |
| ODI | 60.46 ± 25.90 | 5.08 ± 3.42 |

The $SpO_2$ signals were extracted from the night PSG recording, in standard EDF format. All signals had a sampling frequency of 3 Hz. Figure 1 depicts an example $SpO_2$ signal of a PSG study of an OSAS patient with AHI 42/h. In this figure, many mild and some serious desaturations occur throughout the night.

### $SpO_2$ Signal Analysis

The $SpO_2$ signal was processed in Matlab following the procedure described below:

1. Signal preprocessing to remove artifacts and improve signal.

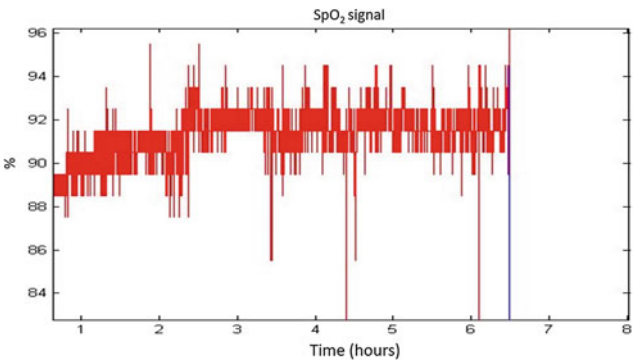



**Fig. 1** The $SpO_2$ whole night time series (6 h) of a subject with AHI 42/h

2. Calculation of hypoxia parameters: setting the threshold to 92 and 94%, and considering the signal values below the thresholds, statistical parameters were calculated for this set up of values. The parameters M92 and M94 refer to the calculated average values, with the respective thresholds employed.
3. Detection of desaturation events (Dev), and related parameters. Desaturation events refer to areas with continuous values below threshold, characterized by the start, end and duration of the event. The number of Dev and their total duration at the time of recording, the Dev duration statistics (mean, standard deviation, min and max), and the percentage of Dev that were long than 6 points (2 s) were calculated as candidate markers.

## Statistical Analysis

A statistical analysis of the calculated parameters was performed in order to investigate potential correlation between the characteristics of the signal and the insulin resistance.

Pearson correlation coefficient was used, between the proposed newly introduced hypoxia markers and metabolic markers, and considering the already used hypoxia markers as measured by PSG.

## Results

As previously reported, a total of 51 patients with an average age of 46.6 ± 8.84 years were studied. The patients were in general obese with a body mass index (BMI) 31.45 ± 5.81.

During the first interview they reported sleepiness during the day, restless sleep, intense snoring and episodes of breathing during their sleep, recorded by relatives who accompanied them on the day of the interview. They all went through a full PSG study, and on average had AHI 34.02 ± 33.93/h. Blood sampling was performed after overnight fasting of at least 8 h as reported. 28 patients were diagnosed with AHI 59.11 ± 26.10/h, while the other 22 patients had normal sleep with AHI/h of 3.47 ± 0.6.

The blood sample values are depicted in Table 2. There was a statistically significant correlation between mean M92 with insulin levels ($r = 0.401$, $p < 0.035$) and HOMA ($r = 0.431$, $p < 0.022$), as well as the amount of desaturations with fasting glucose ($r = 0.400$, $p < 0.035$). In addition, there was a statistically significant correlation between mean duration of desaturations with insulin ($r = 0.378$, $p < 0.047$) and CRP ($r = 0.400$, $p < 0.035$). A strong correlation was also revealed between the total duration of desaturation of SpO2 with fasting glucose ($r = 0.964$, $p < 0.000$), glycosylated hemoglobin ($r = 0.860$, $p < 0.000$) and HOMA ($r = 0.580$, $p < 0.001$) (Table 3).

**Table 2** Blood samples results (mean ± std)

| | OSAS | Controls |
|---|---|---|
| Glucose | 112.71 ± 43.33 | 81.82 ± 10.43 |
| Insulin | 17.34 ± 13.88 | 6.10 ± 3.40 |
| HbA1c | 5.93 ± 1.50 | 5.12 ± 0.27 |
| HOMA | 5.03 ± 4.92 | 1.3 ± 0.91 |
| hsCRP | 0.54 ± 0.54 | 0.28 ± 0.38 |
| Fibrinogen | 7.89 ± 15.98 | 2.31 ± 0.86 |
| Cholesterol | 242.75 ± 44.45 | 223.95 ± 44.53 |
| Triglycerides | 228 ± 194.91 | 149.04 ± 66.05 |

**Table 3** Hypoxia markers OSAS group

| | | INS | $HbA_{1c}$ | GLU | HOMA | hsCRP |
|---|---|---|---|---|---|---|
| Time in event | p | | 0.000 | 0.000 | 0.001 | |
| | r | | 0.860 | 0.964 | 0.580 | |
| Mean | p | 0.047 | | | | 0.035 |
| | r | 0.378 | | | | 0.400 |
| Percentage in event | p | 0.032 | | | | |
| | r | 0.406 | | | | |
| Events | p | | | 0.035 | | |
| | r | | | 0.400 | | |
| M92 | p | 0.035 | | | 0.022 | |
| | r | 0.401 | | | 0.431 | |
| Percent-92 | p | 0.026 | | | | |
| | r | 0.419 | | | | |

In addition, a statistical analysis was performed between metabolic indices and markers of hypoxia deriving from the PSG. A statistically significant positive correlation between $t < 90$ and insulin levels ($r = 0.496$, $p < 0.007$), HOMA ($r = 0.432$, $p = 0.022$) was found, while a negative correlation between insulin and AverSpO2 ($r = -0.552$, $p = 0.022$) and the minSpO2 ($r = -0.486$, $p < 0.009$) was observed. A negative correlation was also found between AverSpO2 and HOMA ($r = -0.379$, $p < 0.047$).

An additional statistically significant correlation was found between minSpO2 and CRP ($r = 0.400$, $p < 0.035$).

## Discussion and Conclusion

There is lack of studies on the effect of IH on insulin resistance and on the development of diabetes mellitus. Most researchers in their attempt to correlate IH with insulin resistance have used hypoxia markers from PSG, such as the

average saturation of oxyhemoglobin- Aver$SpO_2$, the minimum $SpO_2$-min$SpO_2$, time below 90% −t < 90 and the ODI-Oxygen Desaturation Index.

ODI is the hourly average number of desaturation episodes, which are defined as at least 4% decrease in saturation from the average saturation in the preceding 120 s, and lasting 10 s.

The purpose of the present study is to analyze the $SpO_2$ signal from the PSG, in order to investigate the correlation between nocturnal hypoxia and metabolic disorders in otherwise patients. There is lack of studies in literature on the analysis of PSG signals (except EEG) and even less is the data regarding the $SpO_2$ signal. All patients in this study were obese with BMI 31.45 ± 5.81 and at a pre-diabetic stage. 28 of them were diagnosed with OSAS with an AHI 47.14 ± 11.54/h.

Hypoxia markers derived from $SpO_2$ signal analysis that are used in this study, are in agreement with the existing evidence that supports the effect of IH on insulin resistance and on glucose metabolism. The IH index currently employed is ODI. In the present study no statistically significant correlation was found between ODI and the HOMA index, insulin and fasting glucose, but significant correlation was found between Aver$SpO_2$, t < 90 and min$SpO_2$ with the same metabolic markers [7].

Muraki et al. [8] studied the effect of IH in patients with OSAS in the highly sensitive C-reactive protein. The index that was used as a marker to observe the intermittent hypoxia was the ODI. The severity of nocturnal IH was determined by 3% ODI levels. ODI scores of 5–15 episodes/h, corresponding to mild, moderate to severe intermittent hypoxia, were found to be associated with elevated levels of CRP. The results of the study are in agreement with the results of our study. An important correlation was found between CRP and the mean $SpO_2$ signal desaturations.

As regards analysis of the $SpO_2$ signal, in a recent study of Moret-Bonillo et al. [9] analysis of $SpO_2$ and respiratory (airflow, abdominal and thoracic signals) signals was performed, for the detection and quantification of respiratory pauses in the patient's respiratory activity and the generation of diagnostic patterns of sleep apnea syndrome. In a pilot study of Renata Trimer et al., the $SpO_2$ signal variability ($SpO_2$V) derived from $SpO_2$ intervals ($SpO_2$i) between successive $SpO_2$ values of the pulse wave was studied. The aim was to use the signal variability $SpO_2$V as a predictor marker of sleep apnea syndrome [10]. Alvarez et al. analyzed the $SpO_2$ signal using central tendency measure (CTM), measures indicating the point to which values of a data/observations group tend to accumulate, and is considered the "center" of the distribution of observations. The result of their study was that the CTM can be used as a reliable diagnostic tool for the sleep apnea syndrome [11].

General research on the analysis of signals from PSG and specific $SpO_2$ signal, revolves around finding the most reliable indicators for diagnosis of OSAS [12, 13].

This study proposed new markers of IH. The results of this study are in agreement with the data from the international literature about the role of IH in glucose metabolism and insulin resistance. In addition, it is important to mention that the new hypoxic indicators employed in this work, presented a stronger correlation to metabolic indices, than the hypoxic indicators so far used by PSG. These findings can lead to better utilization of $SpO_2$ data and help in the prognosis and treatment of metabolic syndromes by addressing their causes.

**Acknowledgements** We want to thank the sleep lab of the G.H. "G. Papanikolaou", Thessaloniki, Greece, for the data collection.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Romero-Corral A, Caples SM, Lopez-Jimenez F, Somers VK (2010) Interactions between obesity and obstructive sleep apnea implications for treatment. Chest 137(3):711–719. https://doi.org/10.1378/chest.09-0360
2. Malhotra A, White DP (2002) Obstructive sleep apnoea. Lancet 360(9328):237–245. https://doi.org/10.1016/S0140-6736(13)60734-5
3. Touma C, Pannain S Does lack of sleep cause diabetes? Cleve Clin J Med 78:549–58
4. Foster GD, Sanders MH, Millman R, Zammit G, Borradaile KE, Newman AB et al (2009) Obstructive sleep apnea among obese patients with type 2 diabetes. Diabetes Care 32:1017–1019. https://doi.org/10.2337/dc08-1776
5. Drager LF, Li J, Reinke C et al (2011) Intermittent hypoxia exacerbates metabolic effects of diet-induced obesity. Obesity (Silver Spring) 19:2167–2174. https://doi.org/10.1038/oby.2011.240
6. Moon K, Naresh M, Punjabi, NM, Rashmi N, Aurora RN (2015) Obstructive sleep apnea and type 2 diabetes in older adults. Clin Geriatr Med 31(1):139–ix. https://doi.org/10.1016/j.cger.2014.08.023
7. Punjabi NM, Shahar E, Redline S, Gottlieb DJ, Givelber R, Resnick HE (2004) Sleep-disordered breathing, glucose intolerance, and insulin resistance: the sleep heart health study. Sleep heart health study investigators. Am J Epidemiol 15, 160(6):521–530. https://doi.org/10.1093/aje/kwh261
8. Muraki I, Tanigawa T, Yamagishi K, Sakurai S, Ohira T, Imano H, Kitamura A, Kiyama M, Sato S, Shimamoto T, Konishi M, Iso H (2010) Nocturnal intermittent hypoxia and C reactive protein among middle-aged community residents: a cross-sectional survey. CIRCS Invest Thorax 65(6):523–527. https://doi.org/10.1136/thx.2009.128744
9. Moret-Bonillo V, Alvarez-Estévez D, Fernández-Leal A, Hernández-Pereira E (2014) Intelligent approach for analysis of respiratory signals and oxygen saturation in the sleep apnea/hypopnea syndrome. Open Med Inform J 8:1–19. https://doi.org/10.2174/1874431101408010001

10. Mendes RG, Delfino A, Oliveira JR, Arena R, Cabiddu R, Borchi-Silva A (2016) Correlation between desaturation indices of oxygen saturation variability in severe obstructive sleep apnea :a pilot study. J Resp Cardiov Phy Ther 4(1):3–11
11. Alvarez D, Hornero R, García M, del Campo F, Zamarrón C (2007) Improving diagnostic ability of blood oxygen saturation from overnight pulse oximetry in obstructive sleep apnea detection by means of central tendency measure. Artif Intell Med 41(1):13–24. (Epub 2007). https://doi.org/10.1016/j.artmed.2007.06.002
12. Koley BL, Dey D (2014) On-line detection of apnea/hypopnea events using $SpO_2$ signal: a rule-based approach employing binary classifier models. IEEE J Biomed Health Inform. 18(1):231–239. https://doi.org/10.1109/JBHI.2013.2266279
13. Garde A, Dehkordi P, Wensley D, Ansermino JM, Dumont GA (2015) Pulse oximetry recorded from the phone oximeter for detection of obstructive sleep apnea events with and without oxygen desaturation in children. Conf Proc IEEE Eng Med Biol Soc 7692–7695. https://doi.org/10.1109/EMBC.2015.7320174

# Part V
# Biosignal Analysis Methods

# Preprocessing and Filtration Techniques of BSPM Signals in a Small-Scale Study

M. Hrachovina, L. Lhotská, and M. Huptych

**Abstract**

Cardiac resynchronization therapy (CRT) is an accepted therapeutic option in heart failure treatment. Before the treatment it is necessary to have sufficient information about electrical and mechanical dyssynchrony at each patient. In a joint project of CTU and University Hospital Motol in Prague, data from several modalities are collected and analyzed. In the paper we focus on data preprocessing from multichannel ECG. In particular, we describe in detail the filtration step, as the original signals were contaminated by various artefacts and noise. The aim is to have the signals as clean as possible and preserving the useful information for the analysis.

**Keywords**

BSPM • Filtration • Preprocessing • ECG mapping

## Introduction

Cardiac resynchronization therapy (CRT) is an accepted therapeutic option in heart failure treatment clinical effect of which was demonstrated in multiple multicenter trials, similarly as its positive impact on morbidity and mortality [1–3]. Currently, CRT is indicated in form of so-called biventricular pacing in patients with symptomatic chronic heart failure (functional class NYHA II—ambulatory IV) with symptoms lasting despite optimized pharmacological treatment and concurrent presence of systolic dysfunction of the left ventricle with an ejection fraction $\leq 35\%$ [4]. In a joint project with the Department of Cardiology of the University Hospital Motol Prague, we focus on data analysis. Both electrical and mechanical dyssynchrony are studied. Electrical dyssynchrony is studied and quantified from ECG body surface mapping data; in selected patients also from endocardial electrical mapping of the left ventricle using electroanatomical mapping system CARTO. Recent studies [5, 6] have also addressed this topic, but their preprocessing and filtration techniques were either not described or they were not reproducible on our data. Therefore, in the paper we describe an important step in multichannel ECG processing, namely signal filtration, which proved to be crucial for the next steps in signal analysis.

## Background

In the project aimed at assessing the effectiveness of Cardiac Resynchronization Therapy (CRT), Body Surface Potential Mapping (BSPM) is one of the methods chosen to help quantify the therapy progress. BSPM is an extension of the conventional electrocardiography that provides refined

M. Hrachovina (✉)
Department of Cybernetics, FEE CTU Prague, Karlovo náměstí 13, 121 35 Prague 2, Czech Republic
e-mail: matej.hrachovina@gmail.com

L. Lhotská
Faculty of Biomedical Engineering, Czech Technical University in Prague, Prague, Czech Republic

M. Huptych
Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Prague, Czech Republic



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_22

non-invasive characterization of cardiac activity. The data acquisition systems may use 24–300 electrodes. The acquired signals are used as input data in BSPM. Increased spatial sampling on the body surface provides more in-depth information on the potentials generated by the heart, and thus in many cases exhibits better diagnostic value with regard to spatial and temporal information of the electrical activity of the heart [1, 4]. The BSPM has two major advantages over standard 12-leads ECG: (1) it allows exploring the entire chest surface, thus providing all the information on the cardiac electric field available at the body surface; (2) it is more sensitive in detecting local electrical events, such as local conduction disturbances or regional heterogeneities of ventricular recovery. However, the multichannel ECG has its drawbacks as well, e.g. more complicated measurement and electrode positioning. The most frequently used body surface maps are isopotential, giving a distribution of the potential at a specific moment, and isointegral, providing a distribution of the sum of potentials over a specified time interval.

## Aim of Filtration

BSPM was traditionally used for obtaining potential maps and mapping propagation of electrical signal along chest and back. Recently, the idea of using BSPM instead of invasive catheter diagnosis for finding early activation sites of the cardiac muscle is becoming popular again. In both of these applications, the quality of signals is crucial for labeling the ECG and extracting feature timing. Therefore the aim of filtration is to cancel as much noise as possible without altering the ECG components (mainly Q, R, S and T waves).

The main sources of noise in our case are mains hum, motion artefacts and bad electrode contact. The electrode—skin impedance varies slightly during any ECG measurement, but the changes are usually very slow and have negligible effect on the measured signal. If an electrode does not have good contact with the skin, its impedance changes dramatically and the amplitude of the measured signal changes accordingly, but with no correspondence to the ECG signal [7]. If we do not record the impedance of the electrode, there is no way of reconstructing the signal from the lead where electrode impedance is unstable. Unfortunately, it is very common in BSPM measurement that some electrodes do not stick to the skin well enough and loose contact unpredictably. It is impossible to monitor the electrodes constantly.

## Preprocessing

### A. *Wilson's central terminal*

BSPM Signals are recorded from unipolar leads referenced against Wilson's central terminal, which is formed by the mean of the bipolar limb signals.

$$V_{WCT} = \frac{(RA + LA + LL)}{3}$$

The raw signals are saved referenced to a virtual floating reference though, so the WCT signal needs to be subtracted from each lead. The referencing effect is visualized on Fig. 1.

### B. *DC offset*

The recorded signals suffer from DC offset induced during amplification. The easiest and most commonly used method of eliminating DC offset during post processing is subtracting the mean of the signal itself. DC offset normalization is outlined in Fig. 2.

## Filtration

### A. *Motion artefacts and isoline removal*

ECG signals suffer from baseline wander caused by breathing or (slow) movement. The induced noise is a low frequency signal under 1 Hz (AHA guidelines state 0.67 Hz threshold). This noise can be removed by filtration using high pass filters. There has to be a tradeoff between filter strength, signal distortion and filtering duration. The higher the order of the filter, the better the filtration result, but the longer it takes to filter the signal. Also, the type of filter is important, IIR filters need lower order to achieve good attenuation, but their phase is not linear, which distorts the ECG signal. Phase distortion is crucial for ECG, so FIR filters are preferred (although they need higher order and thus take longer time to filter).

But even filters with linear phase characteristics can cause signal distortion as shown in [8]. It shows that even when using cutoff frequency of 0.67 Hz as according to guidelines, the higher frequency components of the QRS complex are

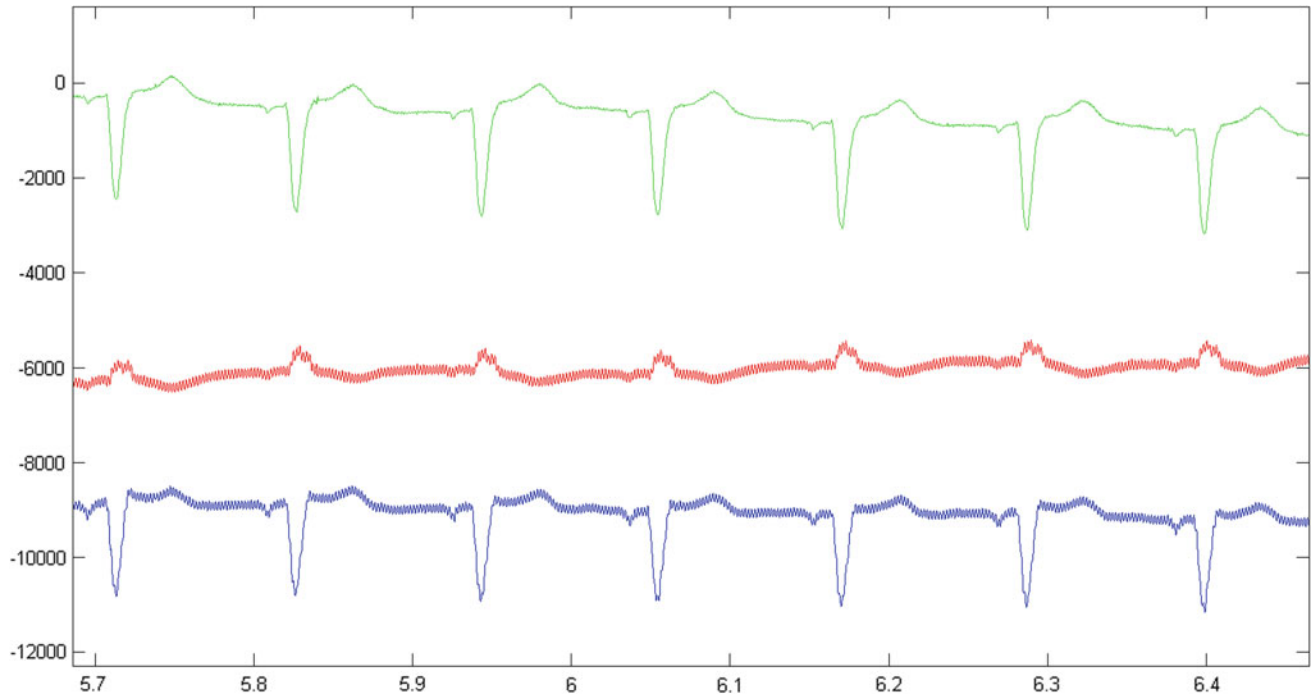



**Fig. 1** Referencing signal against WCT. Blue—Unipolar lead signal. Red—WCT signal. Green—referenced signal

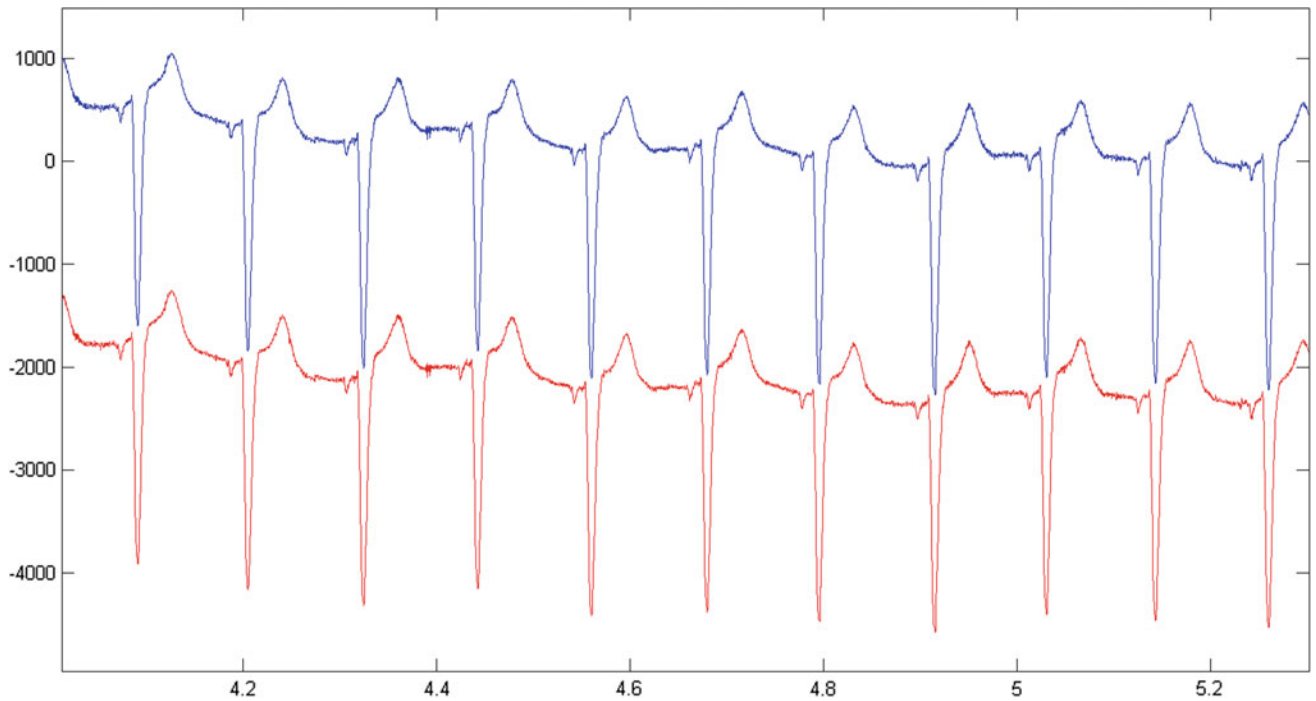



**Fig. 2** DC offset. Red—signal with DC offset. Blue—signal without DC offset

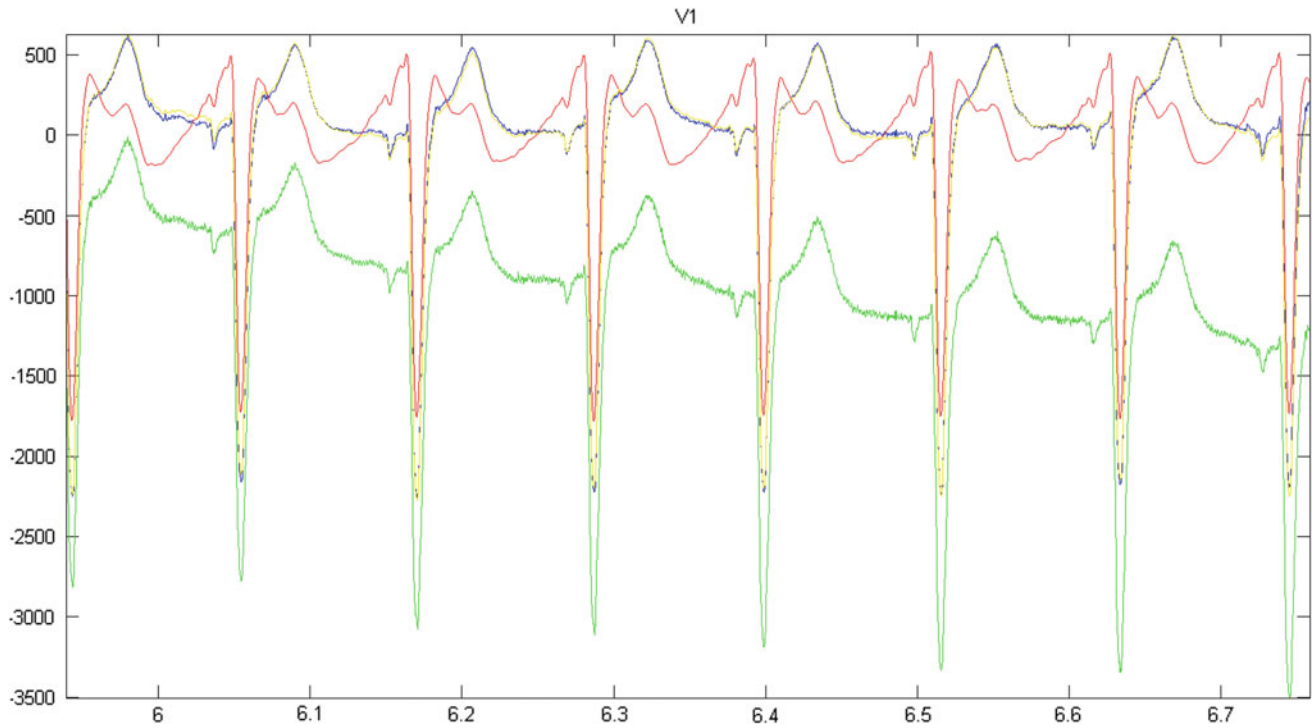



**Fig. 3** Comparison of different filtration techniques. Green—unfiltered signal. Red—isoline removed using FIR highpass, cutoff 0.67 Hz; Q, S and T waves visibly distorted. Blue—isoline removed using FIR highpass, cutoff 0.05 Hz; morphology unchanged, but baseline drift not removed completely

phase shifted and superpone to the S wave, thus causing S wave exageration.

To get rid of motion artefacts and baseline drift we use a sequence of a low-pass and a high-pass filters. The reason for applying them separately rather than using a band-pass filter is that we can use filters of lower order.

Low-pass filter is a FIR filter with order 10000, and cutoff frequency 0.05 Hz. For filtering out motion artefacts we use a high-pass FIR filter with order 500 and cutoff frequency 100 Hz.

For filtering out baseline drift we first estimate the baseline using a low-pass FIR filter with order 1000 and cutoff frequency 0.67 Hz, then we smooth this signal using a 1500-point moving average FIR filter. We then subtract the baseline signal from the original signal.

The effects of different filter settings on ECG signal are outlined in Fig. 3.

### B. *Mains hum filtration*

Removing such a narrow band signal effectively using a FIR filter is impossible, so we use an IIR biquad notch filter. It is centered at 50 Hz with 3 dB bandwidth of 6 Hz.

If we assume that all other noise is random with even probability distribution and that the signal is relatively stationary, i.e. the beats are very similar during recording, and that we only need a single beat of the resulting processed signal, we can use averaging of the signal to enhance its quality.

To do this, the individual beats in the signal must first be annotated. We use R peak detection for now, but we are testing more sophisticated methods to be used in the future. We also need to check that the beats we choose for averaging are similar. We test the R-R intervals, which must be within $\pm$ 5% of the median R-R interval of the whole recording and the interbeat R-R variability must not exceed 3% of the median.

The process of enhancing signal using averaging is outlined in Fig. 4.

### C. *Electrode contact loss detection*

Samples where skin-electrode impedance instability ruined the signal must not get into further processing, so they do not bias the propagation maps. We are currently searching for methods for recovering such signals.

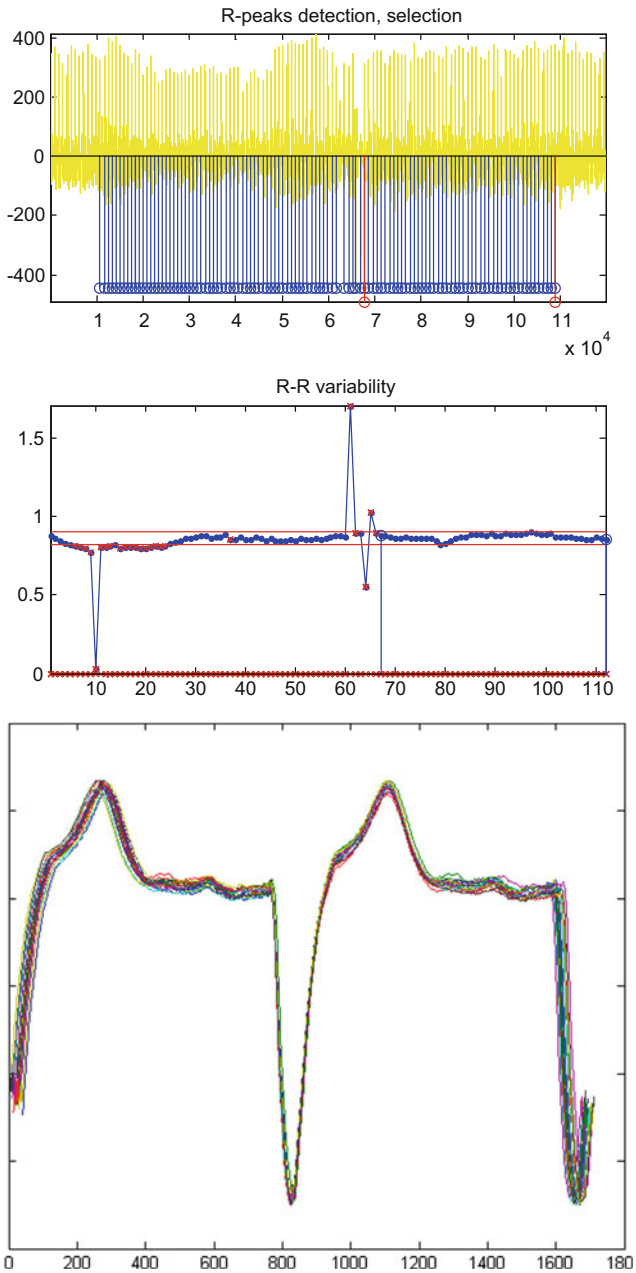



**Fig. 4** Enhancement of signal using averaging. Up: R-peak detection; Middle: R-R variability assessment; Down: individual beats alignment respective to R peak

## Results

The accuracy and effect of the methods on the resulting maps is hard to assess quantitatively. What we can test is the difference between the labeled signal features. However at the time being we do not have enough patient data to compare. The problem is complicated with the fact that the ECG waveforms are pathological and we cannot use simple automatic waveform labelling methods, but they need to be labelled by experienced physicians. However, current results prove that the designed experimental protocol and preprocessing methods are correct and we can acquire rather unique data.

## Conclusions

Multichannel measurement and data collection represent a complex task, in particular in a limited case study. Thus data quality is of high importance. During the measurement many aspects must be considered, in particular electric noise, noise induced from breathing, patient height influencing electrode placement, and proper electrode-skin contact. With the high number of electrodes, data collection is rather difficult task. The next step—data preprocessing is similarly challenging because the artefacts and noise vary in individual channels. Thus one method cannot be applied automatically to all channels. We tried to present in the paper individual steps, which must be performed in the preprocessing phase to acquire useful signals in all channels, or at least most of the channels.

Although the methods were not evaluated quantitatively yet, they have been judged sufficient for accurate labelling of the signals by the physicians that were working on them. The methods have been used in preparing data for medical publications that are in reviewing process at this time.

## Statement of Informed Consent

The study protocol and patient informed consent have been approved by the University Hospital Motol ethical committee.

## Protection of Human Subjects and Animals in Research

The procedures followed were in compliance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects.

**Acknowledgements** The research is supported by the project No. 15-31398A "Features of Electromechanical Dyssynchrony that Predict Effect of Cardiac Resynchronization Therapy" of the Agency for Health Care Research of the Czech Republic and by CVUT institutional resources (SGS grant application No. OHK4-019/17).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Abraham WT, Fisher WG, Al Smith et al (2002) Cardiac resynchronization in chronic heart failure. N Engl J Med 346 (24):1845–1853
2. Cleland JG, Daubert JC, Erdmann E et al (2005) The effect of cardiac resynchronization on morbidity and mortality in heart failure. N Engl J Med 352(15):1539–1549
3. Bristow MR, Saxon LA, Boehmer J, et al (2004) Comparison of Medical Therapy, Pacing and Defibrillation in Heart Failure Investigators. Cardiac resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. N Eng J Med 350:2140–2150
4. Brignole M, Auricchio A, Baron-Esquivias G, et al (2013) ESC guidelines on cardiac pacing and cardiac resynchronization therapy: the task force on cardiac pacing and resynchronization therapy of the European Society of Cardiology (ESC). Developed in collaboration with the European Heart Rhythm Association (EHRA). Eur Heart J 34:2281–2329
5. Giffard-Roisin ST, Jackson L, Fovargue et al (2017) Noninvasive personalization of a cardiac electrophysiology model from body surface potential mapping, IEEE Trans Biomed Eng 64(9):2206–2218
6. Wang Y, Yoram R (2006) Application of the method of fundamental solutions to potential-based inverse electrocardiography. Ann Biomed Eng 34(8):1272–1288
7. Taji B, Shirmohammadi S, Groza V, Batkin I (2014) Impact of skin-electrode interface on electrocardiogram measurements using conductive textile electrodes. IEEE Trans Instrum Meas 63 (6):1412–1422
8. Buendía-Fuentes F, Arnau-Vives MA, Arnau-Vives A et al (2012) High-bandpass filters in electrocardiography: source of error in the interpretation of the ST segment, ISRN cardiology, vol. 2012, pp 706217

# Blood Vessel Segmentation from Microcirculation Images

Bea Lyn M. Virtudazo, Jimmy Hasugian, Wen-Chen Lin, Mei-Fen Chen, and Kang-Ping Lin

**Abstract**

Image segmentation is one of the most important steps in analysis of features in image data. In order to acquire stable image of blood vessel, frame-to-frame matching as a step in preprocessing has be made in solving the problem of motion throughout image acquisition. The purpose of this study is to extract the information of the blood vessel from microcirculation video images for the uses in further processes. This image segmentation process had done by dividing the image into two parts, which are the information of the blood vessel and the background. Through the frame-to-frame image intensity pattern based feature analysis and extraction of blood vessel in microcirculation images had used in obtaining the information needed for segmentation. The results showed solid and consistent with the perspective checking result. These segmented blood vessel images has applied for the further measurement of blood flow velocity.

**Keywords**

Image segmentation • Microcirculation • Blood vessel • Feature extraction

## Introduction

Microcirculation is important in analyzing the local tissue situation of subjects. The condition of microcirculation is responsible for regulation of blood flow in individual organs and for exchange between blood and tissue [1]. This system consists of the heart and the blood vessels running through the entire body [2]. Blood vessels are complex in structure. There are three major types of blood vessels: the arteries, vessels up to ~100 μm inner diameter in arterial system which carries the blood away from the heart to a higher physiologic pressures, the capillaries, the smallest vessel of 4–8 μm inner diameter that enables the actual exchange of water and chemicals between the blood and the tissues, and the veins, vessels up to ~100 μm inner diameter carries the blood from the capillaries back toward the heart at lower physiologic pressures [3, 4]. Moreover, microcirculatory study is important for understanding the amount of the blood supplied in the local tissues [5].

There are numerous methods in segmentation over years. Blood vessel segmentation is one of the important procedure in analysis of microcirculation images. This should be acquired in order to analyze the flow of red blood cells within the vessel [6, 7]. One of the problems in image segmentation by processing from video images is the movement of the focused object during image acquisition. Frame-to-frame matching of microcirculation images as a preprocessing step is a solution for this problem [8]. Also, the obtained images were degraded by some random noise that occurs during acquisition or transmission. An efficient filter has been expected in eliminating or limiting the noise presenting in images.

The aim of this study is to extract a feature of blood vessel from microcirculatory images obtained in sequence images of video. Using the frame-by-frame image intensity

B. L. M. Virtudazo · J. Hasugian · W.-C. Lin · M.-F. Chen
K.-P. Lin (✉)
Department of Electrical Engineering, Chung Yuan Christian University, Cheng-Pei Rd, Taoyuan, Taiwan
e-mail: kplin@cycu.edu.tw



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_23

feature from the information acquired in the blood vessel and the background will be used to segment the two parts for further analyzations. This extracted feature can apply in segmenting the blood vessel in the microcirculation video images.

## Materials and Methods

### Image Acquisition

In order to obtain a microcirculation images to be used in this study, an imaging system which includes a CCD video camera, a LCD monitor, a video recorder, and a reflective microscope without fluorescence labeling were used. The video recorder used to record real-time blood flow images of microcirculation at ×380 magnification, with special resolution of 1.42 μm, and the matrix size of static image pixel is 720 × 480 (∼1 × 0.7 $mm^2$) that can be seen in Fig. 1a. The acquired original image shown in Fig. 1b. The sampling rate of the video is 30 frames per second.

### Vessel Alignment

There are many types of motion blur which are caused by relative motion between the focused object and the camera. Under translation, an object at its original position has shifted to different position creating a blurring distance. Real time images of the blood vessels have always having motion blur generated by regular heartbeats in mice. The image motion problem of blood vessel images has to be corrected by matching images in frame-by-frame. The image alignment was achieved by a block matching method, that moved frame-by-frame images to the center of the vessel in each image and match it with its succeeding frames in order for the vessel to register and align (Fig. 2).

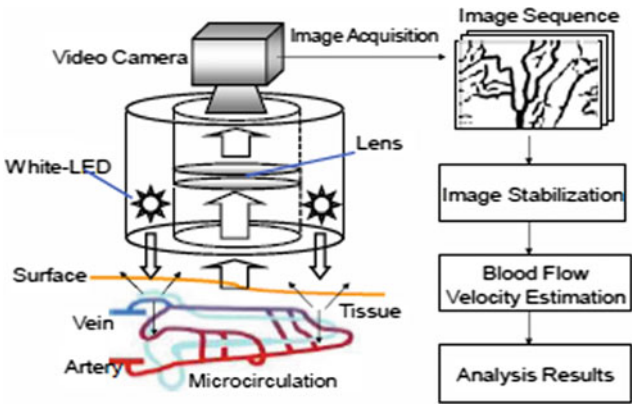



**Fig. 2** Schematic representation of the microscopy system and data analysis flow chart

### Feature Extraction for Blood Vessel from Mouse

The microcirculation blood vessel segmentation was based on the serious components of feature location and integration in microcirculation video images. There were various steps in recognizing of the blood vessel. Histogram equalization was applied as the first step in enhancing the microcirculation video images of 30 frames, i.e. one-second video images. This will allow the 27-frame images processed under the same situation, and obtained the consistent image quality. This enhanced image had smoothed by a 5 × 5 mean filter to reduce the noise caused by the reflective point as seen in the image as seen in Fig. 3.

The second step was to collect the image intensity at the same pixel location from different image frames. The patterns of different image pixel represented by a data train of intensity which was assigned to each image pixel. The total 256 × 256 pixel's patterns will be separated into two classes, that were the blood vessel and background. And then, all corresponding features were defined as below:

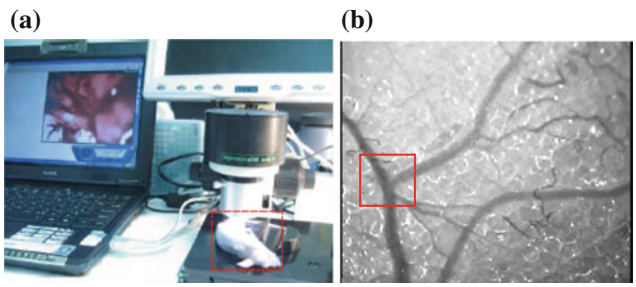



**Fig. 1** **a** experimental setup. **b** Image pixel matrix (1 $mm^2$) with 720 × 480 pixels. The rectangular region shown in RED marker is the ROI (256 × 256 pixels) for image analysis

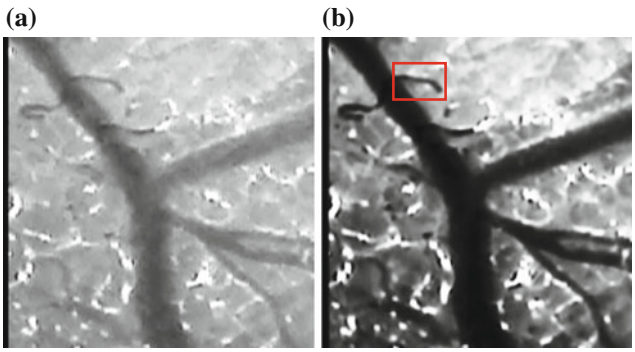



**Fig. 3** Blood vessel image enhancement. **a** original image of microcirculation, and **b** enhanced image at the marked rectangular region focused results

- Mean ($\bar{x}$): The mean of the pixel value of the 27 sample points from all the trained images
- Variance ($\sigma^2$): the variance of pixel value
- Maximum normalized-value (*max*)
- Minimum normalized-value (*min*)

## Linear Discriminant Analysis

Linear discriminant analysis has commonly used as dimensionality reduction technique in the pre-processing step for pattern classification and machine learning applications. This technique have used to maximize a function that represents the differentiations between the class centers of two classes from normalized feature patterns of scatter.

We have two classes that will define to the scatter, which will be an equivalent of the variance as:

$$\tilde{s}_i^2 = \sum_{y \in \omega_i} (y - \tilde{\mu}^2), \tag{1}$$

where $\tilde{s}_i^2$ measures the variability within class $\omega_i$ after projecting it on the feature-space. This formula had further used in Fisher linear discriminant method. Fisher linear discriminant had defined as linear function $w^T x$ that maximizes the criterion in terms of $S_w$ and $S_B$ as:

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{w^T S_B w}{w^T S_W w}, \tag{2}$$

where J(w) is the measure of the difference between class which are the background and the blood vessel (encoded in between-class scatter matrix, $S_B$) normalized by a measure of the within-class scatter matrix, $S_w$. Solving Eq. (2) in term of generalized eigenvalue, by finding the transformation matrix (w). This matrix has used to transform all feature vectors to a new-set of feature vectors that are differentiating the blood vessel and the background.

## Results

There are total of 51 microcirculation images that used in detecting the blood vessel for segmentation. In Fig. 4a and b, the feature plots of the pixel value mean and pixel value variance were depicted. It was noticeable that there was an intersection between background and blood vessel pixel value mean, around 40–110, and also intersection in the variance around 10–50.

These features didn't separate the blood vessel and background very well as it is shown in the Fig. 5. By applying Fisher linear discriminant analysis, all the feature

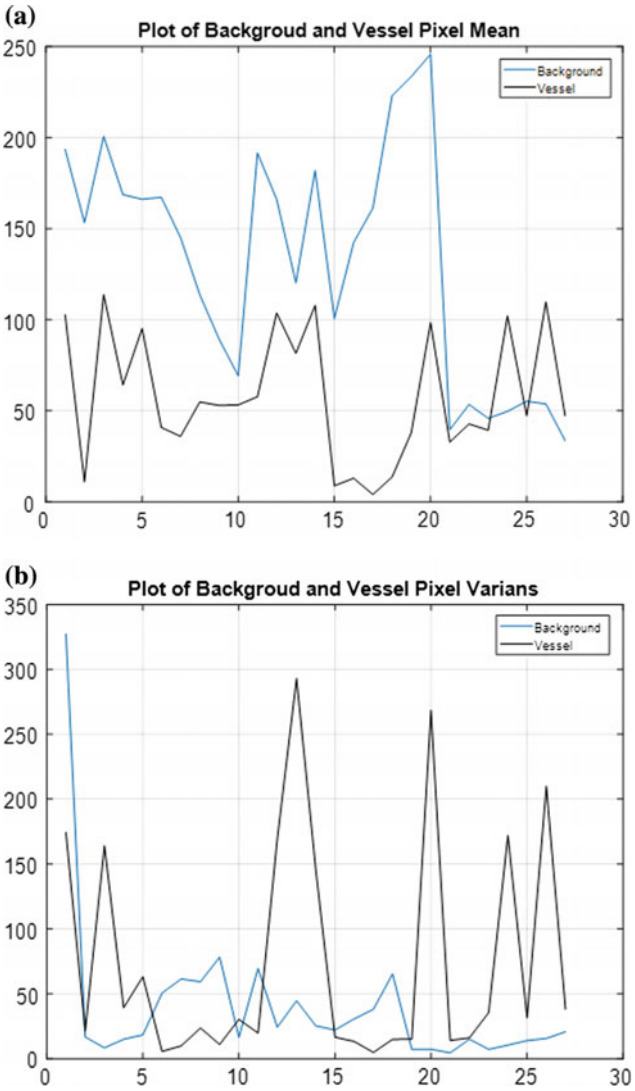



**Fig. 4** Feature plots of two classes—blood vessel (black line) and background (blue line). **a** Mean value of their pixel value and **b** variance used in Fisher discriminant analysis

vectors transformed with transformation matrix *w*, (which is actually an eigenvector from the highest eigenvalue), to be a single value that would be discriminative. In Fig. 6 the transformed feature value was depicted. It was very clear that most of the feature separated though in a few values that the features of the background and the blood vessel have been well discriminative in intersection.

After calculating all of the features of the pixel in the original video images, and transforming them by using the transformation matrix, the feature classifier was applied. In the classification of image segmentation, each pixel assigned to the background or the blood vessel was performing by measure its closeness to the reference features in the sense of Euclidean distance.

**Fig. 5** The image as a result of process without applying LDA

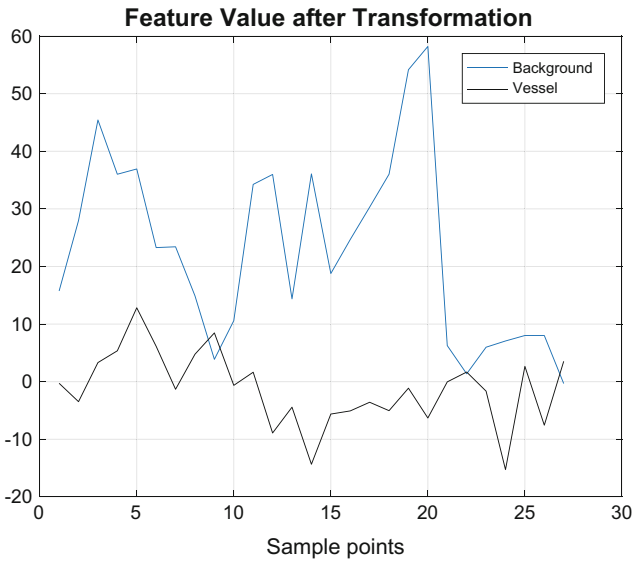



**Fig. 6** Feature plots of two classes' feature values for the background and the blood vessel

As the result that is shown in Fig. 7, some improvements occur compare to the Fig. 5, which more blood vessel can be recognized, particularly the small blood vessel as indicated in the marked red-box. Though some parts of the image, especially the left-bottom still need another approach to recover the blood vessel that has similar pixel value with the background.

## Conclusion

Finding the proper feature for the image are still a challenging issue. In this paper, the using of linear discriminant analysis is still not completely in recognizing all of the background yet, but it is still a promising method that can

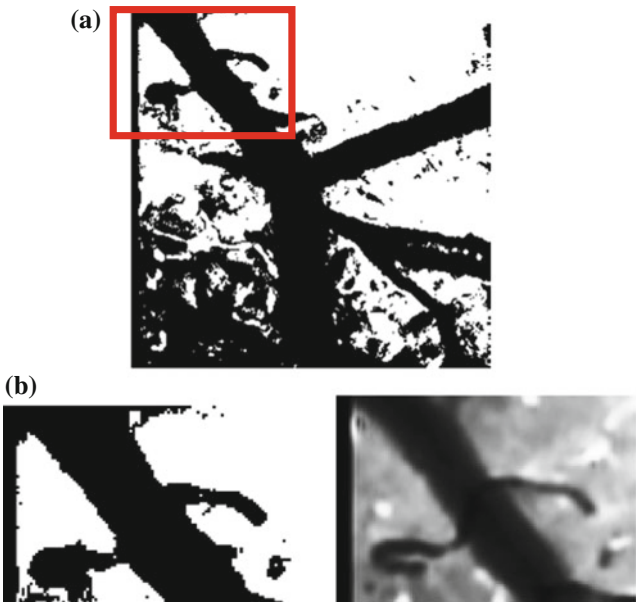



**Fig. 7** **a** Segmented image of the background and the blood vessel. **b** Left: segmented image of the small blood vessel from the marked rectangular region on the enhanced image. Right: zoom-in image for comparison

straightforwardly detect and recognize the blood vessel, even for the small blood vessels that can be acquired by the microscopic imaging system.

It is special and helpful in segmentation of the blood vessel to differentiate the background from a microcirculation video images that obtained only through the white LED lighting as the illumination source and no fluorescence labeling. Using the collected frame-to-frame image intensity patterns as a feature to segment the blood vessel has been successfully developed as an applicable method for the blood flow velocity measurement of the further analysis for each microcirculation imaging information [9].

**Acknowledgements** This work have supported by grants MOST 106-2221-E-033-001 from the Ministry of Science and Technology, R.O.C. (Taiwan).

**Conflict of Interest** All contributing authors declare no conflict of interest.

## References

1. Kirbas Cemil, Quek F (2004) A review of vessel extraction techniques and algorithms. ACM Comput Surv (CSUR) 36(2):81–121
2. Menche N (2012) Blood circulatory system. In: Biologie anatomie physiologie. Munich
3. Thomas B, Sumam KS (2016) Blood flow in human arterial system—a review. ICETEST Proc 4(2016):339–346

4. Wiedman M (1963) Dimensions of blood vessels from distributing artery to collecting vein. J Am Heart Assoc 12:375–378. https://doi.org/10.1161/01.RES.12.4.375
5. Lin WC et al (2011) An approach to automatic blood vessel image registration of microcirculation for blood flow analysis on nude mice. Comput Methods Biomech Biomed Eng 14(4):319–330. https://doi.org/10.1080/10255842.2010.497489
6. Kashef N, Fouad Y, Mahar K (2017) An automatic changeable edge detection model for digital images. J Comput Sci Syst Biol 10:056–060. https://doi.org/10.4172/jcsb.1000249
7. Popel S, Johnson P (2005) Microcirculation and hemorheology. Annu Rev Fluid Mech. 37:43–69. https://doi.org/10.1146/annurev.fluid.37.042604.133933
8. Lin W-C, Lin K-P (2011) Red blood cell velocity measurement in rodent tumor model: an *in vivo* microscopic study. J Med Biol Eng 32(2):97–102. https://doi.org/10.5405/jmbe.875
9. Lin W-C, Lin K-P (2016) Feature analysis and extraction of tumor vascular structure on microscopic images. Biomed Health Inform. https://doi.org/10.1109/BHI.2016.7455929

# Active Learning for Semi-automated Sleep Scoring

N. Grimova, M. Macas, and V. Gerla

**Abstract**

This paper introduces the semi-automatic process using active learning methods which could improve the current state, where a human specialist has to annotate a multiple hours long polysomnographical record to sleep stages. This work is focused on the utilization of density-weighted methods of active learning, one of them turned out to be well-suited for this type of task. Moreover, we proposed several criteria for the comparison of active learning methods. The method saves more than 80% of expert's annotation effort.

**Keywords**

EEG • Active learning • Sleep stages • Classification • Evaluation criteria

## Introduction

An expert annotation of long biomedical signals can be very time and money consuming, moreover its quality is able to be reduced by the expert's fatigue or by the replacement of the expert by a less skilled annotator. The whole annotation process can be improved by semi-automatic methods using the active learning paradigm. Active learning can be used in terms of a selection of several signal segments in order to be manually annotated by the expert followed by the subsequent automatic annotation of the rest of the signal. Another approach is a detection of erroneously annotated parts of the signal and a query for their re-annotation. A typical example of a suitable application is the annotation of polysomnographical signals. This work is focused on active learning for the classification of polysomnographical data to five sleep stages. However, proposed methods can be applied on any suitable application domain.

Polysomnography (PSG) deals with recording of several biological signals during sleep and it is often used for the diagnosis of sleep disorders such as sleep apnea or narcolepsy. According to The AASM Manual for the Scoring of Sleep and Associated Events [1], it is recommended to monitor brain (EEG), heart (ECG) and muscle activities (EMG), eye movements (EOG), the respiration, and the pulse oximetry, in order to detect abnormalities and also to distinguish between individual sleep stages.

Sleep is divided into two main stages which cyclically repeat 4-6 times during the sleep time—non-rapid eye movements (NREM) and rapid eye movements (REM) [2]. NREM sleep is characterized by a low overall activity and can be further subdivided into three stages which differ particularly according to the observed brain activity [2]. As the name suggests, REM sleep is characterized mostly by rapid eye movements and also by irregularities in the respiration, the twitching of muscles and changes in the heart rate [2].

N. Grimova
Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic

M. Macas (✉) · V. Gerla
Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Zikova street 1903/4, Prague, Czech Republic
e-mail: martin.macas@cvut.cz



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_24

**Table 1** Number of instances of each class for each dataset

| # | WAKE | REM | NREM1 | NREM2 | NREM3 | Sum |
|---|---|---|---|---|---|---|
| 1 | 284 | 53 | 87 | 309 | 146 | 879 |
| 2 | 200 | 88 | 64 | 443 | 53 | 848 |
| 3 | 292 | 124 | 61 | 316 | 118 | 911 |
| 4 | 362 | 15 | 103 | 299 | 17 | 796 |
| 5 | 87 | 107 | 57 | 482 | 121 | 854 |

## PSG Data

We have at our disposal five overnight polysomnographs of healthy subjects that consist of channels recommended by AASM: six EEG channels (F4-M1, C4-M1, F3-M2, C3-M2, O1-M2), two EOG channels and two EMG channels. Each record was split to 30 s-long segments and these segments were assigned to a sleep stage by the expert. The number o f segments of all classes is shown in Table 1. Further, a feature extraction was performed. Each signal segment was represented by 82 features which were proposed and verified in [3]. Eleven features were computed on each of EEG channels, five features on EOG channels, and three features were acquired from each of EMG channels.

Finally, all datasets were divided to training and testing data in such way that instances from each class are evenly distributed in both parts. The Parzen classifier was chosen as the classifier.

## Active Learning

In supervised learning, the goal of the task is to train a classifier on a given set of labeled observations, so when an instance of unknown class is observed, the classifier can correctly assign it to its class. In many real-world scenarios including sleep staging, it is relatively easy to collect instances, but it is expensive or difficult to obtain their classes (labels). It typically includes applications where a large amount of data is available as are e.g. the text classification [4], natural language processing [5], or when an expertise is needed.

Instead of standard supervised methods, active learning enables to select only the most informative instances for learning which are further annotated by the oracle (the expert in our case) and used for training of the classifier. We demonstrate here that this kind of semisupervised approach needs less effort from the expert. In the real annotation procedure, the computer agent automatically suggests segments for annotation. After the annotation a sufficient number of these segments, a classifier is trained on labeled data and classifies all remaining unlabeled segments.

The essential part of active learning is to determine in which way the most informative instances will be chosen. Two strategies are addressed in this work: uncertainty sampling and density-weighted methods.

## Query Selection Approaches

In the uncertainty sampling strategy, the instance whose class the classifier is the least certain of, is chosen [6]. Uncertainty sampling methods are further divided into categories according to how the uncertainty is computed, these strategies are more described in [6]. In our work, we used the so-called margin uncertainty sampling strategy (MUS). The instance is queried for which holds:

$$x^* = \underset{x}{\operatorname{argmax}}\ I(x), \tag{1}$$

where $I(x) = 1 - P(\omega_1|x) + P(\omega_2|x)$ is the certainty computed as the complement to the difference of aposterior probabilities of the first and the second most probable class label $\omega_1$ and $\omega_2$.

However, the classifier is typically the most uncertain about instances which are close to the classification border, but these instances do not fully represent the class distribution [6]. Moreover, in the sleep staging problem, these instances are often erroneously classified by the expert. For this reason, Settles et al. [7] proposed the density-weighted approach. When a density-weighted method is used, the instances is queried for which holds:

$$x^* = \underset{x}{\operatorname{argmax}}\ I(x) \times \left( \frac{1}{|S_U|} \sum_{j=1}^{|S_U|} s(x, x_j) \right)^n \tag{2}$$

where $I(x)$ is the certainty of instance $x$, $S_U$ is the set of unlabeled instances, $s(x_i, x_j)$ is the distance between instances $x_i$ and $x_j$ and $n$ is the weight of the distance term.

As it can be seen in Table 1, the number of instances in each class differs significantly, therefore these data are imbalanced. Since density-weighted methods are suitable for learning on imbalanced data [8, 9], we decided to use three methods for our task: the simplest density-weighted method and two other approaches, which we propose.

**DWD*n*** In DWD $n$ (Density Weighted method with Distance as the weight term), the informativeness of each instance is multiplied by the mean distance to all unlabeled instances. The number $n$ denotes the weight of the similarity term in Eq. (1).

**UMDN*l*** In UMND*l* (Uncertainty weighted by the Maximal Distance to the nearest Neighbor) strategy, the instance from $l$ least certain instances, which has the maximal distance from its nearest neighbor, is queried.

**UMDL** In UMDL (Uncertainty weighted by the Minimal Distance to Labeled instances) strategy, the centroid of labeled instances of each class is computed, the unlabeled instance is selected for which the dot product of its certainty and the distance to the nearest centroid is minimal

## Performance Evaluation

Active learning methods are commonly compared to random sampling (RS) when the instance is chosen from the set of unlabeled instances at random, the rest of the procedure remains. The comparison of methods is mostly done visually, i.e. the chosen metric (e.g. the accuracy, the error, the miss rate) is plotted as a function of the number of labeled instances.

Additionally, we propose four evaluation criteria in order to provide a more automatic and objective comparison of active learning methods and random sampling.

**Reaching Threshold Criterion (RT)** deals with specifying of the threshold of a computed metric and determining how many instances are necessary to learn on until this threshold is reached. This criterion is able to be used for the interpretation how many instances will be saved when active learning will be used.

**Mean Rank (MeanR)** is computed as follows: performances of active learning methods are compared and ordered in each step of the algorithm, the mean rank over all iterations is computed for each method.

**Most Common Rank (ModeR)** is similar to the previous criterion, however the most frequent rank is computed over all iterations of active learning.

**Breaking Point (BP)** is the number of iteration after which a certain percentage of classification performance values would be smaller or bigger (according to the chosen metric) than a given threshold

**Table 2** Evaluation criteria for Sleep2 dataset, threshold is set to 0.2

| Methods | RT | MeanR | ModeR | BP |
|---|---|---|---|---|
| RS | 329 | 5.396 | 4.2 | 237.6 |
| MUS | **130** | 3.296 | 6 | 165.8 |
| DWD1 | 270.8 | 4.764 | **2.6** | 301.4 |
| DWD0.5 | 381.2 | 4.26 | **2.6** | 390.6 |
| UMDN10 | *131* | *2.678* | *3* | *139* |
| UMDN50 | 238 | **2.624** | 4.4 | **98.5** |
| UMDL | 293.6 | 4.984 | 3.2 | 402.2 |

## Experimental Results

The mean values of all criteria for individual methods on all datasets are shown in Table 2, the best results are highlighted in bold, the second best ones in italics. Note that when a method did not reach the specified threshold in Reaching Threshold or Breaking Point criteria then the value of this criterion was set as the amount of all instances in the dataset. We can see that UMDN*l* methods together with MUS strategy converged faster than other methods (RT and BP criteria) and both UMDN*l* methods reached the best result in MeanR criterion. The most frequent rank of both DWD*n* methods are the lowest.

We get similar outcomes if we will compare how many times each strategy was better than random sampling. The results are shown in Table 3. The clear winner is the UMDN10 strategy, which reached better results than random sampling in 18 out of 20 cases, followed by UMDN50 and MUS strategies with 16 cases.

The false negative rates for each class on Sleep1 and Sleep2 dataset for MUS and UMDN10 strategy are shown in Fig. 1a–d. It can be observed that the highest miss rate is obtained on NREM1 class, it reached the value 0.5 on Sleep1 dataset and the value 0.7 on Sleep2 dataset. This is able be caused by the transient character of NREM1, which means that some instances from this class are very similar to observations belonging to classes NREM2 or REM. This is interesting problem of noisy oracles [6] and we believe that

**Table 3** Summarization of how many times active learning method was better than random sampling for each criterion and each datasets

| Methods | TRtT | MeanR | ModeR | BP | Sum |
|---|---|---|---|---|---|
| MUS | 5 | 5 | 1 | 5 | 16 |
| DMD1 | 3 | 2 | 4 | 2 | 11 |
| DWD0.5 | 2 | 3 | 4 | 3 | 12 |
| UMDN10 | 5 | 5 | 3 | 5 | 18 |
| UMDN50 | 5 | 5 | 2 | 4 | 16 |
| UMDL | 3 | 3 | 3 | 3 | 12 |

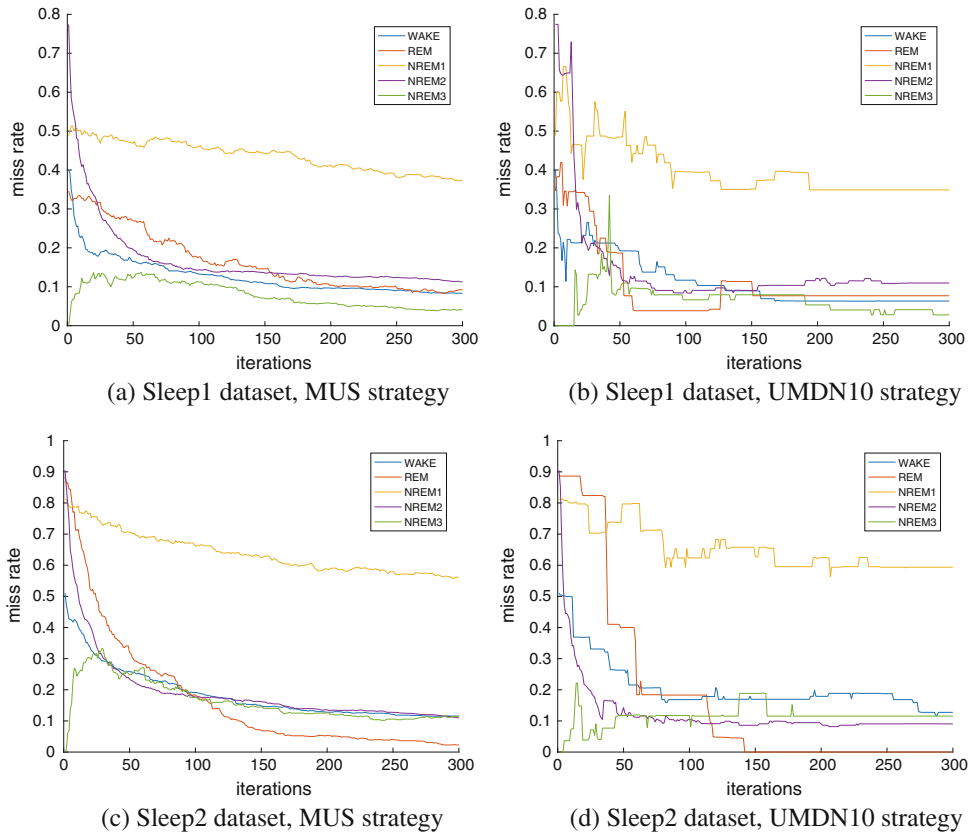



**Fig. 1** Miss rates of all classes on Sleep1 and Sleep2 datasets

(a) Sleep1 dataset, MUS strategy

(b) Sleep1 dataset, UMDN10 strategy

(c) Sleep2 dataset, MUS strategy

(d) Sleep2 dataset, UMDN10 strategy

active learning could be advantageously used for querying re-annotation of transient or mislabeled instances.

Finally, false negative rates averaged over all classes are depicted in Fig. 2 for RS, MUS and UMDN10 strategy. The number 4 denotes that instances of NREM1 were excluded from training data, the number 5 means that all instances were used for training. We find interesting that on Sleep1 dataset MUS strategy had a better performance than UMDN10 when instances from NREM1 class were excluded, on the other hand UMDN10 strategy reached better results than MUS when the classification to all classes was used.

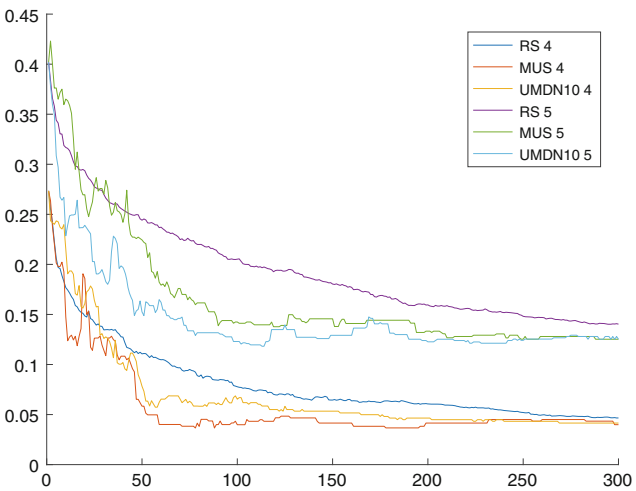



**Fig. 2** Miss rates of all classes on Sleep1 dataset when observations are classified into four (without NREM1 class) or into all five classes

## Discussion

In this work, several active learning strategies were proposed and used for the classification of polysomnographical data, where each 30 s-long segment was represented by 82 features and classified into five sleep stages. Methods were compared to random sampling visually and also by four evaluation criteria. The newly proposed UMDN10 method outperformed random sampling in 18 out of 20 experiments and also provided the second fastest convergence. Active learning saves more than 80% of the annotation effort in average.

We also detected the problem of the classification to stage NREM1. It might be caused by the continuity of original data and the ambiguity of samples at the REM-NREM1 and NREM1-NREM2 transitions. We suppose that when the time-dependency (temporal context) between samples would be considered, the overall performance will reach better results. However in that case, the active learning will be

much more challenging, which will be solved in our future work.

**Acknowledgements** Research has been supported by Czech Technical University SGS grant No. OHK-4-/3T/37 and Czech Science Foundation grant no. 17-20480S "Temporal context in analysis of long-term non-stationary multidimensional signal".

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Brett D, Conchita R, Jasmina M, Craig H (2014) The AASM recommended and acceptable EEG montages are comparable for the staging of sleep and scoring of EEG arousals. J Clin Sleep Med JCSM: Off Publ Am Acad Sleep Med 10:803
2. Rama AN, Cho SC, Kushida CA (2006) Normal human sleep. In: Sleep: a comprehensive handbook
3. Gerla V (2012) Automatic analysis of long-term EEG signals. PhD thesis, Czech Technical University
4. McCallum A, Nigam K (1998) Employing EM and pool-based active learning for text classification. In: Proceedings of the fifteenth international conference on machine learning ICML '98, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc., pp 350–358
5. Tomanek K (2010) Resource-aware annotation through active learning. PhD thesis, Technical University Dortmund
6. Settles B (2009) Active learning literature survey computer sciences. Technical report 1648, University of Wisconsin–Madison
7. Settles B, Craven M (2008) An analysis of active learning strategies for sequence labeling tasks. In: Proceedings of the conference on empirical methods in natural language processing, EMNLP '08, Stroudsburg, PA, USA. Association for Computational Linguistics, pp 1070–1079
8. He H, Ma Y (2013) Imbalanced learning: foundations, algorithms, and applications, 1st edn. Wiley-IEEE Press
9. Zhu J, Wang H, Yao T, Tsou BK (2008) Active learning with sampling by uncertainty and density for word sense disambiguation and text classification. In: Proceedings of the 22nd international conference on computational linguistics, COLING '08, vol 1, Stroudsburg, PA, USA. Association for Computational Linguistics, pp 1137–1144

# Human Fall Detection from Acceleration Measurements Using a Recurrent Neural Network

T. Theodoridis, V. Solachidis, N. Vretos, and P. Daras

**Abstract**

In this work, a method for human fall detection is presented based on Recurrent Neural Networks. The ability of these networks to process and encode sequential data, such as acceleration measurements from body-worn sensors, makes them ideal candidates for this task. Furthermore, since such networks can benefit greatly from additional data during training, the use of a data augmentation procedure involving random 3D rotations has been investigated. When evaluated on the publicly available URFD dataset, the proposed method achieved better results compared to other methods.

**Keywords**

Human fall detection • Recurrent neural network • Data augmentation • Acceleration

## Introduction and Related Work

A fall is defined as an event that results in a person coming to rest inadvertently on the ground or floor or other lower level [1]. The people more susceptible to falls are usually the elders. The frequency of fall events is even higher in elders who suffer from chronic illnesses [2] (e.g. Parkinson, Arthritis, Osteoporosis). Moreover, in many cases, a fall may immobilize a person and make him/her unable to call for assistance. Thus, the presence of carers at home becomes necessary, resulting in increased expenses for the patient and the health-care system.

In this paper, a method that automatically detects a fall is presented. Although a fall detection system does not prevent the fall, the information that provides is valuable and can be used by both carers and medical professionals. The necessity for continuous presence by caregivers of chronic illness patients due to the risk of fall, can be relaxed if a fall detection system is installed, since, in the case of a fall, it can alert them in order to assist the patient. This technology can ease the life of caregivers (professional or patient relatives) and, at the same time, contribute to the decrease of the health-care system expenses. Regarding medical professionals, the detailed reports that the fall system can provide, give valuable information (e.g., frequency of falls per time of the day, increase/decrease of incidents, etc.), since they can be correlated with medication changes and, hence, contribute to the medication scheme definition.

Several methods that detect falls have been presented in the literature using a variety of sensors. The most common sensors that are used are accelerometers [3–8], RGB cameras [9, 10], depth or infrared cameras [11]. Other technologies such as floor-vibration sensors [12, 13] and Wireless Sensor Networks [14] have been employed as well.

The fall detection method presented in [15] uses acceleration measurements from two devices placed at the trunk and thigh of the users. A threshold value on the acceleration magnitude is used in order to determine if a fall has occurred or not. Two variations of this approach are presented: one that signals a fall event if the acceleration magnitude exceeds a certain threshold and another that signals a fall if the acceleration magnitude goes below a different threshold.

T. Theodoridis (✉) · V. Solachidis · N. Vretos · P. Daras
Information Technologies Institute, Centre for Research and Technology Hellas, 6th km Charilaou - Thermi Road, Thessaloniki, Greece
e-mail: tomastheod@iti.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_25

In [16] a fall detection method is proposed that uses data acquired from an accelerometer, placed near the pelvis region of the users, and depth cameras. The system assumes that there is no fall if the acceleration magnitude is below a certain threshold, regardless of the depth camera input. If the acceleration magnitude exceeds the threshold value, then the input from the depth camera is analyzed. The method detects the person from the depth image along with the floor plane equation, and then extracts features related to the person's body position (e.g., distance of body's centroid from the floor, ratio of the person's bounding box dimensions, etc.). A Support Vector Machine (SVM) classifier, based on these features, produces the final decision.

The authors in [17] propose three different methods that rely on a Kinect device and two wearables placed to the person's wrist and waist. The methods use the body skeleton captured by Kinect and the wearable devices' acceleration and orientation. The best method of the three relies on the rapid downward movement of the spine base joint of the human body, on the distance of said joint from the floor and on the acceleration magnitude.

In [18] a variety of machine learning models has been tested on features extracted from the acceleration measurements of a wearable device and a mobile phone. The features extracted in the time domain include the mean, variance, kurtosis, etc., while the frequency domain features were the autocorrelation coefficients and the total spectral power in different frequency bands. The best results were obtained using a Decision Tree ensemble for both the wearable sensor and the mobile phone.

Despite the fact that, generally, the use of RGB or depth cameras increases the accuracy of fall detection methods compared to ones that use only accelerometers, cameras have two significant disadvantages: cost and limited coverage area. Thus, if we wish to apply a method that includes them in a house set-up, we have to install cameras in every room, increasing the total system cost. The proposed method detects falls using only data acquired from a body-worn accelerometer, keeping the total system cost low, and at the same time, having a large spatial range where the method can be applied. Additionally, the proposed method is capable of identifying falls without false positive detections, as indicated by the evaluation results on the UR Fall Detection (URFD) dataset [16]. In section "Proposed Method" the proposed method is described in detail. The experimental evaluation is illustrated in section "Experimental Evaluation", where the proposed method is compared not only to methods that use acceleration as the only modality, but also to ones that employ accelerometer and depth cameras. Finally, in section "Conclusions" are drawn.

## Proposed Method

The proposed fall detection method takes advantage of the Recurrent Neural Networks' (RNNs) ability to process and encode the inherent information contained in sequential data. Traditional machine learning models, such as the Multi-layer Perceptron (MLP) or the Support Vector Machine (SVM), process their input without any notion of sequential order, and thus cannot take advantage of this information. Recurrent Neural Networks process their input in a sequential manner, accumulating more information after each time step about the sequence being presented to them. The Long Short-Term Memory variant of RNNs (LSTM [19, 20]), which is adopted in this work, further improves the basic RNN architecture, by enabling the network to retain information from many time steps back into the past, thus giving the network the ability to encode and learn longer sequences.

In order for a Recurrent Neural Network to process the input signal as a sequence, the sequence length $n$ must be determined from the beginning. Then, the signal is divided into time windows of length $n$ and the network processes each window independently. Figure 1 shows the network architecture used in this work. Variables $\mathbf{X_1}, \ldots, \mathbf{X_n}$ denote the multi-dimensional input signal that spans $n$ time steps. The first LSTM layer (first blue row) processes the input signal and produces an output at each time step $t = 1, \ldots, n$. The second LSTM layer, in the same way, processes the output of the first LSTM layer at each time step, but produces an output only at the last time step $n$. Then, a traditional feed-forward neural network (first cyan rectangle) processes the output of the second LSTM layer and, finally, a second feed-forward neural network produces the final decision $Y$, which in our case is the probability that a fall incidence has occurred. As with all supervised machine learning techniques, in order for the network to learn the output probabilities, besides the input signal at each time step, a label must also be provided, which indicates whether an incidence has occurred or not.

Due to the beneficial impact of additional training data on model generalization and performance, the effectiveness of a data augmentation procedure involving random rotations has been evaluated as well. Given the acceleration vector $\mathbf{a}(t) = [a_x(t), a_y(t), a_z(t)]$ at time $t$, that contains the acceleration along the $x$, $y$ and $z$ axes of the device respectively, a new vector $\mathbf{a}^r(t)$ can be obtained by rotating $\mathbf{a}(t)$ by $\theta$ radians about the $x$ axis, $\phi$ radians about the $y$ axis and $\psi$ radians about the $z$ axis:

$$\mathbf{a}^r(t) = \mathbf{R}_z(\psi) \cdot \mathbf{R}_y(\phi) \cdot \mathbf{R}_x(\theta) \cdot \mathbf{a}(t) \quad (1)$$

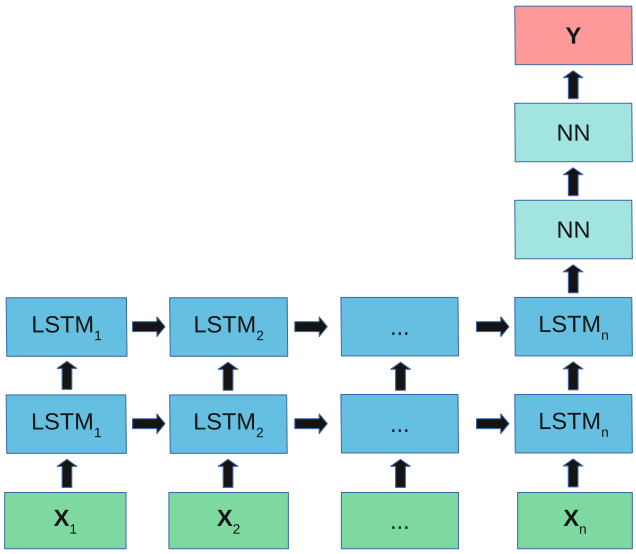



**Fig. 1** Proposed model architecture

where $(\cdot)$ denotes matrix multiplication, $\mathbf{R}_x(\theta)$, $\mathbf{R}_y(\phi)$ and $\mathbf{R}_z(\psi)$ are the rotation matrices about $x$, $y$ and $z$ axes respectively:

$$\mathbf{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & -sin(\theta) \\ 0 & sin(\theta) & cos(\theta) \end{bmatrix}, \theta \in [0, 2\pi) \quad (2)$$

$$\mathbf{R}_y(\phi) = \begin{bmatrix} cos(\phi) & 0 & sin(\phi) \\ 0 & 1 & 0 \\ -sin(\phi) & 0 & cos(\phi) \end{bmatrix}, \phi \in [0, 2\pi) \quad (3)$$

$$\mathbf{R}_z(\psi) = \begin{bmatrix} cos(\psi) & -sin(\psi) & 0 \\ sin(\psi) & cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \psi \in [0, 2\pi) \quad (4)$$

and $\mathbf{a}(t)$ is considered a $3 \times 1$ matrix in (1), in order for the multiplication to be well-defined.

## Experimental Evaluation

### The Dataset

The proposed method has been evaluated on the publicly available UR Fall Detection (URFD) dataset [16, 21], which contains measurements from an accelerometer, placed near the pelvis area of the human body, as well as depth images, and features extracted from those images, acquired from two Kinect cameras. In total, the dataset contains 40 sequences with activities of daily living and 30 sequences with falls.

The accelerometer data from the body-worn device at time $t$ consist of the 3D acceleration $\mathbf{a}(t) = [a_x(t), a_y(t), a_z(t)]$ as well as the norm of the acceleration vector (also called the magnitude or in some cases the total sum vector):

$$\|\mathbf{a}(t)\| = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)} \quad (5)$$

Regarding the Kinect camera measurements, the authors of the dataset have provided features extracted from the Kinect depth images, such as the width to height ratio of the person's bounding box in the depth image, the height of the person's centroid, etc. The measurements from the accelerometer and the Kinect camera have been temporally synchronized, so that at each time step information from all sensors is available.

### Parameter Selection and Evaluation Protocol

The proposed method, denoted as *LSTM-Acc*, consists of the network architecture presented in Sect. Proposed Method that processes sequences of length $n = 30$, which corresponds to a time span of one second. The two LSTM layers and the first feed-forward layer of the network have 200 units each, while the last feed-forward layer has two. The proposed method was trained and evaluated using only acceleration data. Furthermore, we augmented the training data with one rotated version of the original measurements by a random angle of $\theta, \phi, \psi \in [-10, 10]$ degrees about $x$, $y$ and $z$ axes, using the procedure discussed at the end of Sect. Proposed Method. In doing so, we increase the number of available samples for training and also force the model to learn representations that are more robust to rotations. This second approach is denoted as *LSTM-Acc Rot*.

Two methods were chosen and implemented for comparison purposes. The first method was proposed in [15] and is denoted as *UFT*. It uses a threshold on the acceleration magnitude $\|\mathbf{a}(t)\|$ in order to determine if a fall has occurred or not. In training, the threshold is determined as the minimum of the magnitude peaks during the fall instances. In testing, the same threshold is used for separating fall instances from non-fall ones.

The second method was proposed in [16] and is denoted as *Acc + SVM-Depth*. It uses a threshold value of $3g$ on the acceleration magnitude $\|\mathbf{a}(t)\|$ in order to initiate a fall detection procedure, which consists of an SVM model that has been trained on the extracted depth features giving the final decision. The depth features were scaled so that they have zero mean and unit variance.

Regarding the *UFT* and *Acc + SVM-Depth* methods, there was no point in using the augmented dataset discussed previously, since they rely on the acceleration magnitude $\|\mathbf{a}(t)\|$, which is invariant to rotations.

The evaluation was performed using a 10-fold cross-validation procedure, in which the dataset is split

**Table 1** The fall detection results (%) on the URFD dataset

| | *LSTM-Acc* | *LSTM-Acc Rot* | *Acc + SVM-depth* | *UFT* |
|---|---|---|---|---|
| Accuracy | 95.71 | **98.57** | 92.86 | 92.86 |
| Precision | 95.00 | **100** | 94.17 | 90.00 |
| Sensitivity | **96.67** | **96.67** | 90.00 | **96.67** |
| Specificity | 95.00 | **100** | 95.00 | 90.00 |

into ten parts, nine of which are used for training and one for testing. The procedure is repeated ten times, so that all possible test parts have appeared once. Since this dataset consists of 40 sequences with activities of daily living and 30 with falls, each fold contained 4 sequences from the first group and 3 from the second. Finally, the results were evaluated on a sequence level using four metrics: accuracy, precision, sensitivity and specificity. The evaluation on sequence level means that for a non-fall sequence to be correctly classified, the models had to produce zero alerts during the whole sequence. On the other hand, in order for a fall sequence to be classified correctly, the models had to produce at least one fall alert, starting from one second before the beginning of fall and onward, not before.

## Experimental Results

The experimental results on the URFD dataset are shown in Table 1. Starting from the simplest method of the four, *UFT*, it is evident that it produces the highest amount of false positives (detecting a fall when no fall has occurred), as it has the lowest specificity score (90%). Overall, it has the same accuracy as the *Acc + SVM-Depth* method, which however has higher specificity, but lower sensitivity. This means that the *Acc + SVM-Depth* method produces fewer false positives, but also finds fewer actual falls. Next, the proposed *LSTM-Acc* method is equal to or better than the previous methods in all evaluation metrics. It has the same sensitivity as the *UFT* method (96.67%), the same specificity as the *Acc + SVM-Depth* method (95%) and higher precision and accuracy than both. Lastly, the *LSTM-Acc Rot* approach has produced the best results. Even though it relies only on acceleration information, it has not produced a single false positive result (specificity = 100%), while the sensitivity of 96.67% corresponds to not detecting one fall event.

## Conclusions

In this work, a Recurrent Neural Network-based approach to fall detection has been presented. By leveraging the ability of such networks to process sequential data, as well as data augmentation in the form of random rotations of the input acceleration signal, the proposed method was able to find all but one fall event, while at the same time producing no false alarms when tested on the URFD dataset.

**Acknowledgements** This work was supported by the European Project: ICT4LIFE http://ict4life.eu/ Grant no. 690090 within the H2020 Research and Innovation Programme.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. World Health Organization WHO, Falls. http://www.who.int/mediacentre/factsheets/fs344/en/. Accessed 19 Sep 2017
2. Kalache A, Fu D, Yoshida S, et al (2007) World health organisation global report on falls prevention in older age. World Health Organisation
3. Yuwono M, Moulton BD, Su SW et al (2012) Unsupervised machine-learning method for improving the performance of ambulatory fall-detection systems. Biomed Eng 11:9
4. Bourke AK, Ven P, Gamble M, et al (2010) Assessment of waist-worn tri-axial accelerometer based fall-detection algorithms using continuous unsupervised activities. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, pp 2782–278
5. Liu S-H, Cheng W-C (2012) Fall detection with the support vector machine during scripted and continuous unscripted activities. Sensors 12:12301–12316
6. Kangas M, Vikman I, Wiklander J et al (2009) Sensitivity and specificity of fall detection in people aged 40 years and over. Gait Posture 29:571–574
7. Koshmak GA, Linden M, Loutfi A (2013) Evaluation of the android-based fall detection system with physiological data monitoring. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, pp 1164–1168
8. Abbate S, Marco A, Bonatesta F et al (2012) A smartphone-based fall detection system. Perv Mobile Comput 8:883–899
9. Vishwakarma V, Mandal C, Sural S (2007) Automatic detection of human fall in video. Pattern Recognit Mach Intel 616–623
10. Rougier C, Meunier J, St-Arnaud A et al (2011) Robust video surveillance for fall detection based on human shape deformation. IEEE Trans Circuits Syst Video Technol 21:611–622
11. Mastorakis G, Makris D (2014) Fall detection system using Kinects infrared sensor. J Real-Time Image Process 9:635–646
12. Rimminen H, Lindström J, Linnavuo M et al (2010) Detection of falls among the elderly by a floor sensor using the electric near field. IEEE Trans Inf Technol Biomed 14:1475–1476
13. Alwan M, Rajendran P J, Kell S, et al (2006) A smart and passive floor-vibration based fall detector for elderly. In: Proceedings of

information and communication technologies, vol 1, pp 1003–1007
14. Wang Y, Wu K, Ni LM (2017) Wifall: device-free fall detection by wireless networks. IEEE Trans Mob Comput 16:581–594
15. Bourke AK, Obrien JV, Lyons GM (2007) Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. Gait Posture 26:194–199
16. Kwolek B, Kepski M (2014) Human fall detection on embedded platform using depth maps and wireless accelerometer. Comput Methods Programs Biomed 117:489–501
17. Cippitelli E, Gasparrini S, Gambi E, et al (2016) An integrated approach to fall detection and fall risk estimation based on RGB-depth and inertial sensors. In: Proceedings of the international conference on software development and technologies for enhancing accessibility and fighting info-exclusion, pp 246–253
18. Alzubi H, Ramzan N, Shahriar H, et al (2016) Optimization and evaluation of the human fall detection system. In: Proceedings of SPIE 10008, remote sensing technologies and applications in urban environments, p 1000816
19. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780
20. Gers FA, Schmidhuber J, Cummins F (1999) Learning to forget: continual prediction with LSTM. In: Proceedings of international conference on artificial neural networks, pp 850–855
21. UR Fall Detection Dataset. http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html. Accessed 19 Sep 2017

# Optimal Threshold Selection for Acceleration-Based Fall Detection

G. Šeketa, J. Vugrin, and I. Lacković

**Abstract**

In this paper we present the results of an experiment with 16 subjects performing activities of daily living and simulated falls. We used a triaxial accelerometer to track the subjects' movements. From the accelerometer data we calculated five different features that are used for fall detection. Contingency tables were built based on the collected dataset and ROC curves were plotted. Optimal thresholds for every feature and corresponding sensitivities and specificities were calculated based on the ROC curve analysis.

**Keywords**

Fall detection • Thresholds • Acceleration • ROC

## Introduction

Falls are a common problem among the elderly population that can result with serious consequences. Falls may cause various physical injuries as hip fracture, traumatic brain injuries, upper limb injuries, and in some cases death of the falling person [1]. Among physical effects, falls affect the quality of life: fallers have a fear of falling again, suffer from decreased mobility and loss of independence [2].

Various solutions have been proposed for fall detection. Research has mostly been focused on using vision based, ambient or inertial sensors to detect fall events. Wearable systems equipped with inertial sensors (accelerometers, gyroscopes) are suitable for fall detection due to their low cost and portability [3]. They are of small size and weight and can have a low power consumption making them convenient for a prolonged use.

Algorithms for fall detection found in the literature [4–12] can roughly be divided in two groups: threshold-based algorithms and algorithms based on machine learning. Some papers proposed combining those two approaches for better results [9]. These algorithms mostly use accelerometer measurements, but the use of other sensors, like gyroscopes [12] and altimeters [7] has also been explored. A thorough review of fall detection algorithms based on inertial sensing can be found in [13].

In this paper, we present the results of an experiment we have conducted with subjects simulating falls and performing activities of daily living in a laboratory setting. We measured their activities by an accelerometer sensor and used these measurements to extract features for fall detection. We present the algorithm for optimal threshold determination and the corresponding sensitivities and specificities for the chosen features that were calculated from our dataset.

## Methods

### Subjects and Equipment

For the experiment, 16 subjects were recruited to perform simulated falls and activities of daily living (ADL). Subjects age ranged from 15 to 44 (mean 23.1, std 6.7), height from 159 to 192 cm (mean 176.4, std 8.9) and body mass 55–93 kg (mean 75.2, std 12.8).

G. Šeketa · J. Vugrin · I. Lacković (✉)
Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, Zagreb, Croatia
e-mail: igor.lackovic@fer.hr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_26

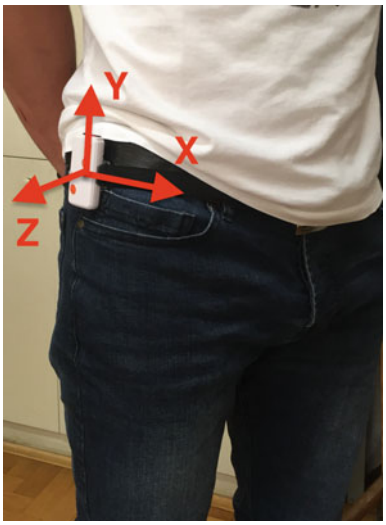



**Fig. 1** Sensor placement—Y axis was pointing up (vertical), X axis was pointing forward (antero-posterior plane) and Z axis was pointing away from the body (medio-lateral plane)

The subjects were equipped with a Shimmer3 (ShimmerSensing, Doublin, Ireland) inertial measurement unit (IMU). Shimmer3 contains a triaxial accelerometer, triaxial gyroscope, triaxial magnetometer and an altimeter [14]. Unit's small size and battery power supply make it suitable for data gathering in free living conditions. Shimmer3 contains two types of accelerometers: a Wide Range Accelerometer (with selectable measurement range from ±2 to ±16 g) and a Low Noise Accelerometer (working only at ±2 g but with less noise). For the purpose of this experiment, only Wide Range Accelerometer was used, at a selected range of ±8 g. During the experiments, data was streamed via Bluetooth to a PC with a frequency of 204.8 Hz.

During experiments, subjects wore the Shimmer3 IMU attached to their waist with a Velcro belt. Waist has been shown to be an optimal position for accelerometer based fall detection [13]. The IMU was fitted to the body with the axis oriented as shown in Fig. 1.

## Experiment Protocol

Subjects were asked to perform 15 tasks: 12 activities of daily living (ADLs) and 3 falls. The activities were: walking, fast walking, running, fast running, jumping, high jumping, sitting, standing up, lying down, getting up from lying position, walking down the stairs and walking up the stairs. Three types of falls were simulated by subject falling on a 2 cm thick tatami mat: forward fall, sideways fall and backward fall. The experiments were conducted in a safe laboratory environment. During each experimental session, data from one subject was recorded.

## Features

From the measured raw acceleration data, five different features were calculated [9]:

- sum vector magnitude (*SVM*)

$$SVM(n) = \sqrt{a_x(n)^2 + a_y(n)^2 + a_z(n)^2} \tag{1}$$

- differential *SVM* (*DSVM*)

$$DSVM(n) = ((a_x(n) - a_x(n-1))^2 + (a_y(n) - a_y(n-1))^2 + (a_z(n) - a_z(n-1))^2)^{\frac{1}{2}} \tag{2}$$

- Euler angle (*theta*) between the vertical device axis and the direction of gravitational field

$$theta(n) = \text{atan2}\left(\frac{\sqrt{a_x(n)^2 + a_z(n)^2}}{a_y(n)}\right) \tag{3}$$

- gravity weighted *SVM* (*GSVM*)

$$GSVM(n) = \frac{theta(n)}{90} SVM(n) \tag{4}$$

- gravity-weighted *DSVM* (*GDSVM*)

$$GDSVM(n) = \frac{theta(n)}{90} DSVM(n) \tag{5}$$

where $a_x(n)$, $a_y(n)$, $a_z(n)$ represent $x$, $y$ and $z$ acceleration components of the $n$-th sample. Each feature was calculated using accelerometer measurements from all subjects and experimental sessions. For the calculation of theta, a four-quadrant inverse tangent function was used that returns values in the closed interval $[-\pi, \pi]$.

## Performance Assessment

For all beforementioned features, optimal thresholds were saught. Data was gathered and processed with a GUI implemented in Matlab [15]. Every feature was compared to a set of thresholds. If the feature value at any time was greater than the threshold, this is classified as a potential fall event. With the knowledge of when actual falls happened, a

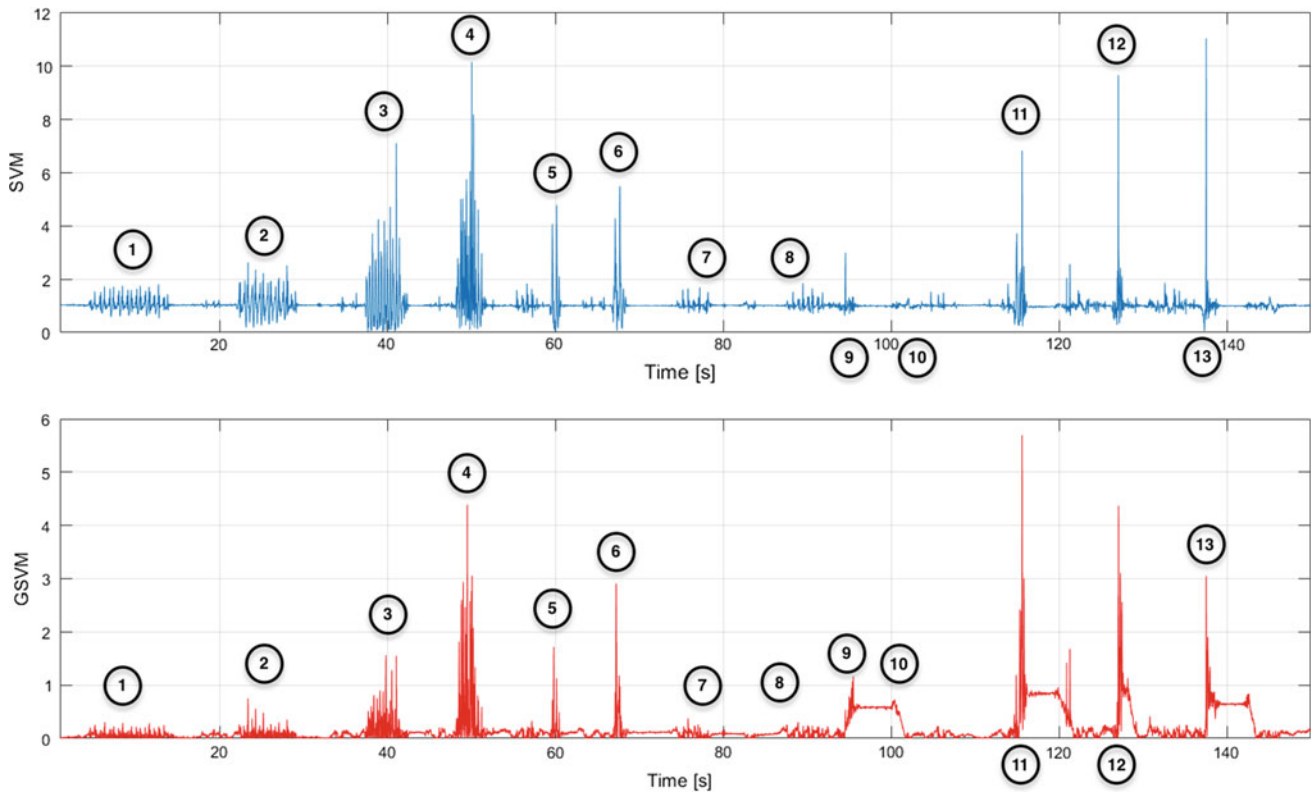



**Fig. 2** An example of *SVM* and *GSVM* signals acquired from one of the subjects. Numbers in circles denote the following activities: (1) walking, (2) fast walking, (3) running, (4) fast running, (5) jumping (6) high jump, (7) siting, (8) standing up from sitting, (9) lying down, (10) standing up from lying position, (11) falling forward, (12) falling sideways, (13) falling backward

contingency table was created for every threshold value containing the number of:

TP True positive events (fall happens, device detects)
FP False positive events (no fall, device detects)
TN True negative events (no fall, not detected)
FN False negative events (fall happens, not detected)

Using contingency table entries, sensitivity and specificity were calculated as:

$$sensitivity = \frac{TP}{TP + FN} \tag{6}$$

$$specificity = \frac{TN}{TN + FP} \tag{7}$$

When *sensitivity* (true positive rate) values are plotted against 1-*specificity* (false positive rate), a ROC curve is obtained for every feature. From the ROC curve, threshold can be chosen that balance between a desired sensitivity and specificity. Lower threshold values result in an increase of false positive alarms while higher thresholds lead to a decreased sensitivity to actual falls.

In this work, optimal threshold was calculated from the ROC curve as the value at which the distance $d$ between the curve and point (0, 1) is minimal. The point (0, 1) corresponds to an ideal case where both sensitivity and specificity

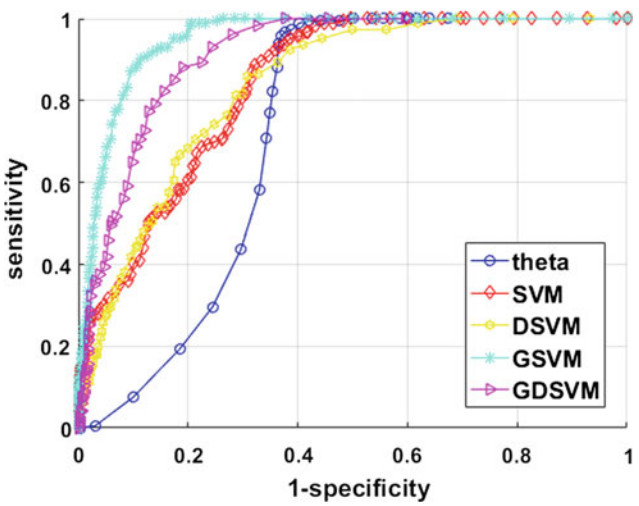



**Fig. 3** ROC curves for all five analyzed accelerometer-based features

**Table 1** Acceleration features with corresponding optimal threshold values and corresponding sensitivities/specificities

| Feature | Optimal threshold (g) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| SVM | 4.3 | 88.7 | 67.8 |
| DSVM | 1.2 | 86.0 | 69.1 |
| theta | 120 | 94.0 | 63.5 |
| GSVM | 3 | 90.3 | 88.5 |
| GDSVM | 0.8 | 88.1 | 80.9 |

would be 100%. Distance $d$ is thus calculated by the expression:

$$d = \sqrt{(1 - sensitivity)^2 + (specificity - 1)^2} \qquad (8)$$

## Results

An example of a typical *SVM* and *GSVM* signal acquired from one of the subjects is shown in Fig. 2. Activities of walking up and down the stairs are shown separately because they were recorded outside of the laboratory where all the other measurements were taken.

For the five features calculated from accelerometer measurements, ROC curves were plotted as shown in Fig. 3. Optimal threshold values were chosen for every feature by minimising the distance $d$ given by Eq. (8). Table 1 contains the values of optimal thresholds for every feature and the corresponding sensitivity and specificity. The values for all feature threshold, except for the angle theta, are expressed as the multiple of gravitation acceleration constant $g$. The threshold for theta is given in degrees. The highest sensitivity and specificity were obtained for the GSVM feature with values of 90.3% and 88.5% respectively when the threshold was 3 g.

## Conclusions and Future Work

This paper presents results of a threshold selection process for fall detection based on features extracted from accelerometer signals. We conducted a series of experiments with subjects performing ADL and simulated falls and created a database of recorded acceleration signals. For five acceleration-based features for fall detection we showed how to obtain optimal thresholds based on ROC analysis. From the five features selected, GSVM achieved the best sensitivity (90.3%) and specificity (88.5%) for the threshold value 3 g. The optimal threshold values were selected with the use of ROC curves.

The authors are aware of the limits of the experiments performed. Mostly young participants were involved in the study, although fall detection system are mostly aimed for the use in elderly population. Falls were simulated in safe laboratory settings, and although the participants were instructed to fall relaxed and avoid any compensatory strategies, it is possible that those falls may differ from real-life falls.

The main goal of this study was to establish a measurement environment for fall detection with results that could be compared to other groups in this research area. In our future work we plan to work with more complex threshold and machine learning algorithms for fall detection and explore measurements obtained from other research groups available in open databases for comparison.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. W.H.O Ageing and L.C. Unit (2008) WHO global report on falls prevention in older age. World Health Organization, Geneva, CH
2. Walker J, Howland J (1991) Falls and fear of falling among elderly persons living in the community: occupational therapy interventions. Am J Occup Ther 45(2):119–122
3. Patel S, Park H, Bonato P et al (2012) A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 9(1):21–38
4. Sabatini AM, Ligorio G, Mannini A, Genovese V (2015) Prior-to- and post-impact fall detection using inertial and barometric altimeter measurements. IEEE Trans Neural Syst Rehabil Eng, accepted for publication
5. Chen KH, Yang J-J, Fu-Shan J (2016) Accelerometer-based fall detection using feature extraction and support vector machine algorithms. Instrum Sci Technol 44(4):333–342
6. He J, Shuang B, Wang X (2017) An unobtrusive fall detection and alerting system based on Kalman filter and Bayes network classifier. Sensors (Basel) 17(6)
7. Pierleoni P, Belli A, Maurizi L et al (2016) A wearable fall detector for elderly people based on AHRS and barometric sensor. IEEE Sens J 16(17):6733–6744
8. Medrano C, Igual R, Garcia-Magarino I et al (2017) Combining novelty detectors to improve accelerometer-based fall detection. Med Biol Eng Comput 55(10):1849–1858
9. Lim D, Park C, Kim NH et al (2014) Fall-detection algorithm using 3-axis acceleration: combination with simple threshold and hidden Markov model. J Appl Math (2014)
10. Bourke AK, O'Brien JV, Lyons GM (2007) Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. Gait Posture 26(2):194–199

11. Kangas M, Konttila A, Winblad I, Jamsa T (2007) Determination of simple thresholds for accelerometry-based parameters for fall detection. In: Proceedings of the 29th annual international conference of the IEEE EMBS, Lyon, France, 23–26 Aug 2007
12. Bourke AK, Lyons G (2008) A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Med Eng Phys 30 (1):84–90
13. Pannurat N, Thiemjarus S, Nantajeewarawat E (2014) Automatic fall monitoring: a review. Sensors (Basel) 14(7):12900–12936
14. Shimmer Sensing Webpage. www.shimmersensing.com. Accessed 15 Sept 2017
15. Vugrin J (2017) Fall detection system for the elderly based on wearable wireless sensors. MS thesis, University of Zagreb, Faculty of Electrical Engineering and Computing (in Croatian)

# Camera Based Real Time Fall Detection Using Pattern Classification

M. Macaš, S. Lesoin, and A. Périn

**Abstract**

A complete real-time fall detection system is presented consisting of camera data acquisition, image processing, pattern recognition, fall alarming, and web interface. Classifiers are trained using only three input features extracted from each video frame using image processing. Among linear and quadratic Bayes, Parzen classifier and 3-nearest neighbors classifier, the last one performed best on a testing set from sensitivity, specificity and area under receiver operating characteristic point of view. Such frame classification is further used in a simple rule triggering the fall alarm process. The fall alarm tested in a real time scenario with 40 falls performed by four persons detected all the falls while having one false positive case.

**Keywords**

Fall detection • Elderly • Pattern recognition • Camera • Feature extraction

## Introduction

The World Health Organization (WHO) published many important statistical facts related to falls.[1] Falls are one of the leading injury deaths causes that lead to more than 600 000 deaths per year. Elderly people are in a special risk of fall related injuries often followed by several complications also caused by a time delay between the fall and subsequent aid. For this reason, it is crucially important to detect the fall as quickly as possible. This is a sufficient motivation for creating systems that detect a fall immediately and automatically send an alarm to a responsible person. This paper focuses on scenario with a single camera in a room connected to an application which detects a fall from the image in real time and eventually sends an alarm to registered emergency contacts. The method extracts only three features and uses simple and fast classifiers for fall detection.

Probably the most popular method for processing the raw camera image is the background subtraction [1], which is used for differentiating the foreground from background. Further, some features are typically extracted characterising the localized silhouette of the person. Different posture representations were tested in literature. Most common is a bounding box fitted on the foreground image [1]. In [2], an ellipse fitting was more accurate than the rectangle box. A representation using three key points were used in [3], where the key points were found as the centroids of three different regions of the foreground. Histogram of maximal optical flow projection was proposed in [4] for crowded scenes. Many different features were proposed in literature [5]: position and behavior of the barycenter (ventre of mass) of the silhouette, geometrical orientation of the silhouette

[1] http://www.who.int/mediacentre/factsheets/fs344/en/.

M. Macaš (✉)
Czech Institute of Informatics, Robotics and Cybernetics in Prague, Czech Technical University, Zikova street 1903/4, Prague, Czech Republic
e-mail: martin.macas@cvut.cz

S. Lesoin · A. Périn
Institut supérieur de l'électronique et du Numérique, Brest, France



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_27

(e.g. angle between an object's bounding box and the ground, projection of the silhouette's shape on the x and y axes, variances in vertical and horizontal directions, ration of the bounding box height and width), projection histogram along the axis of the ellipse (the local features) and the ratio between the major axis and the minor axis, color features, change in illumination. Shape based fall characterization is carried out with a so called curvature scale space features and Fisher vector encoding in [6]. The extracted features are further used for deciding if there is a fall or not. Sometimes, a decision rule is created empirically like in [1]. More often, some automatic classifier trained on collected data is used like support vector machines [7], ELM [8] and MLP decision tree [9] or Naive Bayes classifier [9]. Many different classifiers were tested in [10] and the best among them was support vector machine. Next, the paper describes a proposal of the feature based pattern recognition fall detection system.

## Methods

This section describes all components of the presented approach. HTML, C++, JavaScript and PHP were used for the development of the system. Moreover, the database was based on MySQL. The web site is hosted with the free web hosting 000webhost powered by Hostinger.[2] Materialize front-end framework from Google was used to make the web site development faster and easier.

### Camera and Image Capture

Vivotek IP camera IP8131W was used for testing. The camera is equipped with infra-red LEDs which provide a good video quality irrespectively to luminosity conditions. The resolution is $1280 \times 800$ and the frame rate is 30 fps. The camera can be connected by Wi-Fi or by cable. The protocol used to stream the video from the camera is Real Time Streaming Protocol.

After an initialization, all the processes are performed in an infinite loop, where each iteration takes 30 ms. At each such iteration, the current frame image is captured and stored in a matrix form, replacing the previous frame. For each frame, the person is detected using a background subtraction. Further, the person is represented by its barycenter (center of gravity) and ellipse fitted to the moving pixels. Although bounding rotated rectangle is also computed, it is not used in the presented system. The analysis of parameters of those representations are used to detect falls.

[2] http://smarthomeenvironment.000webhostapp.com/index.php.

### Detection of Moving Pixels

The image processing was performed using Open Source Computer Vision Library version 3.2 (OpenCV) in C++. HighGUI module was used to create and manipulate windows that can display images and "remember" their content, add trackbars to the windows, handle simple mouse events as well as keyboard commands. Video module was used for motion estimation, background subtraction, and object tracking algorithms.

Gaussian Mixture-based Background/Foreground Segmentation [11] was used for background subtraction to detect the movement and moving person. The background "historic" parameter—number of frames from which the background is calculated—was set to 400. The background is subtracted from the current frame to get a binary image describing which pixels are moving and which form the background. Next, two different morphological operators were applied to the binary image to improve its quality—erosion and dilatation. Erosion deletes small groups of pixels which probably do not correspond to the moving person. Dilatation expands the shape of the moving object. After, the mass centre of the moving object and its contour are computed.

### Features

The barycenter was calculated by averaging the coordinates of all white pixels in the binary image obtained before. To asses the dynamic behavior of the barycenter, the trajectory of the barycenter was computed. Velocity vector is computed as the difference between two consecutive positions (on two consecutive frames) of the barycenter. The slope of those vectors are calculated that describe the direction in which the barycenter moves. Five latest slopes are kept in memory. To filter out some noisy positions of the barycenter, an average of those last five values of the slope is calculated and used as the first input feature for classification abbreviated as SLOPE. Further, maximum vertical acceleration of the barycenter computed over last 15 positions is used as the second feature abbreviated as Y-ACC. It is expected that high values of this feature will probably indicate the fall.

Next, a contour is computed from the frame and represented by a vector of coordinates of pixels forming the contour. `findContour()` function from OpenCV was applied to the binary image for this purpose. Only the extreme outer contours were used. If the in-motion contour detected is long enough to be a human (more than 1000 pixels, assuming that the camera is installed indoor within a room of a common size, i.e. less than 100 $m^2$), the ellipse was fitted to the contour. The difference angle between the

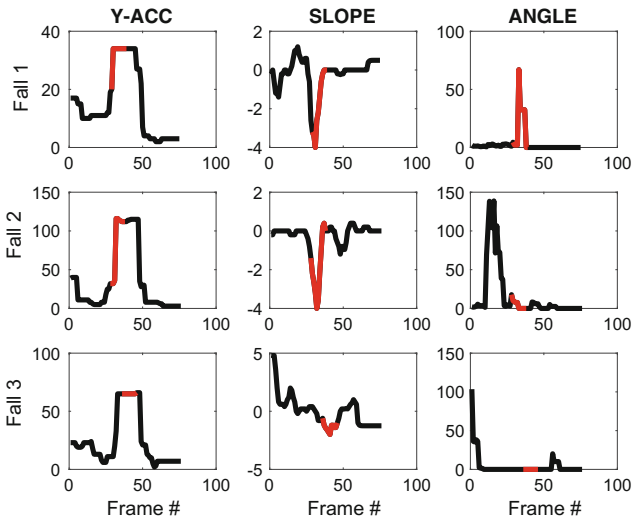



**Fig. 1** Three examples of temporal evolution of particular features during a fall

major axis of the ellipse and vertical axis was used as the third feature abbreviated as ANGLE (Fig. 1).

## Frame Classification

Next task is to detect the fall from the particular frame. This corresponds to a classification task. As described above, only three features were used for the classification—the average velocity slope, maximal vertical acceleration and the ellipse angle. Statistical pattern recognition methods were used for the classification. There are obviously many possible classifiers. We focused on simple and less complex pattern recognition methods. We tested four different classifiers [12].

Linear Bayes classifier (LB) is a simplest Bayesian theory based classifier utilizing the Bayes formula for computing the aposterior probabilities for each class from likelihoods and priors and classifying according to the maximum aposterior probability criterion. It assumes that likelihoods have normal distributions with same covariance matrix for each class. This leads to a linear decision boundary. Quadratic Bayes classifier (QB) is the same like previous one, but different covariance matrices are assumed for different classes. The decision boundary is a quadratic curve. Parzen classifier (PC) estimates the likelihoods using Parzen window density estimation. 3-Nearest Neighbors classifier (3NN) assigns an unlabeled data instance into the same class as is the class of majority of its three nearest neighbors.

## Fall Alarm Triggering

The system continuously analyses the frames and detects if the particular frame corresponds to fall Fig. (2). To prevent disturbing alarming after every false positive classification of frame, one needs to create a robust decision rule which triggers on the alarm event. If a certain number $N_1$ of adjacent frames is classified as fall, the fall alarm event is activated. With low value of $N_1$, the risk of a false positive alarm increases. If $N_1$ is too large, the system can erroneously ignore a fall event. After the fall alarm event is triggered on, the classification of individual frames starts to be ignored for the next $N_2$ frames. If $N_2$ is too low, the systems will unnecessarily provide multiple alarms for one fall. If $N_2$ is too large, the system can miss two adjacent falls. Values of those parameters, which lead to a small error on the training set were set empirically to $N_1 = 3$ and $N_2 = 70$.

## Web Interface

The implementation of the web interface consists of log-in page, home page, emergency contacts management page, fall alarm page and a database.

The aim of the log-in page is to provide a secure access for users. A log-in system enables the system to identify the user and then display information according to the registered


(a) Negative example (no fall) (b) Positive example (fall)

**Fig. 2** Examples of the image from camera, detected barycenter and fitted ellipse. **a** Negative example (no fall). **b** Positive example (fall)

data. Several administrators can be registered in the system. Two different levels of access are available for users. First, the full access is provided for the resident of the home, or any other responsible person. It enables the user to have access to the stream of the camera, to manage the emergency contacts and to switch on or off the alarms. Secondly, a limited access is provided for the emergency contacts that are only informed about the fall and have access to the stream of the camera they are registered for.

The log-in management was implemented in PHP The database has been made to register and manage all the users. After a successful log-in, a home page is displayed. Its main task is to provide to the user an access to the streamed video from the camera. In case of fall detection, an alarm is sent by mail to the emergency contacts. Emergency contacts can then log-into the web site to check what happened via the streamed video. The purpose of the emergency contacts management page is to enable the administrator to manage all the emergency contacts (e.g. adding or deleting contacts, changing e-mail addresses or retrieving forgotten passwords). Finally, the fall alarm page is not available for users and it is only used to send the alarm e-mail when the fall detection algorithm triggers the fall alarm. It sends a request to the web page which sends an e-mail to the emergency contacts in order to inform them about the fall.

## Experiments

To design a suitable classifier, a data set was created by recording 50 falls and extracting one data instance for each frame. The evolution of each feature during 3 different cases of fall is depicted in Fig. 1. To train and validate the classifiers, we used the previously measured data set with three features described above and labeled by 0 for non-falls and 1 for falls. The final data contained 7396 data instances of which 532 were positive (part of a fall) and 6864 were negative. The data set is visualized in Fig. 3, where a scatter plot for each pair of three input features can be found.

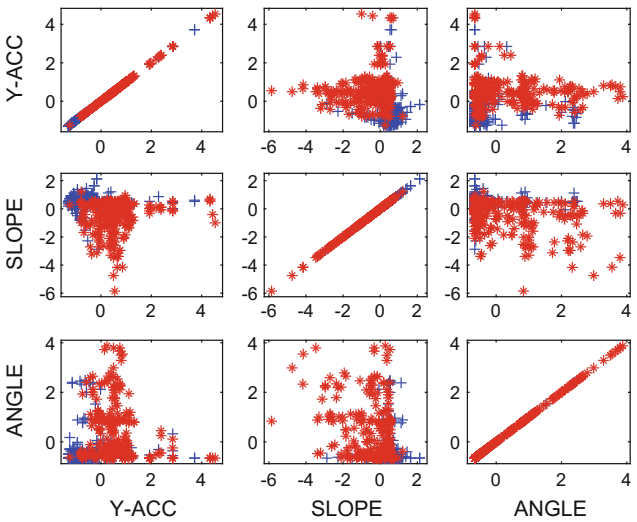



**Fig. 3** Scatter plot of the data set

## Experimental Setup

The whole data were randomly split into two different subsets of the same size. One was used for training the classifier and one was used for evaluation of classification performances. The training set was balanced to have the uniform number of classes by randomly deleting excessive instances from the majority non-fall class. We used sensitivity, specificity and area under receiver operating characteristic (AUC) for evaluation. Before training, a scaling transform was computed on training data (to zero mean and unitary standard deviation) and applied to both training and testing data to prevent domination of features with higher values.

## Results

The final classification performances are summarized in Table 1. One can see that the 3NN classifier performed the best since it had relatively high and balanced sensitivity and

**Table 1** Final testing sensitivity, specificity and area under receiver operating characteristic for four different classifiers

| | SE | SP | AUC |
|---|---|---|---|
| LB | 85.8555 | 79.3103 | 0.7653 |
| QB | 84.7437 | 67.4330 | 0.7727 |
| PC | 81.7789 | 83.9080 | 0.7240 |
| 3NN | 81.0994 | 85.8238 | 0.8987 |

specificity values and the highest AUC measure. Especially from the AUC point of view, all the distribution estimation based classifiers were significantly worse than much simpler nearest neighbor method.

The results provided above are computed on the frame basis. It means that the do not evaluate the detection of fall, but detection of frame that includes one moment of the fall. To evaluate the success of the fall detection, we let the whole system running one hour. During this one hour, four different people were coming into the scene one by one, doing different activities and different types of falls (sideway, forward, backward). In total 40 falls were performed during the hour. It is worth to note that all of them were successfully detected. Moreover, there was one false alarm produced when the target person was quickly running down the steps. This is obviously not problem for scenarios without steps or with elderly people. The system was validated and can be reliably used for real fall detection.

## Conclusions

There are many points that could be improved. For example, many other different features could be extracted and eventually automatically selected. Nevertheless, it seems that machine learning techniques can be useful and hide a great potential within the fall detection area.

**Acknowledgements** The research is supported by Ministry of Industry and Trade of the Czech republic grant no. FV20696.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Vishwakarma V, Mandal C, Sural S (2007) Automatic detection of human fall in video. Pattern Recogn Mach Intell 616–623
2. Yu M, Rhuma A, Naqvi SM, Wang L, Chambers J (2012) A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. IEEE Trans Inf Technol Biomed 16:1274–1286
3. Chua JL, Chang YC, Lim WK (2015) A simple vision-based fall detection technique for indoor video surveillance. Signal Image Video Process 9:623–633
4. Li A, Miao Z, Cen Y, Wang T, Voronin V (2015) Histogram of maximal optical flow projection for abnormal events detection in crowded scenes. Int J Distrib Sensor Netw 11:1–11
5. Igual R, Medrano C (2013) Challenges, issues and trends in fall detection systems. Biomed Eng Online 12:66
6. Aslan M, Sengur A, Xiao Y, Wang H, Ince MC, Ma X (2015) Shape feature encoding via fisher vector for efficient fall detection in depth-videos. Appl Soft Comput 37:1023–1028
7. Fan K, Wang P, Hu Y, Dou B (2017) Fall detection via human posture representation and support vector machine. Int J Distrib Sensor Netw 13:1–21
8. Ma X, Wang H, Xue B, Zhou M, Ji B, Li Y (2014) Depth-based human fall detection via shape features and improved extreme learning machine. IEEE J Biomed Health Inform 18:1915–1922
9. Tapia EM, Intille SS, Haskell W et al (2007) Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: 2007 11th IEEE international symposium on wearable computers, pp 37–40
10. Luštrek M, Kaluža B (2009) Fall detection and activity recognition with machine learning. Informatica 33
11. Zivkovic Z, Van Der Heijden F (2006) Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recogn Lett 27:773–780
12. Lei B, Xu G, Feng M et al (2017) Classification, parameter estimation and state estimation: an engineering approach using MATLAB. Wiley

# Part VI
# Machine Learning and Predictive Models in Medicine

# Epileptic Seizures Classification Based on Long-Term EEG Signal Wavelet Analysis

K. D. Tzimourta, A. T. Tzallas, N. Giannakeas, L. G. Astrakas, D. G. Tsalikakis, and M. G. Tsipouras

**Abstract**

Epilepsy is a complex neurological disorder recognized by abnormal synchronization of cerebral neurons, named seizures. During the last decades, significant progress has been done in automated detection and prediction of seizures, aiming to develop personalized closed-loop intervention systems. In this paper, a methodology for automated seizure detection based on Discrete Wavelet Transform (DWT) is presented. Twenty-one intracranial ictal recordings acquired from the database of University Hospital of Freiburg are firstly segmented in 2 s epochs. Then, a five-level decomposition is applied in each segment and five features are extracted from the wavelet coefficients. The extracted feature vector is used to train a Support Vector Machines (SVM) classifier. Average sensitivity and specificity reached above 93% and 99% respectively.

**Keywords**

Electroencephalogram (EEG) • Seizure • Epilepsy • Discrete wavelet transform (DWT)

## Introduction

One of the most challenging brain disorders that has gained increasing attention the last decades is epilepsy. Epilepsy is characterized by recurrent seizures, which are brief episodes of involuntary movement caused by excessive electrical discharges in a group of brain cells. According to the latest World Health Organization (WHO) reports, epilepsy affects almost 1% of the world's population and is estimated that about 2.4 million people are diagnosed with epilepsy each year [1]. Furthermore, about 30% of children and adults suffering from seizure episodes are left untreated and without anti-epileptic drugs (AED).

The diagnosis and monitoring of seizures is done through neuroimaging and electrophysiological techniques. The electroencephalogram (EEG) is the diagnostic tool that continuously records the brain's electrical activity using electrodes as sensors to detect fluctuations of the emitted electric charges [2]. Based on the location of electrodes, EEG is discriminated in scalp EEG (sEEG), in which the electrodes are placed in the surface, and intracranial EEG (iEEG), in which the electrodes are placed invasively inside the brain.

Computerized methods and automated seizure detection systems have been developed utilizing different EEG databases, owing to the complexity of the disorder coupled with the multiple drawbacks of the visual inspection by neurophysiologists. A variety of methods has been validated with the database of epilepsy center of the University of Bonn, which is consisted of short-term (23.6 s) scalp and intracranial EEG recordings. However, a long-term dataset

K. D. Tzimourta · L. G. Astrakas
Medical Physics Laboratory, University of Ioannina, 45110 Ioannina, Greece

A. T. Tzallas (✉) · N. Giannakeas · M. G. Tsipouras
Department of Computer Engineering, Technological Educational Institute of Epirus, Kostakioi, 47100 Arta, Greece
e-mail: tzallas@teiep.gr

D. G. Tsalikakis · M. G. Tsipouras
Department of Informatics and Telecommunications Engineering, University of Western Macedonia, 50100 Kozani, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_28

such as the one of Epilepsy Center of the University Hospital of Freiburg is closer to clinical recordings and provides more information for further developing seizure prediction algorithms.

Freiburg database has been extensively used by research groups worldwide. Different methodologies have been proposed including Wavelet Transform, Empirical Mode Decomposition (EMD) [3, 4] Principal Component Analysis (PCA) [5], Independent Component Analysis (ICA) [6], Fractal analysis [7, 8] or Fuzzy systems [9, 10]. The majority of them concur with a two-stage procedure, following a pattern recognition approach: feature extraction and classification.

Particularly, the Discrete Wavelet Transform (DWT) has been adopted by many researchers to decompose the recordings in certain sub-bands [11–15]. Then, significant features such as the coastline and Hjorth variance [11], the relative energy, the relative amplitude, the fluctuation index, the coefficient of variation [12], the wavelet variances [14], the lacunarity and the fluctuation index [13] or the diffusion distances [15] were extracted from the resulting signals. The extracted set of features was used as input to train either a Support Vector Machines (SVM) [12] or a Bayesian linear discriminant classifier (BLDA) [13, 15]. Finally, in study [14] Xie and Krishnan evaluated the performance of various classifiers (k-Nearest Neighbor, Fisher's linear discriminant, SVM) whereas in [11] a rule based approach was preferred instead of a classifier.

In this paper, an automated seizure detection methodology is presented based on DWT in order to divide EEG recordings to specific subbands and extract several features. Subsequently, these features are given as an input feature vector to train a SVM classifier. The methodology has been evaluated on 21 long-term intracranial EEG recordings for a binary classification problem and results are presented.

## Materials and Method

The proposed work consists of four stages: segmentation, wavelet analysis, feature extraction and classification. In the first stage, a long-term EEG recording from ictal activity of each patient is segmented into 2 s windows. After that, a 5-level wavelet-decomposition is applied in each EEG segment dividing every signal into several frequency subbands. In the next stage, 5 features are calculated from each subband creating a feature vector. Finally, the feature vector is used to train a SVM classifier. In Fig. 1, a concise diagram of the proposed methodology is presented.

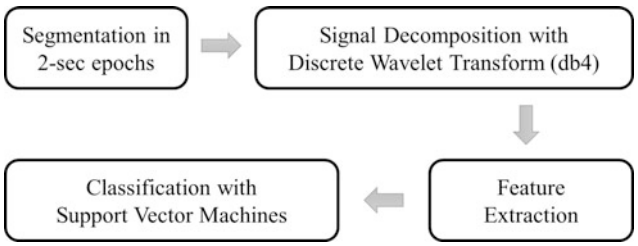



**Fig. 1** A brief diagram of the proposed methodology

## The Database

The methodology has been trained and tested on invasive EEG recordings from 21 patients suffering from medically intractable focal epilepsy. The dataset comes from the Epilepsy Center of the University Hospital of Freiburg and is now available through the EPILEPSIAE project [16]. The available data included six intracranial EEG channels (three focal and three extra-focal electrodes).

The EEG recordings obtained from 21 patients are separated into files of ictal (the period with seizure onset), preictal (the period before seizure onset) and interictal (the period between seizures) activity. Two–five seizure episodes were recorded for each patient lasting from several seconds to a few minutes. A total of 87 seizures, 509 h of interictal and 199 h of both pre-ictal and ictal EEG data are included in this large dataset.

In this methodology, only the first channel of an ictal recording of each patient is used, since the ictal recordings would contain more epileptic components and would provide better discrimination of "seizure" and "non-seizure" activity.

## Preprocessing

The long-term EEG channel of each patient was initially divided into 2 s (512 samples), non-overlapping epochs leading to 1800 segments per patient. This window size proved to be the optimal after testing potential window sizes. Since the seizure duration ranges among [4.21–1071.5] s, the 2 s window was chosen aiming to accurately capture the subtle changes of EEG. Afterwards, the Discrete Wavelet Transform was applied in each one of the resulting segments.

## Discrete Wavelet Transform

Wavelet Transform (WT) has gained significant ground in automated seizure detection scheme and is widely applied in numerous seizure detection studies.

**Table 1** Frequency ranges with the corresponding wavelet decomposition levels

| Frequency range (Hz) | Decomposed signal |
|---|---|
| 64–128 | D1 |
| 32–64 | D2 |
| 16–32 | D3 |
| 8–16 | D4 |
| 4–8 | D5 |
| 0–4 | A5 |

According to Wavelet Analysis [17], a signal can be represented by a linear combination of a particular set of functions, obtained by dilating and translating a single function. This function is called mother wavelet and is used to decompose the initial signal into sub-signals of half its size and spectrum.

In Discrete Wavelet Transform (DWT) the scaling and translating parameters are represented in powers of two. The implementation of the DWT uses a series of quadrature mirror filters (QMF) described as high-pass and low-pass filters. In the first level of DWT, the input signal is simultaneously passed through the conjugate low and high pass filters. The obtained outputs are a set of coefficients called wavelet coefficients. The output of the low-pass filter, namely approximation, is sub-decomposed, whereas the output of the high-pass filter, namely detail, is not. The same procedure is recursively repeated, forming a single-side, pyramid-like architecture.

Choosing the number of decomposition levels and the appropriate mother wavelet is of primary importance. The number of decomposition levels is chosen based on the dominant frequency. The best mother wavelet function was selected, mainly among the Daubechies wavelets, after visual examination. In this study, a 5-level-decomposition transform is used and the family of Daubechies wavelets of order 4 (db4) is selected to decompose the signal. Table 1 shows the corresponding frequencies to the resulting decomposition levels.

## Feature Extraction

In order to minimize the complexity and the computational time of the proposed methodology, the most representative and significant characteristics were extracted. Thus, in the present study five features were calculated in each decomposition level, namely energy, entropy, standard deviation, variance and mean of the absolute values of the wavelet coefficients. The final low-dimensional feature vector was used as an input to train an SVM classifier.

## Classification

Support Vector Machines (SVM) is a machine learning technique for binary classifications problems. According to Machine Learning [18], non-linear instances that need to be classified are mapped to a high-dimension feature space. In this feature space instances are separated by a very clear gap, named hyperplane. The vectors that lie on the margin are called support vectors and they constitute the critical elements of the training set for the classification problem. The basic idea underlying this technique is to locate an Optimal Separating Hyperplane, which maximizes the distance between the margin and the support vectors and minimizes the classification error, in a projection space by solving a quadratic optimization problem. The kernel function that may be a linear, radial basis function (RBF), polynomial, or sigmoid kernels is responsible for the transformation to the higher dimensional space. In this study, RBF kernel function was used. Furthermore, two parameters were optimized in order to optimize the algorithm and therefore the classification results. A grid search was performed on the parameters C and $\gamma$, which are related with margin boundaries and the RBF kernel function respectively, using cross-validation.

## Results

The Freiburg database is one of the most comprehensive, long-term datasets and a variety of classification problems are addressed based on its recordings. In an attempt to identify seizure episodes, one ictal recording from each one of the 21 patients was used. Since the onset and the offset of seizure episodes are known, the epochs between the seizure onset and offset were marked as “seizure” and the rest of the epochs as “non-seizure”, forming the corresponding classes. Also, the 2 s duration epochs adjacent to the onset and offset were excluded from the subsequent processing and no annotation was added.

To validate the experiments, the 10-fold cross-validation technique was employed. In some patients, the non-seizure instances were tremendously more than the seizure ones, leading to unbalanced data and poor training of the classifier. Consequently, the data were rounded up by manually repeating seizure instances, until the seizure instances were approximately 10% of the non-seizure. Sensitivity, specificity and overall accuracy were calculated from the number of correctly/incorrectly classified instances, for evaluation of the classification performance. The obtained statistical results for each patient are described in Table 2. The best sensitivity (100%) was achieved for half of the patients (patients 1, 7, 9, 12, 13, 17, 18, 19, 20 and 21), while the lowest sensitivity was 45.30% obtained from the

**Table 2** Sensitivity, specificity and accuracy results for each patient

| Patient | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| 1 | 100 | 100 | 100 |
| 2 | 95.24 | 99.83 | 99.37 |
| 3 | 97.80 | 99.83 | 99.64 |
| 4 | 98.90 | 100 | 99.90 |
| 5 | 99.44 | 100 | 99.95 |
| 6 | 98.58 | 99.77 | 99.64 |
| 7 | 100 | 99.94 | 99.95 |
| 8 | 96.72 | 100 | 99.69 |
| 9 | 100 | 100 | 100 |
| 10 | 95.03 | 99.65 | 99.21 |
| 11 | 45.30 | 99.20 | 93.78 |
| 12 | 100 | 99.94 | 99.95 |
| 13 | 100 | 100 | 100 |
| 14 | 87.85 | 99.65 | 98.53 |
| 15 | 60.99 | 99.45 | 95.61 |
| 16 | 93.37 | 100 | 99.37 |
| 17 | 100 | 99.94 | 99.95 |
| 18 | 100 | 100 | 100 |
| 19 | 100 | 100 | 100 |
| 20 | 100 | 100 | 100 |
| 21 | 100 | 99.89 | 99.90 |
| Total | 93.77 | 99.86 | 99.26 |

classification of the patient 11. Specificity was also high, reaching above 99% for all patients and ten patients among them reached 100%.

## Discussion and Conclusions

In the present study, a wavelet-based methodology for automated seizure detection is presented. Twenty-one EEG recordings of 1-hour-long duration was initially segmented in 2 s epochs. Then, a 5-level-decomposition transform was applied in each segment using the 'db4' as mother wavelet. Five features namely, energy, entropy, mean of the absolute values of the wavelet coefficients, standard deviation and variance, were extracted in each subband of interest, creating the feature set that trained a SVM classifier. Seizure and non-seizure epochs were adequately classified and the obtained results are presented in Table 1.

Table 3 shows a comparison between the proposed method and other recent DWT-based methods that have been validated on Freiburg database. It can be seen, that this methodology shows really promising results in correctly identifying seizures. Average sensitivity is slightly lower than the one obtained in other approaches. However, this work has been tested only to a small part of the dataset, and more tests should be done to improve the method's performance.

Undoubtedly, epilepsy is a critical brain disorder that can lead to severe and life-threatening conditions, if left uncontrolled. The Freiburg database contains a large and comprehensive amount of data and further extensive study should be conducted. In the direction of developing robust seizure detection and prediction methods and personalized closed-loop treatment systems [19], different approaches, including the evaluation of various classifiers and the combination of linear and nonlinear features, should be examined.

**Table 3** A comparison of performances of the various methods proposed in the literature for the detection of seizures using different data of the Freiburg database

| Related studies | No. of patients/no. of analyzed seizures | Length (h) | Proposed method | Features | Classification problem | Sensitivity | Specificity | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Liu et al. [12] | 21/82 | 80.35 | DWT | Relative energy, fluctuation index, coefficient of variation | Seizure/non-seizure | 94.46 | 95.26 | 95.33 |
| Zhou et al. [13] | 21/81 | 289.14 | DWT | Lacunarity and fluctuation index | Seizure/non-seizure | 96.25 | 96.70 | 96.67 |
| Xie et al. [14] | 4/4 | 8 | DWT | Wavelet variances | Ictal/interictal | – | – | 99.00 |
| Yuan et al. [15] | 21/87 | 597.95 | DWT | Diffusion distances | Seizure/non-seizure | 95.11 | 98.78 | 98.77 |
| Proposed method | 21/21 | 20.57 | DWT | Energy, entropy, mean, variance, standard deviation | Seizure/non-seizure | 93.70 | 99.86 | 99.26 |

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. WHO (2012) World Health Organization, "Epilepsy", Fact sheet N999. http://www.who.int/mediacentre/factsheets/fs999/en/. Accessed 10 Sep 2017
2. Tsiouris KM, Tzallas AT, Markoula S, Koutsouris D, Konitsiotis S, Fotiadis DI (2015) A review of automated methodologies for the detection of epileptic episodes using long-term EEG signals. Handbook of research on trends in the diagnosis and treatment of chronic conditions, pp 231–261
3. Bajaj V, Pachori RB (2013) Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of EEG signals. Biomed Eng Lett 3(1):17–21
4. Parvez MZ, Paul M (2014) Epileptic seizure detection by analyzing EEG signals using different transformation techniques. Neurocomputing 145:190–200
5. Kevric J, Subasi A (2014) The effect of multiscale PCA de-noising in epileptic seizure detection. J Med Syst 38(10):131
6. Liang SF, Chen YC, Wang YL, Chen PT, Yang CH, Chiueh H (2013) A hierarchical approach for online temporal lobe seizure detection in long-term intracranial EEG recordings. J Neural Eng 10(4):045004
7. Yuan Q, Zhou W, Liu Y, Wang J (2012) Epileptic seizure detection with linear and nonlinear features. Epilepsy Behav 24 (4):415–421
8. Zhang Y, Zhou W, Yuan S (2015) Multifractal analysis and relevance vector machine-based automatic seizure detection in intracranial EEG. Int J Neural Syst 25(06):1550020
9. Rabbi AF, Fazel-Rezai R (2012) A fuzzy logic system for seizure onset detection in intracranial EEG. Comput Intell Neurosci 2012:1
10. Geng D, Zhou W, Zhang Y, Geng S (2016) Epileptic seizure detection based on improved wavelet neural networks in long-term intracranial EEG. Biocybern Biomed Eng 36(2):375–384
11. Raghunathan S, Jaitli A, Irazoqui PP (2011) Multistage seizure detection techniques optimized for low-power hardware platforms. Epilepsy Behav 22:S61–S68
12. Liu Y, Zhou W, Yuan Q, Chen S (2012) Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. IEEE Trans Neural Syst Rehabil Eng 20 (6):749–755
13. Zhou W, Liu Y, Yuan Q, Li X (2013) Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG. IEEE Trans Biomed Eng 60(12):3375–3381
14. Xie S, Krishnan S (2013) Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. Med Biol Eng Comput 51(1–2):49–60
15. Yuan S, Zhou W, Yuan Q, Zhang Y, Meng Q (2014) Automatic seizure detection using diffusion distance and BLDA in intracranial EEG. Epilepsy Behav 31:339–345
16. Freiburg seizure prediction project. Freiburg, Germany (2008). http://epilepsy.uni-freburg.de/freiburg-seizure-prediction-project/eeg-database
17. Mallat S (1999) A wavelet tour of signal processing. Academic press
18. Theodoridis S, Pikrakis A, Koutroumbas K, Cavouras D (2010) Introduction to pattern recognition: a matlab approach. Academic Press
19. Nagaraj V, Lee S, Krook-Magnuson E, Soltesz I, Benquet P, Irazoqui P, Netoff T (2015) The future of seizure prediction and intervention: closing the loop. J Clin Neurophysiol: Off Publ Am Electroencephalogr Soc 32(3):194

# Heartrate Variability Comparison Between Electrocardiogram, Photoplethysmogram and Ballistic Pulse Waveforms at Fiducial Points

G. M. W. Janjua, R. Hadia, D. Guldenring, D. D. Finlay, and J. A. D. McLaughlin

**Abstract**

Heart rate variability analysis (HRVA) gives valuable insight to the cardiovascular system. Electrocardiogram (ECG) based HRVA has been assessment gold standard but eavesdropping of wearable technology requires the comparison of its surrogacy to an accepted standard. In this study, optical and mechanical measures at distal artery waveform are compared to the electrical signal of the heart. The sensor data of the six healthy volunteers are collated and compared at fiducial points in various time, frequency and non-linear domains for HRVA. We have found that during early systole fiducial location on waveforms can be surrogate to ECG standard and mechanical sensor 2nd derivative proved to be the best among them. Also, the comparative technology shows enormous potential for cardiovascular diagnostic.

**Keywords**

Heartrate variability • Electrocardiogram • Photoplethysmogram • Ballistic pulse pressure Fiducial point

## Introduction

Cardiovascular diseases caused by a change in heart rhythm are the leading cause of morbidity and mortality in this modern era, globally. It accounts for approximately 31% of deaths worldwide [1]. Non-invasive physiological signs, like blood pressure (BP) variation, HRV and arterial stiffness are few of the widely used, as simplest to acquire vital physiological signs in accessing the cardiovascular health [2]. HRV, the change in the time interval between adjacent heartbeats, is one of the most prominent noninvasive markers to establish the meaningful relationship between the autonomic nervous system and cardiovascular mortality, including sudden cardiac deaths [3]. Most commonly used techniques for HRVA is based on RR interval variability (RRIV) on the ECG waveform and by the advent of wearable technology peripheral photoplethysmograph (PPG) pulse rate variability (PRV) analysis is also becoming famous. The most commonly used and 'Gold standard' HRVA measurement method is the RR interval variability measurement from ECG [4]. The rapid increase in aging population and the increase in demand for home healthcare monitoring have increased research interest in wearable healthcare technologies. Within the last decade, many studies have been carried out to realize the real-time acquisition of different physiological parameters using wearable sensors [5]. Also, in recent times, a lot of work has been done in the direction to implement easy to use, noninvasive single acquisition location wearable technologies aimed at assessing and improving cardiovascular health. These works include the development of novel wearable devices to continuously monitor BP variations, HRVA and sleep apnea. In the market currently, many wearable devices are available to monitor HRV based on PRV and RRIV. In this research, we are presenting a comparative analysis of different waveform based HRVA measurement techniques against the RRIV based HRVA measurement. The signals

G. M. W. Janjua (✉) · R. Hadia · D. Guldenring · D. D. Finlay
J. A. D. McLaughlin
Ulster University, Shore, Belfast, UK
e-mail: g.janjua@ulster.ac.uk



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_29

used are, PPG, which works on the optical absorption phenomenon and a poor man tonometry, which measures the mechanical pressure variation of blood in artery, a distension waveform. The pulse wave rate variability based HRVA measurement techniques studied during this research are PPG & ballistic pulse pressure (BPP) measurement, a poor man tonometer device. The techniques mentioned here used various fiducial points measured on waveform which are the peak, 1st and 2nd derivative. In this experiment, we have used single measurement methods on peripheral pulse which detects the flow pulsation from the blood volume change in the peripheral arteries optically and mechanically, PPG and BPP respectively.

## Material and Method

For this study, six healthy male volunteers were recruited by age group 34 ± 9 years. The subjects were asked to relax for 5 min on an armchair before vital sign data acquisition and informed consent was taken before the study. The Biopac MP36R system was used for simultaneous data acquisition of all signals and subjects were asked to stay calm during the data acquisition to reduce the noise and motion artifacts. The data was sampled at the 2-kHz frequency with a resolution of 24-bit from acquisition unit for data analysis. Three sensors and their placements are: (1) The ECG signal is a gold standard for HRVA and reference for another signal for this comparative study, so to achieve noise-free signals chest position is selected as it provides the most stable location for cardiac activity monitoring. (2) One of the top trends in the medical industry is wristwatch based cardiac activity monitor—due to easy of their used and internet connectivity. To mimic this scenario, we have placed a PPG sensor of 960 nm wavelength on the sensor on subject's wrist position and subjects were asked to refrain from movement to capture noise and motion artifact-free signals. (3) The use of tonometer in the measurement of vascular hemodynamics is very popular in biomedical and clinical research. A poor man tonometer, custom designed device, BPP sensor [6] is placed on a subject palmar artery in index finger to acquire subject's artery distension waveform. The collated dataset was exported in text file format for analysis in MATLAB environment, on a Windows PC (Table 1).

## Fiducial Points Selection

We have selected three key locations for our analysis and they are compared against the RR HRVA of ECG and these fiducial points are described:

1. The peaks of waveform
2. The peaks of 1st derivative of waveform
3. The peak of 2nd derivative of waveform

**Table 1** HRVA dataset during resting heart condition

| BP Start => Stop | Age | H (cm)/W (Kg) => BMI |
|---|---|---|
| 139/98 => 137/96 | 45 | 175/85 => 27.76 |
| 118/77 => 125/77 | 29 | 177/80 => 25.54 |
| 116/71 => 116/73 | 32 | 174/78 => 25.76 |
| 121/71 => 127/72 | 30 | 180/83 => 25.62 |
| 119/66 => 115/73 | 27 | 188/73 => 20.65 |
| 117/67 => 122/68 | 29 | 179/84 => 26.22 |

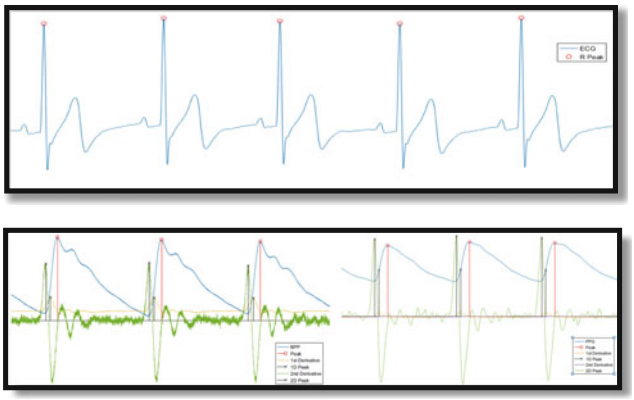


**Fig. 1** Waveforms and fiducial points

**Table 2** Parameter used for analysis

| Time domain | | | | | |
|---|---|---|---|---|---|
| SDNN | pNNx | RMSSD | sdHR | HRVTi | TINN |
| (ms) | (%) | (ms) | (bpm) | (ms) | (ms) |
| Freq domain: Welch periodogram | | | | | |
| Freq domain: auto-regressive periodogram | | | | | |
| Freq domain: Lomb-Scargle periodogram | | | | | |
| pVLF | pLF | pHF | VLF | LF | HF |
| (%) | | | Peak (Hz) | | |
| Nonlinear | | | | Poincare | |
| SampEn | alpha | alpha1 | alpha2 | SD1 | SD2 |
| | | | | (ms) | |
| Time-freq: auto-regressive | | | | | |
| Time-freq: Lomb Periodogram | | | | | |
| Time-freq: wavelet transform | | | | | |
| pVLF | pLF | pHF | VLF | LF | HF |
| (%) | | | Peak(Hz) | | |

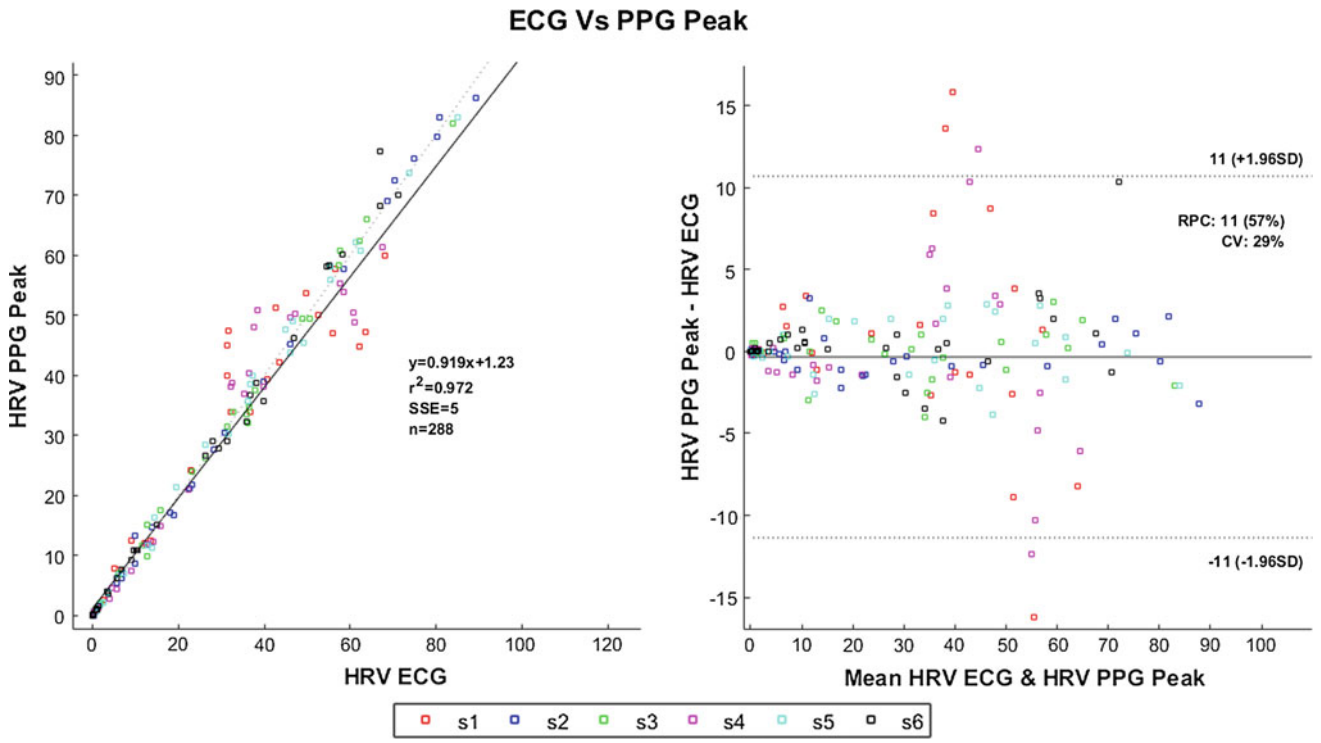



**Fig. 2** R-fit (left) and BA (right) for ECG RR versus PPG peak for HRVA

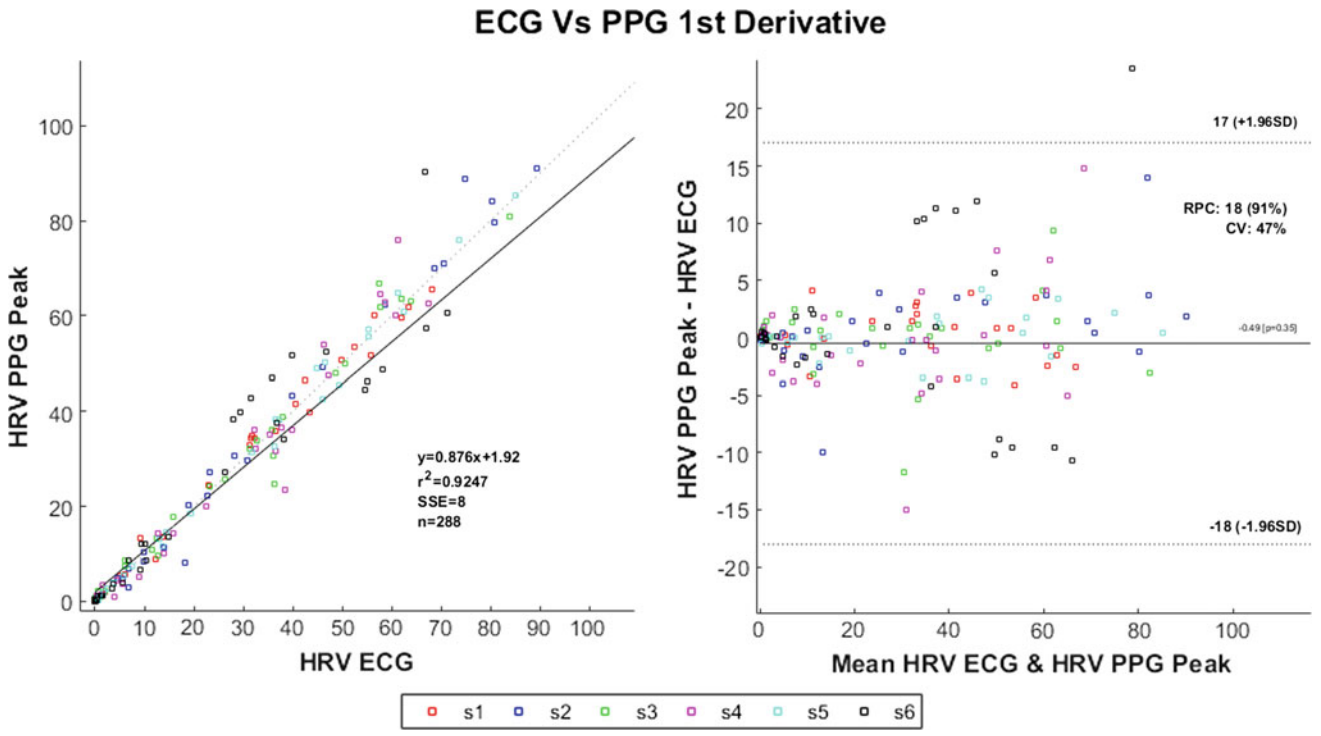



**Fig. 3** R-fit (left) and BA (right) for ECG RR versus PPG 1st derivative for HRVA

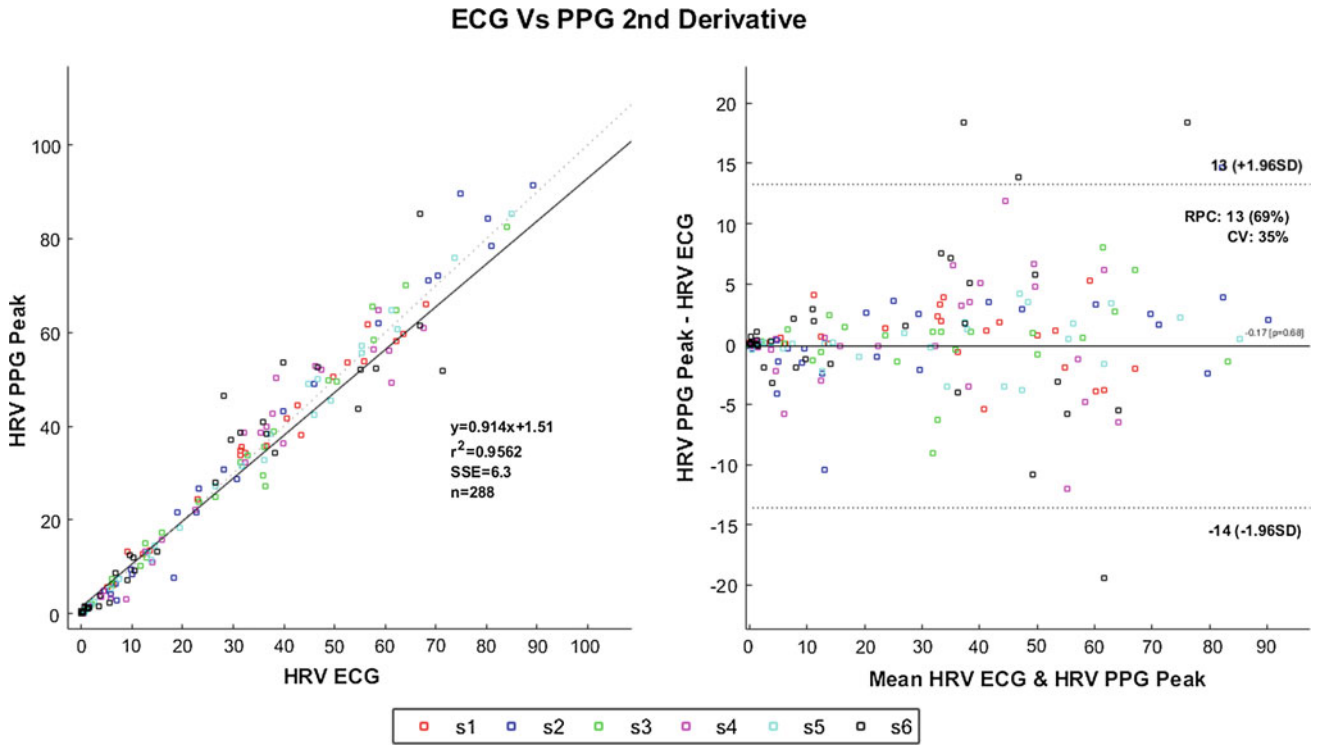



**Fig. 4** R-fit (left) and BA (right) for ECG RR versus PPG 2nd derivative for HRVA

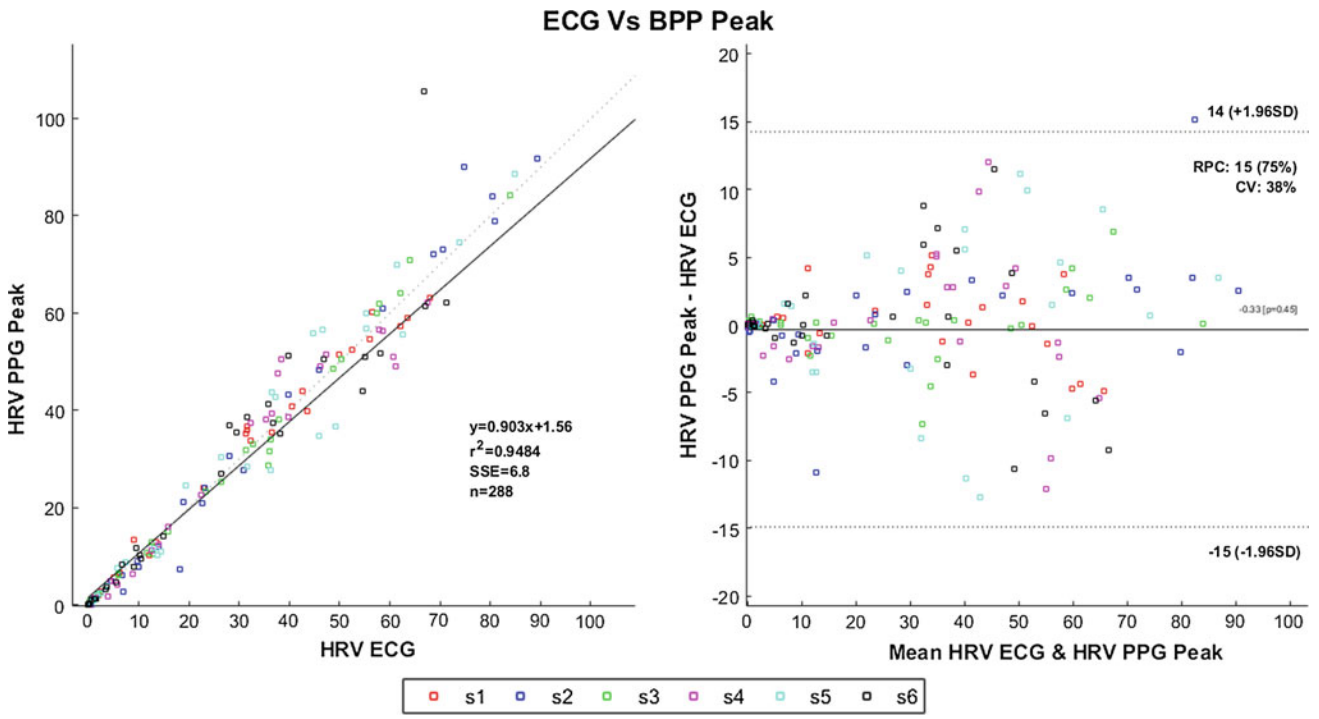



**Fig. 5** R-fit (left) and BA (right) for ECG RR versus BPP peak for HRVA

The significance of these markers has been reported in the literature [6]. During the blood flow in the arterial tree, some of these points are greatly affected by wave reflection due to bifurcation and peripheral resistance. One of the purposes of this study assesses the effect of wave reflection and to highlight the importance of wave reflection; for example, the

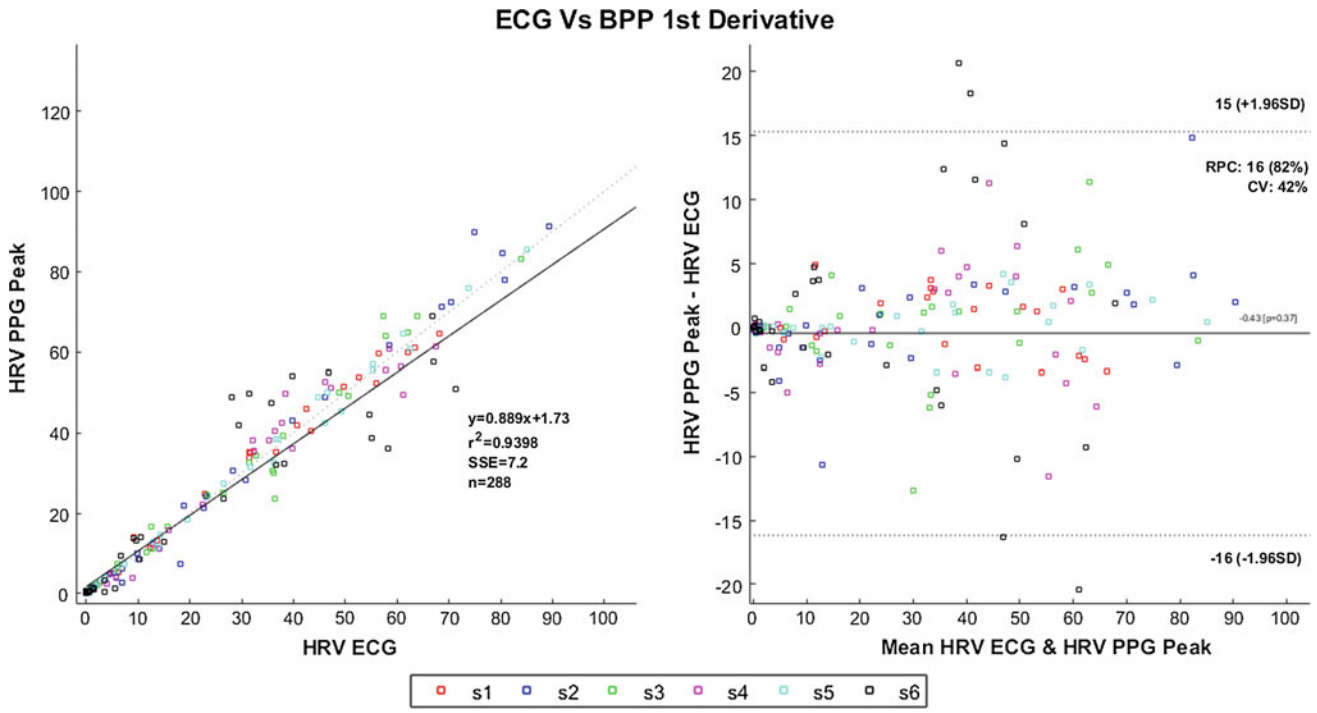



**Fig. 6** Bland Altman for ECG RR versus BPP 1st derivative for HRVA

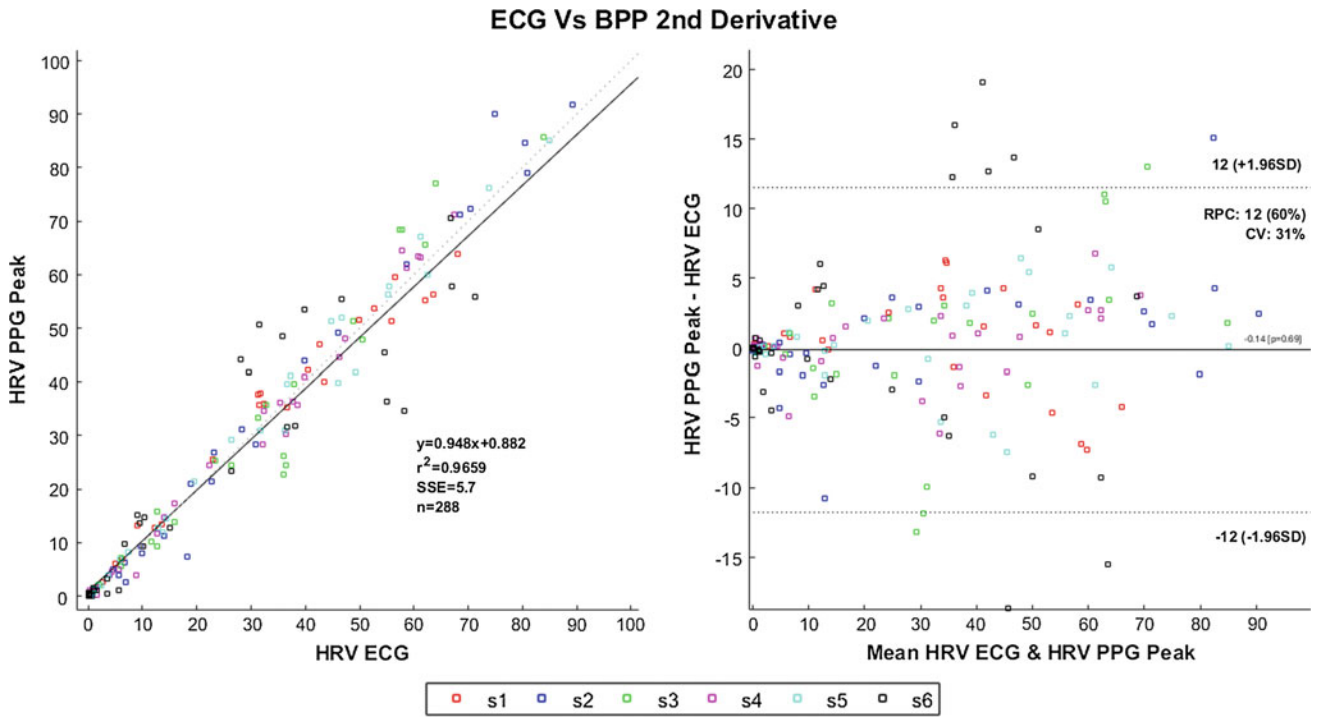



**Fig. 7** R-fit (left) and BA (right) for ECG RR versus BPP 2nd derivative for HRVA

peak of the waveform is greatly affected by a reflected wave in hypertensive subjects [6]. Whereas, waveforms are reflection free in early systole and 1st and 2nd derivatives are the critical points in these aspects. The foot of the waveform is also least affected by wave reflection but it can be impaired by technological limitation like sink color, hairs, and contact to the body in wearable sensors. The intersecting tangent has proved to be a crucial point due to

**Table 3** R-square fit and BA results

| Marker | PPG_P | PPG_1D | PPG_2D | BPP_P | BPP_1D | BPP_2D |
|---|---|---|---|---|---|---|
| $r^2$ | 0.972 | 0.925 | 0.956 | 0.948 | 0.939 | 0.966 |
| RPC/CV | 11/29 | 18/47 | 13/35 | 15/38 | 16/42 | 12/31 |

reproducibility and resilience to artifacts but doesn't have physiological interpretation. The selected location of fiducial points is shown along with waveforms in Fig. 1.

## HRV Algorithm

The HRV is compared across the following aspect:

1. **Time domain**
   a. Standard deviation of the normal-to-normal (NN) interval series (SDNN)
   b. Root mean square of successive differences (RMSSD)
   c. Triangular interpolation of the NN interval (TINN)
2. **Frequency domain**
   a. Power spectrum density (PSD)
      i. VLF, LF, and HF frequency 0.003–0.04 Hz, 0.04–0.15 Hz, and 0.15–0.4 Hz respectively
3. **Time-frequency domain**
   a. Windowed Periodogram
   b. Continuous Wavelet Transform
   c. Discrete Wavelet Transform
4. **Nonlinear domain**
   a. Poincare plot analysis
   b. Sample Entropy
   c. Detrended Fluctuation Analysis

The open source code [7] has been used due to ease of customization and to test calculate HRVA on the above-mentioned domain. The data is preprocessed by fourth order butter worth filter to remove noise from the signals and fiducial points were taken by taking the peaks of waveform and peaks of 1st and 2nd derivative waveforms. The R-peak on ECG signal is detected using Pans-Tompkins algorithm and data was visually inspected for cardiac arrhythmia. All fiducial points were detected and beat to beat interval is calculated for HRVA.

## Results and Discussion

The cardiac chronotropic modulation by autonomic nerves system facilitates end-organ perfusion in coordination to changing blood hemodynamics in arteries, to accommodate the sudden changes in the human circulation system. HRVA is used by biomedical researchers and clinicians to monitor cardiac disease like heart failure, respiratory sinus arrhythmia (RSA), hyperaldosteronism etc. HRVA publicly available resources are Kubios [7], ECGLab [8], KARDIA [9], and PhysioToolkit [10] from which we have selected the HRVAS [6] due to its easy customization. In this study, sensors are compared in the context of e-Health applications as an alternative to chest based ECG systems. The HRVA is conducted on i7 PC in MATLAB environment. The ECG, PPG and BPP waveforms and their fiducial locations based beat to beat interval are shown as below: For agreement analysis, Bland Altman (BA) plot, and regression fit (R-fit) plotted is on following key parameters in respective domains, shown in Table 2. (pNNx = percentage of successive normal cardiac inter beat intervals, sdHR = standard deviation Heart rate, HRVTi = HRV time index, p = peak, VL = very low, L = low, H = high, F = frequency, SampEn = Sample Entropy, alpha, alpha1 & alpha2 = Short term scaling exponent, SD1 = dispersion standard deviation perpendicular to the axis, SD2 = dispersion standard deviation along the axis).

The RR interval based HRVA of all subjects is tested for an agreement for HRVA for all fiducial points. Plots of Bland Altman analysis are shown below: (Figs. 2, 3, 4, 5, 6 and 7).

We have found correlation information which shows that the overall highest R-square parametric fit is found in PPG peak and BPP 2nd Derivative compared to ECG RR HRVA and lowest found in PPG 1st derivative vice versa for the sum of square error. Also, we have found, BA information, that reproducibility and variation coefficient (RPC & CV) is lowest in PPG Peak and BPP 2nd Derivative, results shown in Table 3.

## Conclusions

In this comparative study, we have found that BPP 2nd derivative and PPG peak can be used as a surrogate for ECG RR HRVA. Also, BPP, novel sensor, showed competitive effectiveness to the optical sensor for HRVA.

**Acknowledgements** This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 676201.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology (1996) Heart rate variability. Circulation 93(5):1043–1065
2. Huikuri HV, Makikallio T, Airaksinen KEJ, Mitrani R, Castellanos A, Myerburg RJ (1999) Measurement of heart rate variability: a clinical tool or a research toy? J Am Coll Cardiol 34 (7):1878–1883
3. Politano L, Palladino A, Nigro G, Scutifero M, Cozza V (2008) Acta Myol 27(3):114–122
4. Jeyhani V, Mahdiani S, Peltokangas M, Vehkaoja A (2015) Comparison of HRV parameters derived from photo plethysmography and electrocardiography signals. In: Proceedings of annual international conference on IEEE Engineering in Medicine and Biology Society (EMBS), pp. 5952–5955, Nov 2015
5. Nenova BD, Iliev IT (2009) Non-invasive methods of peripheral pulse detection: advantages and disadvantages. Annu J Electron 57–60
6. Janjua G, Mclaughlin J, Finlay D, Guldenring D (2017) Wireless chest wearable vital sign monitoring platform. In: 39th Annual international conference of the IEEE, EMBC
7. de Carvalho JLA et al (2002) Development of a matlab software for analysis of heart rate variability. In: 6th international conference on signal processing, vol. 2, pp. 1488–1491
8. Perakakis P et al (2010) KARDIA: a matlab software for the analysis of cardiac interbeat intervals. Comput Methods Programs Biomed 98:83–89
9. Goldberger AL et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101:215–220
10. Ramshur JT (2010) Design, evaluation, and application of heart rate variability analysis software

# Wavelet ECG Analysis in Time-Frequency Domain of the QRS-Complex in Individuals with Left Bundle Branch Block

Kalliopi Papathoma, Stavros Chatzimiltiadis, Nikolaos Maglaveras, Ioanna Chouvarda, Efstratios Theofilogiannakos, Dimitrios Konstantinou, and Vassilios Vassilikos

**Abstract**

LBBB in heart failure patients is a negative predictor for survival. This pattern is also recorded in individuals without significant structural heart diseases. The LBBB morphology has not been previously analysed using wavelet analysis in *time-frequency* domain in order to identify markers which distinguish these patients. The purpose of this analysis is to investigate if there are any differences in LBBB morphology between normal individuals with LBBB and patients with heart failure and LBBB. Signal-averaged electrocardiograms were recorded, in orthogonal leads and QRS decomposition in nine (9) time-frequency segments was performed using the 'cmor' wavelet transformation. Seventy (70) patients (mean age 66, 47 male) were studied. The mean and maximum energies of the QRS complexes were calculated in each of the 9 time-frequency segment. Wavelet parameters of the QRS complex in all segments were higher for normal individuals without LBBB. In the Z lead, the differences in mean QRS energies are significant between the two groups especially in the high frequency band (150–200 Hz) in all time segment whereas the mean energy in X lead presents significant difference in median frequency band (100–150 Hz) in the first time segment. In conclusion wavelet transformation of the QRS complex could differentiate normal individuals from heart failure patients with LBBB.

**Keywords**

Left bundle branch block • Signal processing • Cmor • Wavelet transform • QRS wavelet analysis

## Introduction

Left bundle branch block occurs in up to 30% of patients with heart failure [1]. The effect of LBBB on cardiac function leading to left ventricular dyssynchronization is a well established poor prognostic factor in HF patients [2]. However, LBBB can also be seen in asymptomatic patients with a structurally normal heart. Our objective in this study is to apply time-scale analysis of the QRS signal to identify QRS elements that could predict the prognosis of LBBB people without heart failure based on the cmor wavelet transformation [3].

K. Papathoma (✉) · N. Maglaveras · I. Chouvarda
Laboratory of Medical Informatics, Medical School, Aristotle University, 1,St.Kyriakidi, Thessaloniki, Greece
e-mail: kpapathom@gmail.com

S. Chatzimiltiadis · D. Konstantinou
First Department of Cardiology, AHEPA University Hospital, Aristotle University Medical School, Thessaloniki, Greece

E. Theofilogiannakos · V. Vassilikos
Third Department of Cardiology, HIPOKRATIO University Hospital, Aristotle University Medical School, Thessaloniki, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_30

**Table 1** Demographic characteristics of individuals with and without a form of heart failure

| | Total n = 70 | Group A n = 12 | Group B n = 58 | P value |
|---|---|---|---|---|
| Age (years) | 66<br>(42–87) | 69<br>(58–87) | 65<br>(42–82) | 0.586 |
| Male gender | 47 (67.14%) | 0<br>(0%) | 47<br>(81.03%) | *0.000* |
| NYHA_class, median (range) | 3<br>(1–4) | 1<br>(1–3) | 3<br>(1–4) | *0.000* |
| NYHA II (%) | 8<br>(11.43%) | 3<br>(25%) | 5<br>(8.62%) | |
| NYHA III (%) | 52 (74.29%) | 2<br>(16.67%) | 50<br>(86.21%) | |
| NYHA IV (%) | 2<br>(2.86%) | 0 | 2<br>(3.45%) | |
| QRS_duration (ms) | 163<br>(137–192) | 154<br>(137–163) | 168<br>(145–192) | *0.002* |
| Ejection_fraction | 30<br>(14–64) | 58<br>(53–64) | 25<br>(14–44) | 0.190 |

NYHA, New York Heart Association. P-values in bold indicate significance at 5% or lower. Values are rounded to the closest integer

## Materials and Methods

A. *Study population*

Data of seventy (70) patients (mean age 66, 47 male) have been analyzed (Table 1). Group A consisted of twelve (12) people with LBBB without heart failure and Group B consisted of fifty-eight (58) patients with LBBB and heart failure. QRS duration (QRS > 130 ms for women and QRS > 140 ms for men) were used for this study [4].

B. *ECG recordings*

ECGs recorded with a 3-channel digital recorder (Galix Biomedical Instrumentation) in 3 axes (X, Y, Z) for 10 min at rest. The sampling frequency was 1000 Hz, and the duration of each signal was 90 s in order to ensure as much as possible artefact-free signal of same length for all patients.

C. *ECG signal processing*

A signal averaging algorithm was developed to enhance the QRS signal and reduce the noise level. The following steps were taken:

(1) Baseline wander removal: amplitude normalization, baseline correction and application of a denoising wavelet filter.
(2) Frank leads: three orthogonal leads were constructed from the 12-leads.
(3) QRS template: all QRS complex in every signal were considered and manual inspection of all QRS complexes by experts was pre-followed.
(4) Wavelet transform: wavelet transform of the QRS complex was performed using 'cmor' wavelet analysis. QRS complex was divided into 3 equal parts in time domain and 3 frequency bands (low frequency range: 50–100 Hz, medium frequency range: 100–150 Hz and high frequency range: 150–200 Hz).
(5) Calculating features: the mean and maximum wavelet energies of the QRS complexes were calculated in each of the 9 time-frequency band [5].

## Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics version 23.0 (Chicago, IL, USA). We use mean and range of minimum and maximum values to express normally distributed variables, and median and interquartile range for non-normally distributed variables (i.e., all wavelet parameters). Differences between groups were explored with Student's t–test or Mann–Whitney U test for normally and non–normally distributed variables, respectively. Univariate logistic regression analysis was performed in order to identify the contribution of each variable in the prediction of response to CRT. We also engaged a stepwise fashion multivariate logistic regression model to empirically test our inferences. Specifically, we included QRS duration (the only clinical variable that was found significant in the univariate analysis) and each wavelet parameter separately, due to

**Table 2** Baseline wavelet parameters between individuals with and without a form of heart failure

| | | Total n = 70 | Group A n = 12 | Group B n = 58 | P Value |
|---|---|---|---|---|---|
| X | mean_a4X | 9.1 (6.1) | 14 (12.2) | 8.7 (4.6) | *0.038* |
| Z | max_a2Z | 61.4 (42.8) | 91.4 (108.8) | 60.4 (40.3) | *0.040* |
| | mean_a1Z | 56.5 (60) | 81.3 (95.7) | 51.2 (51.4) | *0.029* |
| | mean_a2Z | 38.9 (39.6) | 70.6 (86) | 37.9 (27.3) | *0.035* |
| | mean_a3Z | 25.5 (16.2) | 38 (19.8) | 24.5 (14.1) | *0.046* |
| | mean_a4Z | 16.7 (18.2) | 28.9 (33.4) | 16 (15) | *0.046* |

P-values in bold indicate significance at 5% or lower. Values are rounded to the closest integer

collinearity between wavelet parameters. We considered as statistically significant all p values less than 5% [6].

## Descriptive Statistics

Table 1 provides the baseline characteristics of the entire sample (70 patients, mean age 66 and 47 male), of individuals without a heart failure (12 patients) and individuals with a form of heart failure (58 patients). The baseline mean QRS duration was 163 (range 137–192 ms) and mean Ejection Fraction was 30%. The differences in means of QRS duration and NYHA class are statistically significant ($p < 0.005$) between the two groups, as expected. Another statistically significant difference was that of the gender of participants, as all individuals without a heart failure were female. Finally, almost all individuals with a form of heart failure belong to NYHA class III, while approximately 12% were distributed between NYHA class II (5 patients) and NYHA class IV (2 patient). Wavelet parameters in each of three orthogonal leads in responders and non-responders are presented in Table 2. In both groups the biggest proportion of the QRS energy was distributed in the high frequency band (150–200 Hz) in lead Z and medium frequency band (100–150 Hz) in lead X. All wavelet energies of the QRS complex in all frequency bands in lead Z and the lead X were higher for individuals without heart failure. Finally, no statistically significant differences were observed in lead Y.

Terminology used for wavelet parameters: mean or max {frequency band-time}, e.g. mean{m-t1} represents the mean energy of the QRS complex recorded in the medium frequency band (100–150 Hz) in the first time segment, mean{l-t3} represents the mean energy of the QRS complex recorded in the low frequency band (50–100 Hz) in the third time segment and mean{h-t2} represents the mean energy of the QRS complex recorded in the high frequency band (150–200 Hz) in the second time segment (Fig. 1).

## Results

Wavelet parameters of the QRS complex in all frequency bands in all leads were *higher* for people with LBBB without heart failure. No significant differences were noted in the Y lead. In the Z lead, the differences in mean QRS energies are significant between the two groups in the high and medium frequency band (100–200 Hz) in all time segment whereas in the lead X the mean energy is significant in medium frequency band (100–150 Hz) in the first time segment.

## Discussion

LBBB is a common ECG finding in patients with heart failure. The presence of LBBB is often an indicator of primary structural or ischemic heart disease. LBBB seems to be associated with an increased risk of cardiovascular mortality. Moreover, patients with heart failure and LBBB have increased in-hospital mortality and higher disease severity in comparison with heart failure patients without left bundle branch block [7]. However, LBBB can also be seen in asymptomatic patients with a structurally normal heart. There is evidence that LBBB is associated with the development of progressive LV systolic dysfunction in the absence of coronary artery disease or other identifiable etiologies. Moreover, it is already known the existence of the LBBB-associated cardiomyopathy. In this case of LBBB-associated cardiomyopathy, the isolated LBBB causes cardiac remodeling due to mechanical desynchronize that can be reversible by biventricular pacing [8].

We hypothesized that heart failure patients with LBBB have specific features from the individuals without heart failure. Our results confirm that there are signal components inside the QRS complex (energies) which could diversify these groups. Yet, we do not have a clear link between these components and further investigations are required.

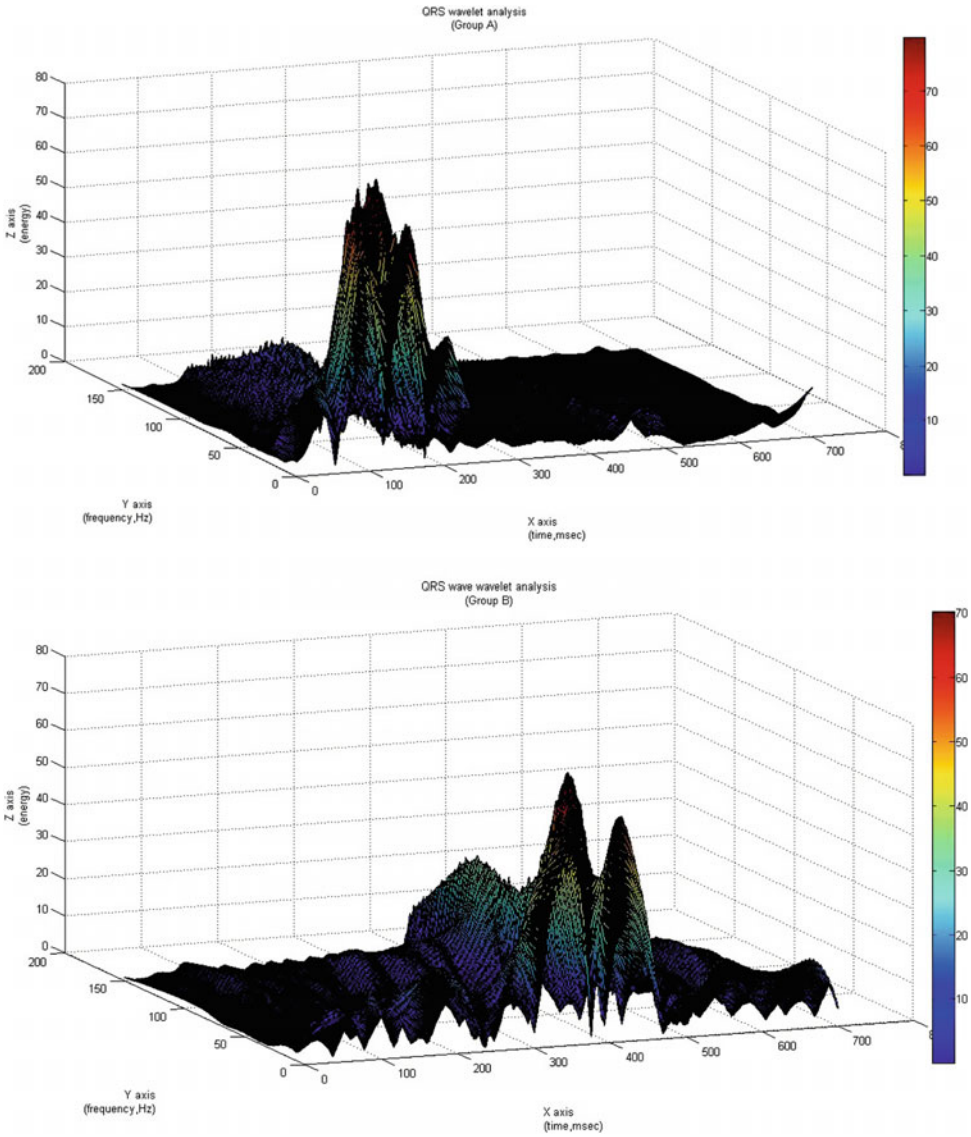



**Fig. 1** Representative example of QRS wave wavelet transformation at Z orthogonal lead for group A and B (top and bottom respectively). Time (QRS wave duration, msec) is shown at X axis, frequency (Hz) at Y axis and QRS wave energy values (mV2) at Z axis

However, in the future, wavelet transformation methods could be implemented as a marker for the detection of these normal individuals with LBBB that may eventually develop LBBB-associated cardiomyopathy.

## Conclusion

In conclusion wavelet transformation of the QRS complex could differentiate normal individuals from heart failure patients with LBBB. Our preliminary results indicate that time-scale analysis of the QRS complex contains information on predicting individuals diagnose.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Shamim W, Francis DP, Yousufuddin M, Varney S, Pieopli MF, Anker SD, Coats AJ (1999) Intraventricular conduction delay: a prognostic marker in chronic heart failure. Int J Cardiol 70(2):171–178.
2. William T, Abraham MD, Westby G, Fisher MD, Andrew L, Smith MD, David B, Delurgio MD, Angel R, Leon MD, Evan Loh MD, Dusan Z, Kocovic MD, Milton Packer MD, Alfredo L, Clavell MD, David L, Hayes MD, Myrvin Ellestad MD, Robin J, Trupp MSN, Jackie Underwood BSN, Faith Pickering BSN, Cindy Truex BSN, Peggy McAtee MSN, John Messenger MD (2002) Cardiac resynchronization in chronic heart failure for the MIRACLE study group. N Engl J Med 346:1845–1853
3. X Xia, JP Couderc, S McNitt, W Zareba (2010) Predicting effectiveness of cardiac resynchronization therapy basen on QRS decomposition using the Meyer orthogonal wavelet transformation. Heart Research Follow-up Program, Cardiology Unit, Computing in cardiology, vol 37. University of Rochester, Rochester (NY), USA, pp 983–986

4. Strauss DG, Selvester RH, Wagner GS (2011) Defining left bundle branch block in the era of cardiac resynchronization therapy. Am J Cardiol 107(6):927–934
5. Vassilikos V et al (2011) Novel non-invasive P wave analysis for the prediction of paroxysmal atrial fibrillation recurrences in patients without structural heart disease. Int J Cardiol 165–172
6. Vassilikos V et al (2014) QRS analysis using wavelet transformation for the prediction of response to cardiac resynchronization therapy. J Electrocardiol 59–65
7. Bouqata N, Kheyi J, Miftah F, Sabor H, Bouziane A, Bouzelmat H, Chaib A, Benyass A, Moustaghfir A (2015) Epidemiological and evolutionary characteristics of heart failure in patients with left bundle branch block—a moroccan center-based study. J Saudi Heart Assoc 27(1):1–9
8. Vaillant C, Martins RP, Donal E, Leclercq C, Thébault C, Behar N, Mabo P, Daubert JC (2013) Resolution of left bundle branch block-induced cardiomyopathy by cardiac resynchronization therapy. J Am Coll Cardiol 61(10):1089–1095

# Adaboost Classifier with Dimensionality Reduction Techniques for Epilepsy Classification from EEG

S. K. Prabhakar and H. Rajaguru

**Abstract**

Epilepsy is a serious neurological disorder affecting the human community and this problem has to be dealt with utmost importance. In this disorder, the activity of the neurons in the human brain becomes abnormal and it is witnessed by recurrent seizures. As it is the second most commonly occurring neurological disorder, next to stroke, it affects the quality of life to a great extent. For the clinical evaluation of the activities of the brain, the most commonly used instrument is Electroencephalography (EEG). For dealing with various disorders like classification of epileptic seizures, assessment of mental fatigueness and coma, sleep disorders and schizophrenia, EEG is widely used. As the recordings of the epileptic EEG signal have a very long duration, a lot of data is generated by it. In this paper, the dimensionality of the data is initially reduced with the help of two dimensionality reduction techniques such as Hilbert Transform and Hessian Local Linear Embedding (HLLE). The dimensionally reduced values are then classified with the help of Adaboost Classifier for epilepsy classification from EEG signals. Results show that when Hilbert Transform is classified with Adaboost, a classification accuracy of 93.92% is obtained. When HLLE is classified with Adaboost, a classification accuracy of 92.83% is obtained.

**Keywords**

Epilepsy • EEG • Hilbert • HLLE

## Introduction

Epilepsy is considered as a common and chronic neurological disorder affecting millions of people across the globe [1]. This disorder is witnessed by recurrent seizures and it is highly unpredictable in nature. The characteristic features of the human brain with respect to epilepsy are detected by a famous non invasive technique called EEG [2]. In the cerebral cortex of the human brain, due to the synaptic excitation of the dendrites of various neurons, the current flows and the measurement of it is EEG. When the neurons are activated, within the dendrites the synaptic currents are produced which generates a secondary electrical field over the scalp of the human brain. By placing the multiple electrode EEG machines either over the cortex or from inside the brain, the recordings can be done in various formats. The most significant and traditional way of analysis of epileptic EEG signals is by means of visual inspection by an expert in clinical sciences [3]. As it is prone to a longer duration with certain other drawbacks, automated seizure detection and classification techniques has to be developed. A lot of works has been published in the past few decades with respect to epilepsy classification from EEG signals.

For the long term intracranial EEG recordings, the seizures were detected automatically by using wavelet transforms and Support Vector Machines (SVM) by Liu et al. [4]. Probabilistic Mixture Models (PMM) was used for the classification of epilepsy by Prabhakar and Rajaguru [5].

S. K. Prabhakar (✉) · H. Rajaguru
Department of ECE, Bannari Amman Institute of Technology, Alathukombai Post, Sathyamangalam, India
e-mail: sunilprabhakar22@gmail.com



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_31

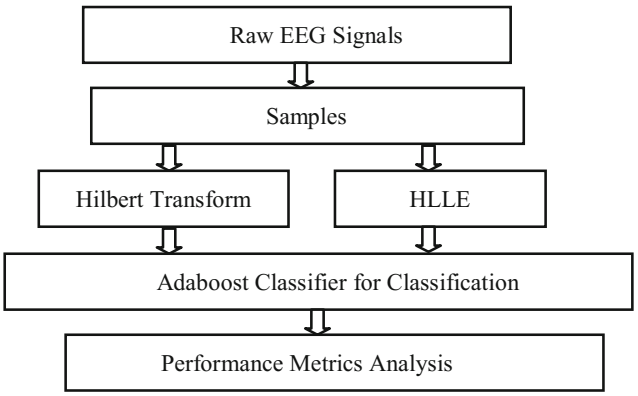



**Fig. 1** Block diagram of the work

A hierarchical methodology of EEG classification for epileptic detection based on the concept of wavelet packet entropy feature extraction was done by Wang et al. [6]. For detection of epilepsy, a neural network approach was developed by Nigam and Graupe [7]. The time-frequency analysis and Artificial Neural Network (ANN) was utilized by Tzallas et al. for automatic seizure detection [8]. A patient remote monitoring system was developed by Prabhakar and Rajaguru for epilepsy classification from EEG signals [9]. In this paper, the concept of Hilbert Transform and HLLE are utilized as dimensionality reduction techniques and then it is classified with Adaboost Classifier. The Fig. 1 shows the block diagram of the work.

In this paper, the materials and methods are discussed in Sect. 2 along with the usage of Hilbert Transform and HLLE as dimensionality reduction techniques. The classification using Adaboost Classifier is given in Sect. 3 followed by results and discussion in Sect. 4. The paper is concluded in Sect. 5.

## Materials and Methods

The cleansing of the electrodes is done initially before placing the electrodes on the scalp of the epileptic patients. In our case, twenty epileptic patients were considered and the readings were obtained from them by placing the 16 channel electrodes on the scalp of these patients according to the standard 10–20 International system. These patients were admitted at Sri Ramakrishna Hospital, Coimbatore, India, for their epilepsy treatment and upon explain the necessity of research to them and with the consent of both the doctors and patients, the EEG data was collected in European Data Format (EDF). The EEG data was split into three epochs for all the 16 channels of every epileptic patient. Each epoch had around 400 values and so the total number of values for all the channels in 20 patients was too high to process. Here the values represent the instantaneous amplitude level of the EEG signals. As there is a huge amount of data present here, dimensionality reduction is absolutely necessary and so Hilbert Transform and HLLE is used.

### A. *Hilbert Transform:*

In the context of signal processing, Hilbert Transform has played an important role always. Especially, if a particular sequence is causal in nature, then the real and imaginary part of its Hilbert Transform relation has an enormous role to play in it [10]. From the analytic function properties, the development of Hilbert Transform relations can be obtained formally. It is developed on the view that the real and imaginary sections of Z-transform of a particular causal sequence are the transforms of even and odd component of the sequence. Based on the arbitrary continuous signal, Hilbert transform is heavily dependent on it and majority of the signals are time limited and can be digitally sampled. For analyzing the frequency content of the signal with respect to time, Hilbert Transform is used widely.

The Hilbert Transform $q(t) = H\{w(t)\}$ is expressed as

$$q(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{w(\tau)}{\tau - t} d\tau$$

The complex analytic signal represented as $h(t)$ is defined as

$$h(t) = w(t) + jq(t)$$
$$h(t) = B(t).e^{j\psi(t)}$$

where $q(t)$ is the Hilbert Transform of $w(t)$. The envelope of $h(t)$ is $B(t)$ and the instantaneous phase of $h(t)$ denoted as $\psi(t)$ is expressed as

$$\psi(t) = \tan^{-1} \frac{q(t)}{w(t)}$$

The power and instantaneous frequency of the signal can be determined easily with the help of Hilbert Transform.

### B. *Hessian Local Linear Embedding:*

As a variant of LLE technique, it is widely used. The conceptual framework is observed as a modification of Laplacian Eigen maps framework [11]. In such modifications, Laplacian is replaced by Hessian in the quadratic form. If the Laplacian is replaced by a Hessian, then a global embedding is found out which is linear in all the sets of local tangent coordinates. A dataset $Y$ is considered which consists of $A$ data points $y_1, y_2, \ldots, y_A$ in a high dimensional space $\Re^D$. It is hypothesized that they lie on a manifold which is smooth $A \subset \Re^D$ which has an intrinsic dimensionality $d < D$.

Assume $Y = [y_1, y_2, .., y_A] \in \Re^{D\times M}$ and $Z = [z_1, z_2, \ldots, z_A] \in \Re^{d\times M}$ indicate the nature of all the sampled points along with their embedding coordinates. Recovering the corresponding low-dimensional co-ordinates $z_i \in \gamma$ of every $y_i \in A$ is the main objective of this manifold based dimensionality reduction technique. $A$ is the signal of every coordinate space $\gamma \subset \Re^d$ under a smooth mapping criteria $\psi : \gamma \to \Re^D$. The mapping $\psi$ is a local-isometric embedding and its inverse mapping $\phi = \psi^{-1} : A \to \Re^d$ gives the local isometric coordinate. If on the interior of $A$, let $y$ be a fixed point and the tangent space defined at $y$ is $T_y(A)$. Then on the tangent space, orthogonal coordinates are utilized to define the Hessian in order to compute the Hessian of a particular function $f : A \to \Re$. Assuming the tangent coordinate of $y \in N(y)$ is expressed as $v$, where $N(y)$ denotes the local patch built by the point $y$ and its $k$-nearest neighbours. The rule defines a function $h = V \to \Re$, where $V$ is the neighborhood of $v \in \Re^d$ and is found by the tangent coordinate of $N(y)$. The Hessian of $q$ at $y$ in tangent coordinates is expressed as the ordinary Hessian of $h$ and is expressed as

$$\left(H_q^{\tan}\right)_{i,j}(y) = \frac{\partial}{\partial v_i}\frac{\partial}{\partial v_j}h(v)\Big|_{v=\phi(y)}$$

Various tangent Hessians are obtained by various local coordinate systems. All the Hessians share the similar Frobenius norm $\left\|H_q^{\tan}(y)\right\|_Q^2$, so that the quadratic form is well defined as

$$H(q) = \int_A \left\|H_q^{\tan}(y)\right\|_Q^2 dy$$

where the canonical measure which corresponds to the volume form on $A$ is expressed as $dy$. The average curviness of $'q'$ over the manifold $A$ is expressed by the functional $H(q)$.

## Adaboost Classifier for Classification

The dimensionally reduced values are then fed inside the Adaboost Classifier for epilepsy classification from EEG signals. The main idea of the classifier ensemble is to combine the outputs obtained from a number of weak learners and one such famous classifier is Adaboost Classifier [12]. For every dimensionally reduced value, the optimal threshold classification function is determined by the weak learner such that the minimum numbers of samples are misclassified. Therefore a weak classifier $h(c, g, q, \theta)$ comprises of a feature $g$, a threshold $(\theta)$ and a polarity $q$ that mentions the direction of the inequality sign as follows

$$h(c, g, q, \theta) = \begin{cases} 1 & if \quad qf(c) < q\theta \\ 0 & if \quad otherwise \end{cases}$$

For classifier learning, the boosting process and the Adaboost algorithm is as follows:

**First step**: Given the sample signals $(c_1, d_1)\ldots(c_n, d_n)$, where $d_i = 0$ and 1for negative and positive samples.

**Second step**: The weight $v_1$ initialization is done as follows $\begin{matrix} i = 1/2s & for & d_i = 0 \\ i = 1/2l & for & d_i = 1 \end{matrix}$

**Third step**: For $t = 1, ..T$

(i) The normalization of the weights are done as $v_{t,i} = \frac{v_{t,j}}{\sum_{j=1}^{n} v_{t,j}}$, where the probability distribution is known as $v_t$

(ii) With respect to the weighted error, the best weak classifier is selected as

$$\varepsilon_t = \min_{g,q,\theta} \sum v_i |h(c_i, g, q, \theta) - d_i|$$

(iii) $h_t(c) = h(c, g_t, q_t, \theta_t)$ is defined where $g_t$, $q_t$ and $\theta_t$ are the minimizers of $\varepsilon_t$.

(iv) The weights are updated as $v_{t+1,i} = v_{t,i}\beta_t^{1-j_i}$, where $j_i = 0$ if example $c_i$ is properly classified otherwise, $j_i = 1$ and $\beta_t = j_t/1 - j_t$

**Fourth step**: The final strong classifier is expressed as

$$Z(c) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(c) \geq 1/2 \sum_{t=1}^{T} \alpha_t \\ 0 & otherwise \end{cases}$$

where, $\alpha_t = \log 1/\beta_t$.

Thus using the Adaboost algorithm, the strong classifier can be trained. The best performing weak classifier is chosen from a collection of weak classifiers over numerous sequences which act on a single sample of dimensionally reduced values.

## Results and Discussion

When Hilbert Transform and HLLE are employed to reduce the dimensions of the EEG data and when it is classified with the Adaboost Classifier, parameters like Performance Index, Accuracy, Specificity, Sensitivity, Time Delay and Quality Values are evaluated and the average results are computed in Tables 1 and 2. The mathematical formulae for the Performance Index (PI), Sensitivity, Specificity and Accuracy are given as follows

$$PI = \left(\frac{PC - MC - FA}{PC}\right) \times 100$$

**Table 1** Hilbert transform with adaboost classifier

| Name | Epoch 1 | Epoch 2 | Epoch 3 | Average |
|---|---|---|---|---|
| PC (%) | 87.24 | 86.30 | 90.00 | 87.85 |
| MC (%) | 4.53 | 3.28 | 1.24 | 3.02 |
| FA (%) | 8.21 | 10.37 | 8.72 | 9.10 |
| PI (%) | 84.34 | 83.54 | 88.60 | 85.49 |
| Sensitivity (%) | 91.77 | 89.58 | 91.25 | 90.87 |
| Specificity (%) | 95.46 | 96.71 | 98.75 | 96.97 |
| Time delay(s) | 2.01 | 1.92 | 1.87 | 1.93 |
| Quality values | 19.04 | 18.40 | 19.61 | 19.02 |
| Accuracy (%) | 93.62 | 93.15 | 95.00 | 93.92 |

**Table 2** HLLE with adaboost classifier

| Name | Epoch 1 | Epoch 2 | Epoch 3 | Average |
|---|---|---|---|---|
| PC (%) | 86.04 | 82.76 | 88.23 | 85.67 |
| MC (%) | 7.55 | 7.70 | 4.53 | 6.59 |
| FA (%) | 6.39 | 9.52 | 7.21 | 7.71 |
| PI (%) | 82.78 | 77.24 | 86.05 | 82.02 |
| Sensitivity (%) | 93.59 | 90.46 | 92.76 | 92.27 |
| Specificity (%) | 92.44 | 92.29 | 95.46 | 93.40 |
| Time delay (s) | 2.17 | 2.11 | 2.03 | 2.10 |
| Quality values | 18.89 | 18.21 | 19.48 | 18.86 |
| Accuracy (%) | 93.02 | 91.38 | 94.11 | 92.83 |

where PC = Perfect Classification, MC = Missed Classification and FA = False Alarm. The Sensitivity, Specificity and Accuracy measures are mathematically formulated by the following

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2}$$

The Quality Value $Q_V$ is mathematically expressed as

$$Q_v = \frac{C}{(R_{fa} + 0.2) * (T_{dly} * P_{dct} + 6 * P_{msd})}$$

where C denotes the scaling constant, Rfa expresses the number of false alarm per set, $T_{dly}$ indicates the average delay of the onset classification in s, $P_{dct}$ mentions the percentage of perfect classification and $P_{msd}$ specifies the percentage of perfect risk level missed.

The time delay is mathematically given as follows

$$Time \quad Delay = \left[2 \times \frac{PC}{100} + 6 \times \frac{MC}{100}\right]$$

If Hilbert Transform is used as dimensionality reduction technique and when it is classified with Adaboost Classifier, an average classification accuracy of 93.92% is obtained. An average quality value of 19.02 along with an average time delay of about 1.93 s is also obtained.

If HLLE is utilized as dimensionality reduction technique and when it is classified with Adaboost Classifier, an average classification accuracy of 92.83% is obtained. An average quality value of about 18.86 along with an average time delay of 2.10 s is also obtained.

## Conclusion

As epilepsy is recognized as a chronic and serious neurological disorder, a very good research on this disorder is the need of the hour. In this paper, the dimensionality of the epileptic EEG signals was reduced with the help of Hilbert

Transform and HLLE and later it was classified with the Adaboost Classifier. The results show that an average classification accuracy with the help of Adaboost Classifier obtained is 93.92% when Hilbert Transform is employed and 92.83% when HLLE is employed. Future works aim to utilize various other versatile post classification schemes for epilepsy classification.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Liang S-F, Wang H-C, Chang W-L (2010) Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection. Eura J Advanc Signal Process 2010, Article ID 853434
2. Liu Y, Zhou W, Yuan Q, Chen S (2012) Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. IEEE Trans Neural Syst Rehabil Eng 20 (6):749–755
3. Srinivasan V, Eswaran C, Sriraam N (2007) Approximate entropy-based epileptic EEG detection using artificial neural networks. IEEE Trans Inf Technol Biomed 11(3):288–295
4. Liu Y, Zhou W, Yuan Q, Chen S (2012) Automatic Seizure Detection Using Wavelet Transform and SVM in Long-Term Intracranial EEG. IEEE Trans Neural Syst Rehabil Eng 20:749–755
5. Prabhakar SK, Rajaguru H (2017) Conceptual analysis of epilepsy classification using probabilistic mixture models. In: 5th IEEE winter international conference on brain-computer interface, January 9–11, South Korea
6. Wang D, Miao D, Xie C (2011) Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, Expert Syst. Appl 38:14314–14320
7. Nigam VP, Graupe D (2004) A neural-network-based detectionof epilepsy. Neurol Res 26(1):55–60
8. Tzallas AT, Tsipouras MG, Fotiadis DI (2007) Automatic seizure detection based on time-frequency analysis and artificial neural networks. Comput Intelligen Neurosci 2007 13 Article ID 80510
9. Prabhakar SK, Rajaguru H (2016) Development of patient remote monitoring system for epilepsy classification. In: 16th International conference on biomedical engineering (ICBME), Singapore, December 7–10
10. Rajaguru H, Prabhakar SK Hilbert transform with elman back-propagation and multilayer perceptrons for epilepsy classification. In: IEEE (ICECA 2017), Coimbatore, India, pp 571–576
11. Donoho DL, Grimes C (2003) Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. Proc Nat Acad Sci 100(10):5591–5596
12. HRajaguru H, Prabhakar SK Analysis of Adaboost classifier from compressed EEG features for epilepsy detection. In: IEEE proceedings of ICCMC 2017, Erode, India

# Performance Analysis of Factor Analysis and Isomap with Hybrid ABC-PSO Classifier for Epilepsy Classification

S. K. Prabhakar and H. Rajaguru

**Abstract**

In human beings, one of the prevalent and most disturbing neurological disorder is epilepsy which is observed by recurrent seizures due to the abnormal electrical activities of the brain. Due to epilepsy, various symptoms like loss of consciousness, muscle jerks and spasms, fatigueness and so on occur. As the seizures occur randomly, the patients cannot be aware of it and so the physical risk and injury associated with it is quite high. To diagnose and analyze epilepsy, Electroencephalography (EEG) signals are used as an important tool. The recorded scheme of electrical activity due to the excessive firing of neurons within the brain can be measured with the help of EEG. The entire visual analysis of EEG signals is difficult due to its lengthy range of recordings. Therefore, in this paper Factor Analysis and Isomap are used to reduce the dimensions of the EEG data and then the dimensionally reduced values are classified with the help of Hybrid Artificial Bee Colony—Particle Swarm Optimization (ABC-PSO) algorithm to classify the epilepsy from EEG signals. The results show that an average classification accuracy of 96.90% along with an average time delay of 2.23 s is obtained when Factor Analysis is classified with Hybrid ABC-PSO. When Isomap is used along with Hybrid ABC-PSO, a classification accuracy of 97.87% is obtained along with an average time delay of 2.07 s.

**Keywords**

EEG • Epilepsy • Factor analysis • Isomap

## Introduction

One of the long lasting and severe neurological disorder is epilepsy and it is characterized by electrophysiological disturbances and seizures in the human brain [1]. The seizure may range from short period of muscle bumps to prolonged periods of robust shaking in the human body. In the nerve cells of the cerebral cortex, when unusual paroxysmal discharges occur the seizure occurs and so the brain can become dysfunctional. With the help of physiological EEG signals, the status of the information about the brain status can be known easily and so it is widely utilized to understand and analyze the various activities of the brain. Most significantly, a lot of vital information is provided by EEG to the doctors and clinicians regarding the epileptic seizure disorder. EEG signals on visual inspection require a very high level of keen concentration because of its lengthy nature. As EEG signals are highly non linear and non stationary in nature, processing it is quite difficult. By fixing an electrode on the scalp of the subject with the help of placement of standardized electrodes, the EEG recordings can be done easily. Various artifacts present in the signals severely affect the quality of the signal and so elimination of it is important. Due to the power line electrical noise, muscular activities and eye blinking during the signal acquisition procedure, the noise can happen easily. As the EEG

S. K. Prabhakar (✉) · H. Rajaguru
Department of ECE, Bannari Amman Institute of Technology, Alathukombai Post, Sathyamangalam, India
e-mail: sunilprabhakar22@gmail.com



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_32

recordings are quite long, it will possess a huge number of data along with noise and to process it is quite impractical. Therefore dimensionality reduction is used to reduce the dimensions of the data and then it is classified with post classifiers. In literature, very good works have been proposed for epilepsy classification from EEG signals. Some of the famous works are as follows.

An overview about the various techniques for seizure detection and prediction was done by Giannakakis et al. [2]. With the help of linear and non linear EEG analysis methods, the absence seizure epilepsy detection was proposed by Sakkalis et al. [3]. The code converters technique and the kernel maximum uncertainty discriminant analysis for epilepsy classification were done by Rajaguru and Prabhakar [4]. An evidence theory based approach for detection of epileptic seizures using EEG signals was done by Mohamed and Shaban [5]. The detection of epileptic seizures was done with the help of neural fuzzy networks by Sadati et al. [6]. For the implementation of the EEG automatic seizure detection, an adaptive structure neural network was used by Weng and Khorasawi [7]. The Bayesian Linear Discriminant Analysis (BLDA) along with Hybrid ABC-PSO was implemented for epilepsy classification in [8]. In this paper, Factor Analysis and Isomap are used as dimensionality reduction techniques and then it is classified with Hybrid ABC-PSO Classifier. The organization of the paper is as follows. In Sect. Materials and Methods, the materials and methods are explained followed by the usage of Hybrid ABC-PSO as the post classifier in Sect. Hybrid ABC-PSO for Classification. The results and discussion are given in Sect. Results and Discussion and it is concluded in Sect. Conclusion. The pictorial representation of the work is depicted in Fig. 1 as follows.

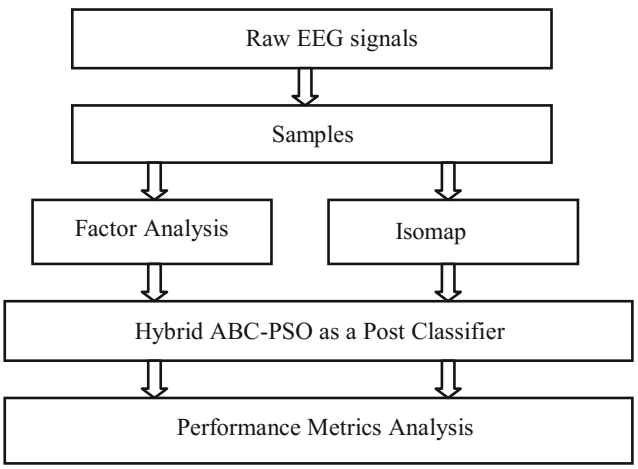



**Fig. 1** Block diagram of the technique

## Materials and Methods

In this study, a total of 20 epileptic patients were considered. The epileptic patients were admitted at the Neurology Department of Sri Ramakrishna Hospital, Coimbatore, India for their treatment and with the permission of both the doctors and patients the EEG recordings were taken from them. For various stages like eyes closed/open, fully sleep, partially awake and fully awake, the recordings was measured. According to the standard 10–20 International system, the 16 channel electrodes were kept on the scalp of the epileptic patients and then in the European Data Format (EDF) the recordings were obtained. Each patient had 16 channels and each channel was split into epochs, probably 3–4. The recordings were done for more than 55 min and later using EEG amplification equipments, the signal was amplified. Approximately, 400 values are embedded in a single epoch of a single channel for a patient. As the total values in all the epochs of all the channels for all the patients exceeds a very high level, processing it is quite difficult and so dimensionality reduction techniques are absolutely necessary.

A. *Factor Analysis for Dimensionality Reduction:*

For data reduction, one of the commonly used techniques is factor analysis and it is a multivariate technique [9]. A set of variables is represented by a smaller number of variables called factors. Such factors are assumed to be underlying constructs which cannot be easily measured by a single variable. Factor analysis is designed for both internal data and ordinal data. There should always be a linear relation for both the different variables used in Factor Analysis. There must be at least a moderate correlation among the variables utilized. The algebraic representation of the factor analysis model is done as follows:

If there are $'q'$ variables $Y_1, Y_2, \ldots, Y_q$ measured on a sample of $'m'$ subjects; then the variable $j$ is expressed as a linear combination of $s$ factors $F_1, F_2, \ldots, F_s$, where $s < q$. Therefore,

$$Y_j = b_{j1}F_1 + b_{j2}F_2 + \ldots. + b_{js}F_s + h_i$$

where the $b_j$ is known as the factor loadings or scores for variable $j$ and $h_i$ is the part of the variable $Y_j$ that cannot be explained by the factors.

The 3 vital steps in factor analysis are as follows:

**First Step**: The initial factor loadings are calculated with the help of 2 techniques such as principal component method or principal axis factoring method. Principal Component Analysis (PCA) is used in principal component method and

the obtained factors will not be the actual principal components though the coefficients of the $k$th principal component is proportional to the loadings of the $k$th factor. In the principal axis factoring method, it aims to seek the lowest number of factors which amounts for the variability in the original variables. Here the principal component method is chosen, and the resulting factors produced at this stage are uncorrelated.

**Second Step**: The rotation of factors is done in this step. The factors are rotated once the calculation of the initial factor loadings is done. This is implemented so that the factors which can be interpreted easily are found out. If there are several clusters of variables, then the rotation aims to make the variable present in a subgroup as high as possible. Oblique and Orthogonal rotation are the two types of rotations used. In oblique rotation, the resulting factors will be correlated and in orthogonal rotation, the resulting factors will be uncorrelated and so orthogonal rotation is used.

**Third Step**: In this step, the calculation of Factor Scores is done. For the calculation of the final factor scores, the value of the $'s'$ factors $F_1, F_2, \ldots, F_s$ is determined so that a decision can be made so as to know the inclusion of as many wanted factors as possible.

B. *Isomap for Dimensionality Reduction:*

In Isomap or isometric feature mapping, a pair of curvilinear distances is preserved in between the data points [9]. Firstly, on the manifold $M$, the determination of which points are neighbours are done based on the distances $d_Y(a, b)$ between the pair of points $(a, b)$ in the input space $Y$. Two simple techniques are present to connect each point to the rest of the points with some fixed radius $\varepsilon$. On the data points, a weighted graph $G$ is used to represent a neighborhood relation with edges of weight $d_Y(a, b)$ in between the neighboring points. The estimation of the geodesic distances $d_M(a, b)$ is done in between all pairs of points on a particular manifold $M$. It is done by the computation of the shortest path distances $d_G(a, b)$ in the graph $G$. To the matrix of graph distances $D_G = \{d_G(a, b)\}$, the classical Multidimensional Scaling (MDS) is applied so that the best preservation of the manifold's estimated geometry can be done. The global coordinates in the isomap gives a simplified methodology to manipulate high dimensional observations in terms of non linear degrees of freedom.

## Hybrid ABC-PSO for Classification

The dimensionally reduced values by means of Factor Analysis and Isomap are then classified with the help of Hybrid ABC-PSO Classifier for epilepsy classification. To avoid the local optima, ABC is widely used so that it can easily make up the drawback of PSO. Optimal solutions can be found out in PSO and so it is used widely. To make use of both the algorithms and to improve the computation process, the Hybriding of these 2 algorithms are done as follows [8].

**First step**: The parameters are initialized at the first step. The number of individuals of the swarm, search range of the solution, maximum circle-index and the necessary constants are set.

**Second step**: The colony is initialized as the second step. With a specific number of individuals, a colony is generated. As a bee colony, it is split into 2 types such as unemployed foragers and employed foragers based on the fitness value of each individuals. As a particle swarm, the fitness value of every particle is calculated and the best location is chosen as the global best location. The *iter* is used to represent the cyclic number and $iter = 1$.

**Third step**: In bee colony stage, for every forager which is employed, it searches for better sources while gathering honey from the current source. If the new source is found, then it switches to the new one from the current. For every forager which is unemployed, it tries to follow the employed forager based on the quality of sources available. If the source quality is better, then more unemployed foragers will turn to it. The employed one searches for the new sources found around and then the fitness value of each and every source is determined. Once the choices have been made, the generation of the best solution is done which is named as *Glob.Min*.

**Fourth step**: In the particle swarm stage, the location is found out by comparing the solutions of each particle experienced and it is named as *Ta_best*.

**Fifth step**: The minimum value of *Glob.Min* and the value of *Ta_best* which can be called as *Glob.Mins* is expressed by the following expression as

$$Glob.Mins = \begin{cases} Glob.Min & if\ Glob.Min \leq Ta\_best \\ Ta\_best & if\ Ta\_best \leq Glob.Min \end{cases}$$

As both these values of *Glob.Min* and *Ta_best* are equal to the value *Glob.Mins*, it is substituted to the next iteration, $iter = iter + 1$.

**Sixth step**: It is denoted that the number of circles is greater than the maximum of the circle-index. If it is not, then return to step 2 or else the computing process is ended and the value of *Glob.Mins* is recorded.

## Results and Discussion

When Factor Analysis and Isomap are utilized as dimensionality reduction techniques and when it is classified with Hybrid ABC-PSO Classifier, parameters like Performance Index, Accuracy, Specificity, Sensitivity, Time Delay and Quality Values are evaluated and the average results are

**Table 1** Factor analysis with hybrid ABC-PSO classifier

| Name | Epoch 1 | Epoch 2 | Epoch 3 | Average |
|---|---|---|---|---|
| PC (%) | 94.37 | 94.16 | 92.91 | 93.81 |
| MC (%) | 5.20 | 5.62 | 7.08 | 5.97 |
| FA (%) | 0.41 | 0.20 | 0 | 0.20 |
| PI (%) | 94.22 | 94.03 | 92.64 | 93.63 |
| Sensitivity (% | 99.58 | 99.79 | 100 | 99.79 |
| Specificity (%) | 94.79 | 94.37 | 92.91 | 94.02 |
| Time delay(sec) | 2.2 | 2.22 | 2.28 | 2.23 |
| Quality values | 22.33 | 22.34 | 21.96 | 22.21 |
| Accuracy (%) | 97.18 | 97.08 | 96.45 | 96.90 |

**Table 2** Isomap with Hybrid ABC-PSO classifier

| Name | Epoch 1 | Epoch 2 | Epoch 3 | Average |
|---|---|---|---|---|
| PC (%) | 96.03 | 96.03 | 95.20 | 95.76 |
| MC (%) | 1.87 | 2.29 | 3.95 | 2.70 |
| FA (%) | 2.08 | 1.66 | 0.83 | 1.52 |
| PI (%) | 95.86 | 95.89 | 95.05 | 95.60 |
| Sensitivity (%) | 97.91 | 98.32 | 99.16 | 98.47 |
| Specificity (%) | 98.12 | 97.70 | 91.04 | 97.29 |
| Time delay(sec) | 2.03 | 2.05 | 2.14 | 2.07 |
| Quality values | 22.40 | 22.59 | 22.52 | 22.50 |
| Accuracy (%) | 98.01 | 98.01 | 97.60 | 97.87 |

computed in Tables 1 and 2. The mathematical formulae for the Performance Index (PI), Sensitivity, Specificity and Accuracy are given as follows

$$PI = \left(\frac{PC - MC - FA}{PC}\right) \times 100$$

where PC = Perfect Classification, MC = Missed Classification and FA = False Alarm. The Sensitivity, Specificity and Accuracy measures are mathematically formulated by the following

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2}$$

The Quality Value $Q_V$ is mathematically expressed as

$$Q_v = \frac{C}{(R_{fa} + 0.2) * (T_{dly} * P_{dct} + 6 * P_{msd})}$$

where C indicates the scaling constant, $R_{fa}$ defines the number of false alarm per set, $T_{dly}$ mentions the average delay of the onset classification in seconds, $P_{dct}$ specifies the percentage of perfect classification and $P_{msd}$ indicates the percentage of perfect risk level missed.

The time delay is mathematically given as follows

$$Time\ Delay = \left[2 \times \frac{PC}{100} + 6 \times \frac{MC}{100}\right]$$

If Factor Analysis is considered as dimensionality reduction technique and when it is classified with Hybrid ABC-PSO Classifier, an average classification accuracy of 96.90% is obtained. An average quality value of 22.21 along with an average time delay of about 2.23 s is also obtained.

If Isomap is considered as dimensionality reduction technique and when it is classified with Hybrid ABC-PSO Classifier, an average classification accuracy of 97.87% is obtained. An average quality value of about 22.50 along with an average time delay of 2.07 s is also obtained.

## Conclusion

As epilepsy is a serious neurological menace to the human community, detection and classification of this disorder plays a vital role in medical research. Thus in this paper, two dimensionality reduction techniques were successfully implemented and later it was classified with ABC-PSO Classifier. Results show that an average classification accuracy of 96.90% is obtained when Factor Analysis is used with ABC-PSO and an average classification accuracy of 97.87% is obtained when Isomap is used with ABC-PSO. Future works aim to discuss different dimensionality reduction techniques with ABC-PSO Classifier.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Tzallas AT, Tsipouras MG, Fotiadis DI (2009) Epileptic seizure detection in EEGs using time-frequency analysis. IEEE Trans Inf Tech Biomed 13:703–710
2. Giannakakis G, Sakkalis V, Pediaditis M, Tsiknakis M (2015) Methods for seizure detection and prediction: an overview. Modern Electroencephalogr Assess Tech 91:131–157
3. Sakkalis V, Giannakakis G, Farmaki C et al (2013) Absence seizure epilepsy detection using linear and nonlinear EEG analysis methods. In: 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC). Osaka, Japan
4. Rajaguru H, Prabhakar SK (2017) Code converters methodology with kernel maximum uncertainty discriminant analysis for epilepsy

classification. In: 2nd IEEE international conference on communication and electronics systems. Coimbatore, India
5. Mohamed A, Shaban KB (2013) Evidence theory-based approach for epileptic seizure detection using EEG signals. In: Proceedings of the 12th IEEE international conference on data mining workshops, pp 79–85. Brussels, Belgium
6. Sadati N, Mohseni HR, Maghsoudi A (2006) Epileptic seizure detection using neural fuzzy networks. In: Proceedings of the IEEE international conference on fuzzy systems, pp 596–600. Vancouver, Canada
7. Weng W, Khorasani K (1996) An adaptive structure neural networks with application to EEG automatic seizure detection. Neural Netw 9(7):1223–1240
8. Rajaguru H, Prabhakar SK (2017) Bayesian linear discriminant analysis with hybrid ABC-PSO classifier for classifying epilepsy from EEG signals. In: IEEE proceedings of the international conference on computing methodologies and communication (ICCMC 2017). Erode, India
9. Harikumar R, Kumar PS (2015) Dimensionality reduction techniques for processing epileptic encephalographic signals. Biomed Pharmacol J 8(1):103–106

# Performance Analysis of Breast Cancer Classification with Softmax Discriminant Classifier and Linear Discriminant Analysis

S. K. Prabhakar and H. Rajaguru

**Abstract**

The early detection and classification of cancer is very important in order to save the life of a person. One of the dreadful diseases affecting ladies is breast cancer and it is a major concern in the medical field. The breast cancer arises from the tissues of the breast cells. Similar to other parts of the human body, breast comprises of numerous microscopic cells. In the case of breast cancer, the multiplication of the cells happens rapidly in the breast and spreads to other parts of the human body. In the recent years, various machine learning and soft computing techniques were employed to classify various medical issues including breast cancer. In this paper, the breast cancer was classified with the aid of two techniques such as Softmax Discriminant Classifier (SDC) and Linear Discriminant Analysis (LDA). Results show that an average classification accuracy of 97.75% is obtained when LDA is used and an average classification accuracy of 100% is obtained when SDC is used.

**Keywords**

Breast cancer • SDC • LDA • Accuracy

## Introduction

Due to the rapid advancements in engineering and medical fields, the main reason of death has seen a paradigm shift from infectious diseases to cardiovascular problems, tumours and cancers. Among the cancers, the breast cancer is one of the fatal cancers affecting women and the second most prominently occurring cancers next to lung cancer in women [1]. If the cancer is detected at a much earlier stage, then it easily reduces the risk of death. The Breast Cancer affects women of various ages and the breast tissue; lobules are severely affected by this disease. Based on the origin, the breast cancer can be classified into two types as ductal carcinoma and lobular carcinoma. If the cancer originates from milk ducts then it is known as ductal carcinoma and if it originates from lobules, then it is known as lobular carcinoma. The exact causes of breast cancer are not known but a lot of risk factors attribute to this cancer. The environmental and genetic factors play a vital role in this cancer. The environmental factors like poor diet, obesity, lack of physical activities and alcohol consumption are the main causes. In genetic factors, the family history, personal health status, density of breast tissue, menstrual history, age, gender, genome changes etc. are the main causes [2]. A lump formation is the initial symptom of a breast cancer and it is due to the tiny calcium deposits. These tumours can be malignant or benign. The malignant tumours are very dangerous as it can be quite aggressive and the cancer can spread over all the body cells. The benign cells are generally non cancerous and non aggressive in nature. During ancient periods, without the advent of effective and efficient techniques, the treatment and diagnostic procedures proved to be quite fatal. But today due to the automated computer aided detection and classification techniques, the breast cancer can be easily detected and classified. Some of the famous works done in this field are as follows.

S. K. Prabhakar (✉) · H. Rajaguru
Department of ECE, Bannari Amman Institute of Technology, Alathukombai Post, Sathyamangalam, India
e-mail: sunilprabhakar22@gmail.com



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_33

With the help of Gaussian Mixture Model (GMM) and Radial Basis Function (RBF), the breast cancer classification was done by Rajaguru and Prabhakar [3]. A computer aided system for an in-depth analysis of breast cancer classification was done by Shanthi and Bhaskaran [4]. Fuzzy C means technique was used to detect breast cancer in Wisconsin prognostic breast cancer data sets by Tintu and Paulin [5]. With cascaded ensemble classifiers, a highly reliable breast cancer diagnosing methodology was done by Zhang et al. [6]. For the diagnosis of breast cancer on various datasets, the neural network approaches was utilized by Chunekar and Ambulgekar [7]. For the breast cancer diagnosis problem, the application of K-nearest neighbours algorithm was implemented by Sarkar and Leong [8]. In this work, SDC and LDA are used to classify the breast cancer. The organization of the paper is as follows. In section "Materials and Methods", the materials and methods are discussed followed by the classification of breast cancer using SDC and LDA in section "Classification Using SDC and LDA". The results and discussion are given in section "Results and Discussion" and the conclusion is given in section "Conclusion". The block diagram of the work is given in Fig. 1.

## Materials and Methods

For totally 82 breast cancer patients, the dataset was obtained from the Oncology Department of Sri Kuppuswamy Naidu Hospital, Coimbatore, India. The patients were taking their treatment at the hospital and with consent of both the cancer patients and the oncologists the data was collected from them. The data was obtained from both reports and charts.

The main reports analyzed are pathological, radiological and operative while the main charts analyzed was both hospital and referral charts. The primary parameters analyzed for breast cancer detection and classification are lump size and location, the size of the breast, the position of the nipples and the nipple level, oral hygiene, family planning, history of breast feeding, menopause, menstrual cycle, menstrual duration, number of abortions, age, marital status, general habits, food habits, lifestyle, family history of breast cancer, previous surgeries and various tests reported from the current treatment. Into four different stages (T1–T4), the stages of breast cancer was classified according to the standard International Union Against Cancer (UICC) [3]. T4 means the severity level is higher and T1 means the severity level is lower. T3 is less severe than T4 and more severe than T2. T2 is more severe than T1 and less severe than T3. In T1 stage, 20 patients are present, in T2 stage 28 patients are present, in T3 stage 11 patients are present and in T4 stage 23 patients are present. Once the clinical values are obtained, it is converted into its respective numerical values. The Chi square test is performed to the numerical values while simultaneously classifying it with both SDC and LDA classifier. Then it is compared with the Tumour Node Metasis (TNM) stage to evaluate the results.

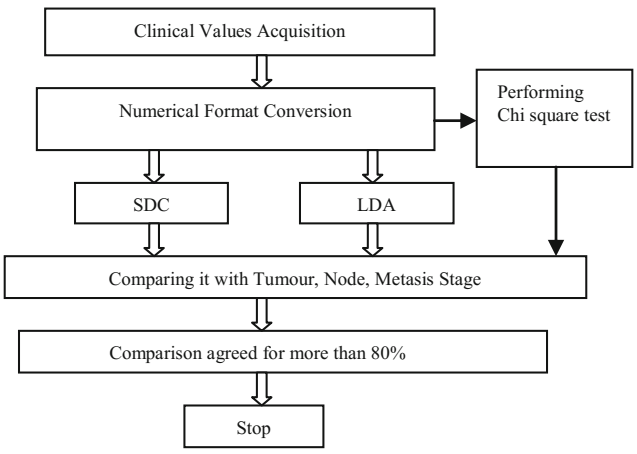



**Fig. 1** Block diagram of the work

A. *Chi Square Test:*

A very significant test regarding the non-parametric analysis is the Chi square test. It is widely used as it is easy for interpretation and analysis of any nominal data. It is commonly used in studies dealing with discrete data, demographics and so on. In one or more category, to determine a vital difference between the expected frequencies and the observed frequencies, Chi square test is utilized. The primary Chi square test requirements are that it should have a quantitative data, must have one or more categories, should possess adequate sample size, must have independent observations and should have a simple random sample. The Chi square formula utilized here is as follows

$$\chi^2 = \frac{(O_q - E_q)^2}{E_q}$$

where $O_q$ is the observed number of cases in category $q$, and $E_q$ is the expected number of cases in category $q$. By calculating the difference between the observed number of cases and the expected number of cases in each and every category, the Chi square test is performed. Then the squaring of difference is done and it is divided by the expected number of classes. After the calculation of Chi square values, sometimes degree of freedom is also computed because it gives us the total number of values that are free to differ after the restriction has been imposed on the data. For the breast cancer patient, the results of the Chi Square Test are computed in Table 1.

**Table 1** Chi square test measures for the breast cancer patients

| | Performance of chi square test | | |
|---|---|---|---|
| | Minimum value | Maximum value | Average value |
| T1 | 3.84 | 19.52 | 11.7 |
| T2 | 21.2 | 32.96 | 26.9 |
| T3 | 30.72 | 40.85 | 35.5 |
| T4 | 41.92 | 90.08 | 66 |

With respect to the data distribution, Chi square test is very robust and its ease of computation helps to handle the data in a flexible manner.

## Classification Using SDC and LDA

The breast cancer classification analysis was done with the help of both SDC and LDA classifiers. The primary inputs were fed inside the classifiers to detect and analyze the breast cancer classification process.

### A. *Softmax Discriminant Classifier*

Assume the training set $D = [D_1, D_2, \ldots, D_l] \in \Re^{f \times g}$ is obtained from '$l$' distinct classes. $D_l = \left[D_1^l, D_2^l, \ldots, D_{g_l}^l\right] \in \Re^{f \times g_l}$ indicates $g_l$ samples from the $l$th class, where $\sum_{i=1}^{l} g_i = g$. The primary intention of SDC is to determine the specific class to which a testing sample belongs to [9]. By using the weighting distance between the testing sample and the training sample it can be done easily. Assuming $d \in \Re^{f \times 1}$ is the testing sample. In SDC, the $l$ class samples are used to represent the test sample so that the achievement of minimum reconstruction error is obtained [9]. Between the $l$ class samples and the testing samples, the non linear transformation value is maximized and it seems to be the primary intention of SDC. The SDC is defined as follows

$$z(d) = \arg\max_j h_d^i$$

$$z(d) = \arg\max_j \log\left(\sum_{n=1}^{g_i} \exp\left(-\lambda \left\| d - d_n^j \right\|_2\right)\right)$$

where $h_d^j, m(d)$ represents the distance in between the testing sample and *jth* class, so that $d$ is identified. The parameter $\lambda > 0$ can provide a relative penalty cost. If $d$ belongs to a *jth* class, then $d$ and $d_n^j$ would have same features and so $\left\| d - d_n^j \right\|_2$ will be chosen to zero. The achievement of maximum value is done asymptotically by $h_d^j$ so that the value is maximized.

### B. *Linear Discriminant Analysis*

The optimal solution vector is selected as the best projection direction dependent on the Fisher criterion function in LDA [10]. After the projection is done in this direction, the original dataset has homogeneous samples closely related to each other and the heterogeneous samples are separated quite far away from each other. Assuming a set $G$ consisting of '$M$' $d$-dimensional samples $g_1, g_2, \ldots, g_m$ in which $M_1$ samples belonging to class $v_1$ is indicated as subset $G_1$ and $M_2$ samples belonging to class $v_2$ are indicated as subset $G_2$. The linear combination of the $G_m$ components is expressed as

$$z_M = v^T g_m, m = 1, 2, \ldots, M$$

In such a manner, a set of '$M$' one-dimensional samples $z_m$ is obtained which can be divided into 2 classes such as $z_1$ and $z_2$. The absolute value of $v$ has generally no influence on the classification process. The degree of separability is affected by the direction of $v$ after the projection of samples, thereby the LDA algorithm looks for the best transformation vector. The mean and prior probability of a particular class $v_i$ are $\mu_i$ and $Q_i$, then the distribution matrix $T_v$ and the inter-class scatter matrix $T_c$ of the samples are defined as follows

$$T_v = \sum_{i=1}^{2} Q_i T_i$$

$T_i = E\left\{(g - \mu_i)(g - \mu_i)^T \middle| g \in v_i\right\}$ is the covariance matrix of class $v_i$.

$$T_c = \sum_{i=1}^{2} Q_i (\mu_i - \mu_0)(\mu_i - \mu_0)^T$$

The mean of sample is represented as $\mu_0 = \frac{1}{N}\sum_{i=1}^{M} g_i$.

The maximization of the inter-class matrix and the intra-class matrix should be done so that the inter-class dispersion is maximized and the intra-class dispersion is minimized. The construction of the Fisher criterion function is done as follows

$$J(\eta_{opt}) = \arg\max_{\eta} \frac{|\eta^T T_c \eta|}{|\eta^T T_v \eta|}$$

The above equation is the generalized Rayleigh quotient of matrix $T_c$ which is relative to the matrix $T_v$. When $T_v$ is singular, the column vector $\eta_{opt}$ of the best projection matrix is simply the Eigen vector which corresponds to the largest Eigen Values.

**Table 2** Performance analysis SDC classifier for the different stages of breast cancer

| | T-1 | T-2 | T-3 | T-4 | Average |
|---|---|---|---|---|---|
| PC (%) | 100 | 100 | 100 | 100 | 100 |
| MC (%) | 0 | 0 | 0 | 0 | 0 |
| FA (%) | 0 | 0 | 0 | 0 | 0 |
| PI (%) | 100 | 100 | 100 | 100 | 100 |
| Sensitivity (%) | 100 | 100 | 100 | 100 | 100 |
| Specificity (%) | 100 | 100 | 100 | 100 | 100 |
| Accuracy (%) | 100 | 100 | 100 | 100 | 100 |

## Results and Discussion

For the Breast Cancer classification analysis, the bench mark parameters taken into consideration are Perfect Classification, Missed Classification, False Alarm, Sensitivity and Specificity measures followed by Accuracy measures and are represented by the following mathematical formulae as

$$PI = \left(\frac{PC - MC - FA}{PC}\right) \times 100$$

PC = Perfect Classification, MC = Missed Classification and FA = False Alarm. The Sensitivity, Specificity and Accuracy measures are mathematically expressed by the following formulae

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2}$$

Table 2 shows the Performance Analysis of SDC Classifier for the various stages of Breast Cancer based on the severity level. Table 3 shows the Performance Analysis of LDA Classifier for the various stages of Breast Cancer based on the severity level. Table 4 shows the Consolidated Accuracy Analysis of Clinical Values with the LDA classifier and SDC Classifier for the various stages Classification of breast cancer.

**Table 3** Performance analysis LDA classifier for the different stages of breast cancer

| | T-1 | T-2 | T-3 | T-4 | Average |
|---|---|---|---|---|---|
| PC (%) | 100 | 100 | 89.47 | 92.59 | 95.51 |
| MC (%) | 0 | 0 | 0 | 7.41 | 1.85 |
| FA (%) | 0 | 0 | 10.53 | 0 | 2.63 |
| PI (%) | 100 | 100 | 88.23 | 91.99 | 95.05 |
| Sensitivity (%) | 100 | 100 | 89.47 | 100 | 97.36 |
| Specificity (%) | 100 | 100 | 100 | 92.59 | 98.14 |
| Accuracy (%) | 100 | 100 | 94.73 | 96.29 | 97.75 |

**Table 4** Consolidated accuracy measures comparison

| S. No | Breast cancer stages | Classification by clinical means (%) | Classification by LDA (%) | Classification by SDC (%) |
|---|---|---|---|---|
| 1 | T1 | 98 | 100 | 100 |
| 2 | T2 | 100 | 100 | 100 |
| 3 | T3 | 97 | 94.73 | 100 |
| 4 | T4 | 100 | 96.29 | 100 |

## Conclusion

Thus breast cancer is one of the serious medical issues for women. Early detection and classification is a must to improve the quality of life of breast cancer patient. In this paper, the breast cancer was detected and classified with the help of SDC and LDA classifiers. The results show that for T1 Stage and T2 Stage, an average classification accuracy of 100% is obtained for both the classifiers. In T3 and T4 stages, the classification accuracy was 100% for SDC while for LDA it was 94.73% in T3 stage and 96.29% in T4 stage respectively. The overall average classification accuracy for LDA is 97.75% while for SDC it was 100%. Future works aim to utilize various other techniques for efficient breast cancer classification.

**Conflict of Interest** The authors declare that they have no conflict of interest

## References

1. Senthilkumar B, Umamaheswari G (2013) Combination of novel enhancement technique and fuzzy c means clustering technique in breast cancer detection. Biomed Res 24(2):252–256
2. Dheeba J, Albert Singh N, Tamil Selvi S (2014) Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J Biomed Inform 49 (2014):45–52
3. Rajaguru H, Prabhakar SK (2016) A comprehensive analysis on breast cancer classification with radial basis function and gaussian mixture model. In: 16th international conference on biomedical engineering (ICBME), 7–10 December
4. Shanthi S, Bhaskaran VM (2012) Computer aided system for detection and classification of breast cancer. Int J Inf Technol Control Autom (IJITCA) 2(4)
5. Tintu PB, Paulin R (2013) Detect breast cancer using fuzzy c means techniques in wisconsin prognostic breast cancer (WPBC) data sets. Int J Comput Appl Technol Res 2(5):614–617

6. Zhang Y, Zhang B, Coenen F, Lu W (2012) Highly reliable breast cancer diagnosis with cascaded ensemble classifiers. pp 10–15
7. Chunekar VN, Ambulgekar HP (2009) Approach of neural network to diagnose breast cancer on three different data set. In: International conference on advances in recent technologies in communication and computing. pp 893–895
8. Sarkar M, Leong TY (2000) Application of k-nearest neighbors algorithm on breast cancer diagnosis problem. In: AMIA annual symposium proceedings archive. pp 759–763
9. Zang F, Zhang J (2011) Softmax discriminant classifier. In: Third international conference on multimedia information networking and security. pp 16–19
10. Rajaguru H, Prabhakar SK (2017) Time frequency analysis (dB2 and dB4) for Epilepsy classification with LDA classifier. In: 2nd IEEE international conference on communication and electronics systems. Coimbatore, India, 19–20 October

# Part VII

# Behavioural Informatics and Connected Health Technologies

# Emotion Recognition from Haptic Touch on Android Device Screens

C. Maramis, L. Stefanopoulos, I. Chouvarda, and N. Maglaveras

**Abstract**

The recognition of the emotional state of an individual at a given time point provides valuable information that can find numerous health-related applications, e.g., interventions treating mental or physiological health problems. However, efficient emotion recognition remains a difficult task, often attempted subjectively, with obtrusive means and/or using specialized hardware. The present work uses haptic touch data acquired from Android smartphones to takes the first step towards the development of an objective, unobtrusive and real-life emotion recognition method that exploits the association between emotion and haptic touch. Focusing on four basic emotions (Excitement, Relaxation, Boredom and Frustration) the proposed method achieves very promising classification accuracy using a mixture of feature extraction and machine learning based classification techniques. A well-sized haptic touch dataset has been collected to support the method development and performance evaluation.

**Keywords**

Emotion recognition • Haptic touch • Android OS

## Introduction

The recognition of an individual's emotional state at a given point in time is of great value for numerous health-related settings. Monitoring and analyzing the emotional status in conjunction with treatment progress can provide actionable information concerning, for example, patients facing mild mental issues, orthopedic rehabilitation and almost any case of long-term recovery or treatment [1, 2]. In the context of medical intervention development, emotion recognition can also provide useful insight about the patients' satisfaction level with respect to the intervention, facilitating, for example, the prevention of drop-outs.

Despite the variety of its potential applications, the actual task of accurately recognizing a subject's emotion remains challenging. Several different techniques have been put to use for emotion recognition, including self-reporting, facial expression investigation [3], vocal pattern recognition [4] and biosignal analysis (EEG, ECG, etc.) [5]. However, the aforementioned emotion recognition techniques are either obtrusive and user-unfriendly or not applicable in real life. Instead, *non-obtrusive* methods that are able to *objectively* recognize the emotion of a subject in *real-time* and in *real-life settings* are missing.

The present work explores the well-known *association between emotion and haptic touch* [6, 7] to make the first step towards the development of an emotion recognition method with the aforementioned using haptic touch data acquired from Android devices. This step concerns algorithm development, where several feature extraction and machine learning (ML) based classification techniques are tested for their emotion recognition performance on a haptic

C. Maramis (✉) · L. Stefanopoulos · I. Chouvarda · N. Maglaveras
Department of Medicine, Aristotle University of Thessaloniki, KEDIP Building, University Campus, Thessaloniki, Greece
e-mail: chmaramis@med.auth.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_34

touch dataset collected through a well-designed data acquisition experiment. To the best of our knowledge, this is the first time haptic touch data acquired from Android devices are ever used for the emotion recognition.

## Emotion Recognition Background

The present work adopts the *Circumplex model of affect* [8], one of the most widely used Psychology-based models for emotions categorization. In brief, Circumplex is a 2-dimensional spatial model that presents the interrelations of affective concepts (emotions) in a circle. The vertical and horizontal axes of the circle show the intensity and the pleasantness of the emotions, respectively. The "representative" emotions of each circle quadrant and main recognition targets of this work are *Excitement* (intense and pleasant), *Relaxation* (mild and pleasant), *Boredom* (mild and unpleasant) and *Frustration* (intense and unpleasant). In the framework of the present work, the basic emotions have been stimulated by an openly accessible *Affective Picture System* (APS) [9], i.e., a dataset of images that, when displayed, triggers specific emotions. The employed APS has been validated against the International Affective Picture System [10].

## Methods

### Haptic Touch Data Acquisition

In order to develop the proposed emotion recognition method, an appropriate dataset *associating haptic touch information with the emotional state* of an individual had to be acquired. This subsection describes the designed data acquisition experiment and the acquired dataset itself.

(a) *Data Acquisition Experiment*

An Android application was developed for the purpose of data acquisition. Its main functionality was to evoke an emotion to the user and then ask the user to carry out a neutral *haptic activity* so as to record a variety of haptic touch data. The aforementioned APS (see section Emotion Recognition Background) was employed to stimulate the desired emotion. The participants were exposed to 4 sets of 10 randomly selected pictures from the APS, each set associated with a different basic emotion. Right after the presentation of a picture set, the participants were asked to carry out a simple haptic activity involving 10 touches to the Android device screen.

Twenty technology-competent subjects (*S1–S20*) were recruited for the experiment (10 males/10 females, mean age of 34). The experiment took place in a neutral, isolated environment with no external stimuli. Each participant used the developed application twice with a distance of 2 h between the iterations.

(b) *Acquired Dataset*

The `onTouchEvent()` callback method of the Android framework was used to retrieve information about the touch events constituting a haptic activity. For each touch event, 3 pieces of information were recorded: (1) the applied pressure on the touch screen, (2) the start-time of the touch event, and (3) the end-time of the touch event. The start- and end-times of a touch event were used for calculating the duration of the event as well as its spacing (i.e., time-distance) from the following touch event.

Since a haptic activity is completed in 10 touches, each activity includes 10 ordered pressure values, 10 ordered duration values, and 9 ordered spacing values. Table 1 demonstrates the schema of the acquired dataset and also provides a sample of the acquired data. In total, data concerning 40 haptic activities (20 subjects × 2 iterations) associated with specific basic emotions were collected.

### Emotion Recognition Algorithms

This section describes the development of emotion recognition algorithm from the acquired haptic touch data, starting

**Table 1** Haptic touch data acquired for the 1st iteration of S1. Distance and spacing are in milliseconds, while pressure is normalized to the range [0, 1]

| Emotion | Pressure vector | Duration vector | Spacing vector |
|---|---|---|---|
| Excitement | [0.243, 0.247, 0.247, 0.243, 0.243, 0.239, 0.243, 0.239, 0.243, 0.239] | [76, 73, 59, 43, 58, 44, 43, 45, 43, 29] | [0, 346, 360, 404, 405, 360, 375, 405, 404, 345] |
| Relaxation | [0.235, 0.235, 0.235, 0.231, 0.239, 0.235, 0.235,0.227, 0.235, 0.235] | [65, 103, 132, 87, 87, 72, 87, 59, 58, 58] | [0, 420, 451, 331, 362, 345, 422, 405, 404, 421] |
| Boredom | [0.220, 0.221, 0.227, 0.229, 0.227, 0.220, 0.225, 0.225, 0.216, 0.221] | [50, 74, 89, 87, 148, 89, 89, 119, 115, 82] | [0, 732, 762, 738, 808, 672, 719, 687, 691, 621] |
| Frustration | [0.253, 0.257, 0.253, 0.257, 0.257, 0.257, 0.251, 0.253, 0.257, 0.257] | [57, 60, 60, 27, 71, 57, 54, 69, 57, 57] | [0, 212, 225, 255, 244, 227, 244, 246, 375, 330] |

from the formulation of the problem, moving to the extraction of suitable features from the acquired data and ending up with the classification process.

(a) *Problem Formulation*

Taking into account the potential applications of an efficient real-time recognition system, the recognition of each one of the 4 basic emotions from haptic touch data has been identified as the primary objective of this work. However, the classification of the pleasantness of a perceived emotion is also of value to an emotion recognition system. Based on these facts, the present work has set two—prioritized—*classification targets*:

1. **Single Emotion Recognition**: The top priority classification target of this work is expressed as a 4-class classification problem, where each class corresponds to a basic emotion (Excitement, Relaxation, Boredom and Frustration).
2. **Pleasantness Recognition**: This classification target is seemingly less ambitious yet easier, translated into a binary classification problem whose classes are the pleasant emotions (Excitement and Relaxation) and the unpleasant ones (Boredom and Frustration).

The complexity of emotion recognition from the collected data suggests the exploitation of state-of-the-art *machine learning (ML) classifiers*. On top of that, the existence of subject specific haptic touch patterns associated with different emotions renders the *personalized training* of the ML classifiers necessary.

(b) *Feature Selection*

From the description of the acquired haptic touch dataset, it becomes clear that 3 attributes of a haptic touch have been deemed to be useful for emotion recognition: the applied **pressure (*P*)**, the touch **duration *(D)***, and its **spacing (*S*)** from the next touch event. However, the selected attributes might not perform equally—or at all—well in discriminating between emotions or emotion categories, meaning that various combinations of attributes were tested in the feature selection process.

Given a combination of touch event attributes, the way of constructing a feature vector to be fed to the ML classifiers is not straightforward. Two main options were investigated: (1) The *summarization* of the values of each attribute using certain properties of a probability distribution as features (e.g., the mean and standard deviation of the normal distribution); (2) The utilization of *series of attribute values* as features to exploit the sequential nature of the haptic touch events. In the second option, the ordered attribute values constitute the individual features of the feature vector. The number of haptic touch events to be employed for building the feature vector is an open issue. Since the analysis of the acquired dataset has provided strong indications in favor of exploiting the sequential nature of the touch events (second option), the first option was abandoned.

Based on the described feature selection methods, the feature vector extracted from a haptic activity when all 3 attributes are used is given by the expression:

$$[P_1, P_2, \ldots, P_N, D_1, D_2, \ldots, D_N, S_1, S_2, \ldots, S_{N-1}] \quad (1)$$

c) *Classification*

Six well-known ML classifiers, namely *Naive Bayes*, *Decision Tree*, *Logistic Regression*, *Multi-Layer Perceptron* (MLP), *Support Vector Machines* (SVM) and *Random Forest*, were tested on the acquired haptic touch dataset with respect to their performance (i.e., classification accuracy) on the two aforementioned classification targets. The classifiers were fed with feature vectors of the form given by (1) using different combinations of touch attributes. Adopting the personalized training approach, for each subject the haptic touch data acquired from the 1st iteration of the experiment were used to train the ML classifiers and the data from the 2nd iteration to evaluate the classifiers' performance. For those classifiers that rely on a random initialization step (e.g., Random Forest), the training and evaluation process was repeated 5 times and the highest obtained accuracy was recorded.

## Results

The various aspects of the proposed emotion recognition method were thoroughly tested on the acquired haptic touch dataset with the final goal being, of course, the maximization of the classification accuracy. This section provides the most important conclusions drawn from the testing process along with some examples and highlights.

The experimentation with the touch event attributes has shown that the inclusion of all 3 of them (*P*, *D*, *S*) is beneficial for the classification accuracy. Moreover, the utilization of the first five touch events in the construction of the feature vector is a solid choice, since the effect of the affective pictures on the target emotion seems to fade out near at the second half of the haptic activity. Figure 1 provides an example in support of the above statements.

Regarding the performance of the tested ML classifiers, the results are very promising. Table 2 presents the

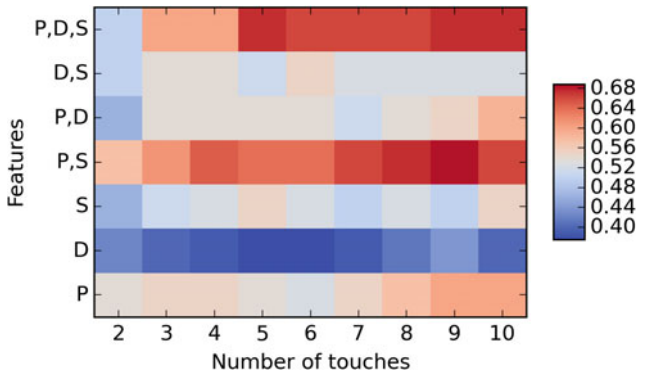



**Fig. 1** Classification accuracy heatmap of the SVM classifier on the single emotion target for various attribute sets and touch event numbers

**Table 2** Classification accuracy of all 6 classifiers per classification target [Feature selection: all 3 attributes, first 5 touch events]

| | Classification target | |
|---|---|---|
| Classifier | Pleasantness (%) | Single emotion (%) |
| Naive bayes | 67.50 | 67.50 |
| Decision tree | 70.00 | 56.25 |
| Logistic regression | 78.75 | 63.75 |
| MLP | 68.75 | 60.00 |
| SVM | 75.00 | 67.50 |
| Random forest | 77.50 | 55.00 |

classification accuracy of the six classifiers when feature vectors that incorporate all 3 attributes and the first 5 touch events are used. Naive Bayes and SVM achieve the best accuracy in the single emotion target, while the Logistic Regression is the winner of the pleasantness target.

The performance of the 6 classifiers on each subject under the previous feature selection configuration for the single emotion classification target is visualized in Fig. 2, which

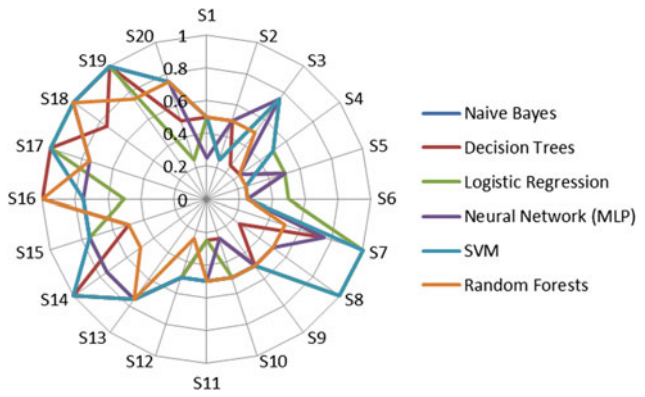



**Fig. 2** Per subject accuracy of the 6 classifiers for the single emotion classification target [Feature selection: all 3 attributes, first 5 touch events]

**Table 3** Confusion matrix of the SVM-based classification for the single emotion target (R: real; P: predicted) [Feature selection: all 3 attributes, first 5 touch events]

| R\P | Excit. | Relax. | Bored. | Frustr. |
|---|---|---|---|---|
| Excit. | 14 | 1 | 0 | 5 |
| Relax. | 1 | 14 | 4 | 1 |
| Bored. | 2 | 3 | 13 | 2 |
| Frustr. | 5 | 0 | 2 | 13 |

demonstrates noticeable variability in the classifiers' performance among the subjects. This fact could be exploited to enhance the proposed approach in the future: the selection of the classifier that is expected to maximize the classification accuracy on a per-subject case, if possible, is going to significantly increase the method's performance.

Going a bit deeper into the classification results, Table 3 provides the confusion matrix of the SVM-based classification on the single emotion target using the previous feature selection configuration. This very indicative example shows that the SVM as well as the other classifiers (*data not shown*) find it particularly difficult to discriminate between intense emotions (Excitement versus Frustration) and—to a lesser extend—between mild emotions (Relaxation versus Boredom). This observation provides a promising research direction for improving the proposed method.

## Discussion and Conclusions

This paper has presented a novel emotion recognition method from haptic touch on screens of Android devices. For the development and the evaluation of the method a 2-iteration data collection experiment involving 20 subjects has been conducted. The proposed method proposes a personalized training approach and has tested several combinations of haptic touch related features and state-of-the-art classifiers. The obtained classification results have been very satisfactory. The proposed method has been able to classify the 4 main emotions (Excitement, Relaxation, Boredom, and Frustration) with an accuracy of 67.5% and to distinguish between pleasant and unpleasant emotions with accuracy of 78.75%.

The analysis of the classification results has also provided a couple of directions for improving the efficiency and robustness of the emotion recognition method. These are associated with (1) the per-subject selection of the most suitable ML classifier and (2) focusing of the better discrimination of emotions that lie at the same end of the intensity spectrum. As soon as the algorithmic part of the proposed method is finalized, the proposed method is going

to be implemented and distributed as an Android library to be exploited in 3rd party Android application development.

**Acknowledgements** This research has been realized in the framework of a State Scholarships Foundation (IKY) Scholarship funded by the Action "Strengthening Post-Doctoral Research"—Co-financed by the ESF and the Greek government.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Danermark BD (1998) Hearing impairment, emotions and audiological rehabilitation: a sociological perspective. Scand Audiol 27:125–131. https://doi.org/10.1080/010503998420757
2. Laederach-Hofmann K, Roher-Gubeli R, Messerli N, Meyer K (2007) Comprehensive rehabilitation in chronic heart failure—better psycho-emotional status related to quality of life, brain natriuretic peptide concentrations, and clinical severity of disease. Clin Invest Med 30:54–62. https://doi.org/10.25011/cim.v30i2.980
3. Busso C, Deng Z, Yildirim S et al (2004) Analysis of emotion recognition using facial expressions, speech and multimodal information. In: Proceedings of the 6th international conference on multimodal interfaces. ACM, New York, NY, USA, pp 205–211
4. Yacoub SM, Simske SJ, Lin X, Burns J (2003) Recognition of emotions in interactive voice response systems. In INTERSPEECH
5. Haag A, Goronzy S, Schaich P, Williams J (2004) Emotion recognition using bio-sensors: first steps towards an automatic system. Affective dialogue systems. Springer, Berlin, Heidelberg, pp 36–48
6. Collier G (2014) Emotional expression. Psychology Press
7. Hertenstein MJ, Holmes R, McCullough M, Keltner D (2009) The communication of emotion via touch. Emotion 9:566
8. Posner J, Russell JA, Peterson BS (2005) The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev Psychopathol 17:715–734. https://doi.org/10.1017/S0954579405050340
9. Machajdik J, Hanbury A (2010) Affective image classification using features inspired by psychology and art theory. In: Proceedings of the 18th ACM international conference on multimedia. ACM, New York, NY, USA, pp 83–92
10. Lang PJ, Bradley MM International affective picture system (IAPS): technical manual and affective ratings

# Objective Smoking: Towards Smoking Detection Using Smartwatch Sensors

C. Maramis, V. Kilintzis, P. Scholl, and I. Chouvarda

**Abstract**

Smoking is arguably one of the most harmful lifestyle behaviors with well-known relations to dozens of diseases. As a consequence, smoking monitoring and cessation support have been popular targets for the emerging field of Behavioral Informatics. Although smoking monitoring has been attempted in the past (e.g., self-reporting), the revolution of smart wearable devices has provided the necessary tools for developing truly objective and unobtrusive smoking detection methods applicable in real-life settings. This work takes the first step towards the development of such a method using smartwatch inertial sensory data, with the end-goal of integrating the proposed method in a novel just-in-time intervention for smoking cessation. The paper explores the detection of smoking instances and the constituting puffs from 3-axis gyroscope data that are routinely acquired by any Android Wear smartphone. The feasibility of the method is tested on actual real-life data, yielding very promising preliminary results.

**Keywords**

Smoking behavior • Smoking and puffing detection • Smartwatch sensors • Android wear

## Introduction

*Smoking* is an extremely harmful lifestyle behavior associated with an enormous list of diseases [1]. Because of its well-known implications to health, smoking behavior has been a popular target of study for *Behavioral Informatics (BI)*, an emerging interdisciplinary scientific field that leverages information technology to develop interventions for the modification of harmful/unwanted behavioral patterns.

Successful *smoking monitoring*, i.e., the efficient detection and/or prediction of a smoking event, can find multiple applications, including the support of just-in-time interventions [2] that aim at smoking cessation. Although various techniques have been employed in the past (e.g., *self-reports* of smoked cigarettes), solutions that unobtrusively and accurately detect smoking instances without input from the smoker are missing. The technological revolution of wearable computing provides the necessary tools for building the aforementioned solutions: Today's smart wearables (e.g., smartwatches) are equipped with dozens of sensors that can be used for effectively detecting smoking instances.

This work takes the first step towards the development of a novel method for unobtrusive and objective smoking detection with the help of smartwatch inertial sensory data; this method is meant to form the basis of a just-in-time intervention for smoking cessation. More specifically, the paper explores the detection of smoking instances and the constituting puffs from 3-axis gyroscope data that are

C. Maramis (✉) · V. Kilintzis · I. Chouvarda
Lab of Computing & Medical Informatics, Department of Medicine, Aristotle University of Thessaloniki, KEDIP Building, University Campus, Thessaloniki, Greece
e-mail: chmaramis@med.auth.gr

P. Scholl
Embedded Systems, Department of Computer Science, University of Freiburg, Freiburg, Germany



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_35

routinely acquired by any Android Wear smartphone. The feasibility of the method is tested on actual real-life data.

## Related Work

The related work ecosystem of the present study includes approaches that employ *inertial measurements units (IMU) of wrist-worn sensors to detect smoking*. Several promising studies have shown evidence that smoking detection is feasible when using sensory data acquired from *custom* [3–5] or *commercial* [6, 7] wrist-worn devices. In some cases, the wrist-worn device is coupled with other wearable sensory devices [5, 6]. Another important distinction is associated with the data collection setting for the evaluation of the proposed methods, which can either be real-life [3, 4, 7] or some controlled laboratory environment [5, 6].

## Methods

### Design Considerations

As it has already been mentioned, the proposed smoking detection method is part of an effort to develop a just-in-time BI intervention for smoking cessation. When seen under this prism, the design of the introduced method had to take into account 3 main objectives:

(a) **Efficient Detection**: This is the primary objective of smoking detection, mainly expressed via the maximization of detection accuracy. However, in the context of a smoking cessation intervention it is probably just as important to ensure the highest sensitivity possible (i.e., to ensure that as many smoking instances as possible are detected).
(b) **Real-life Operation**: A smoking cessation intervention delivered via smart wearable devices is, of course, intended to work in real-life settings. This means that (1) the necessary input to the associated smoking detection method needs to be provided by everyday usage devices, and (2) the smoker should put minimum effort for the smoking detection method to work. The latter impacts the smoking instance annotation in the method development phase, as the manual annotation of the smoking instances by the smokers and the observation of the smokers for annotation purposes might interfere with the actual real-life behavior of the smoker.
(c) **Applicability**: The best way to maximize the potential of a BI intervention is to ensure that it is addressed to as many people as possible. This can be achieved by making sure that the intervention is not built upon expensive specialized hardware.

In order to optimize the applicability of the intended intervention, the inertial sensor measurements acquired from *Android Wear smartwatches* were chosen as input to the detection method. Choosing a family of popular commercial devices has several advantages over the custom or specialized hardware used by previous related studies. Concerning the real-life operation, the data collection experiment for the development of the proposed method was carefully designed to eliminate any need for input from the smokers and, thus, any interference with their real smoking behavior.

### Data Acquisition

In order to collect the sensory data that are necessary for the development of the proposed smoking detection method *a novel mobile computing system* has been developed. This comprises 3 hardware devices: (1) one Android Wear smartwatch to provide the inertial sensor measurements via the Android sensor framework; (2) one Spark Lighter, a smart sparking device to automatically and objectively annotate the smoking instances [8]; and (3) one Android smartphone to drive the other devices and store the sensory measurements and the smoking instance annotations. Software-wise, an Android app with two modules was developed. The first module runs on the smartphone to orchestrate the data collection process, while the second is responsible for acquiring measurements from the built-in 3-axis accelerometer and gyroscope of the smartwatch. Figure 1 outlines the components of the system.

The data collection system operates as follows: Every time the subject lights the Spark Lighter, the timestamp of the spark event is recorded by the smartphone and a trigger is sent to the smartwatch to start capturing the measurements of the 3-axis accelerometer and gyroscope at the highest possible sampling frequency for 5 min. The resulting batch of sensory data corresponds to a *smoking instance* and it is sent to the smartphone to be stored. Simultaneously, the smartphone periodically (approximately every 1 h) sends to the smartwatch triggers for recording 5 min *control* or *non-smoking instance* batches of measurements from the same sensors at the same frequency.

The data collection experiment took place in August and September 2017 in Thessaloniki (Greece). Four smokers (*S1-S4*; 3 males and 1 female between 32 and 36 year old) were asked to use the data collection system in their *real-life* for at least one day. The Spark Lighter and an LG G Watch W100 (to be worn on the hand that is used for smoking)


**Fig. 1** Hardware and software components of the mobile computing system that was implemented for data collection

were handed to the subjects; the mobile module of the app was installed on the personal smartphone of the subjects.

The data collection process was completed with no problems resulting in a total of 23 smoking instance batches and 107 control batches. The numbers of collected smoking instances/control batches per subject were as follows: *S1* = 4/4, *S2* = 5/16, *S3* = 9/73, *S4* = 5/14.

## Method Development

The easily manageable dataset collected from *S1* was explored and paradigmatic batches of smoking instances and controls were used as guides for developing the signal processing and feature extraction steps of the method. After visual comparison, the gyroscope signal was selected as input to the method, owing to its small noise levels. The steps of the proposed method, visualized by Fig. 2, are outlined below.

(a) **Signal Decomposition**: The timestamped 3-axis gyroscope measurements of each batch were decomposed along the X-, Y- and Z-axis to generate 3 1-d gyroscope signals.
(b) **B-Spline Based Resampling**: To ensure a steady sampling rate, the 3 signals were modeled as cubic B-splines and the models were then resampled at 200 Hz.
(c) **Smoothing**: The resampled signals were smoothed by applying a Savitzky-Golay filter with a window of 15 samples so as to remove potential high-frequency noise.
(d) **Convolution with Puff Pattern**: The visual inspection of the signals extracted from smoking instances reveals patterns corresponding to individual puffs. In order to highlight the puff events each smoothed signal was convoluted with an abstract 5 s puff pattern (see Fig. 3).
(e) **Negative Value Zeroing**: As the convolution of the previous step is expected to generate eminent positive peaks near the actual puff events, any negative values are of no use and are defaulted to zero.
(f) **Weighted Superposition**: In this step the 3 1-d signals are superimposed in a weighted summation process. The weights of each dimension have been assigned empirically as follows: $(w_x, w_y, w_z) = (0.4 \cdot M_z / M_x, 0.5 \cdot M_z / M_y, 1)$, where $w_i$ and $M_i$ are the weight and

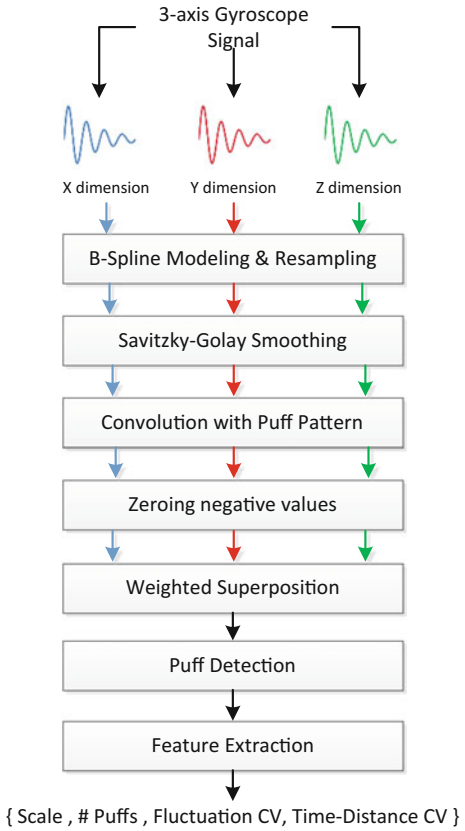



**Fig. 2** The steps (signal processing and feature extraction) of the proposed smoking detection method

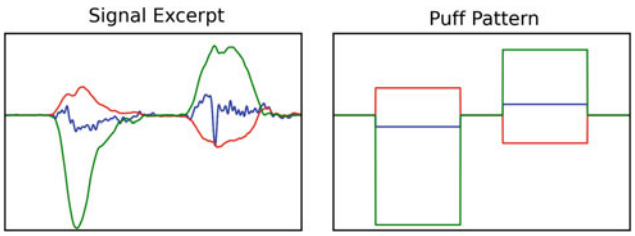



**Fig. 3** Excerpt of 3-d gyroscope signal corresponding to puff event (left) and abstract 3-d puff pattern to be convoluted with the signal (right)

maximum value of the convoluted signal for the *i* dimension, respectively.

(g) **Puff Detection**: In the case of a smoking instance, the local maxima of the superimposed signal correspond to the puff events. Local maxima detection is performed with the use of the 1st derivative of the signal followed by a refinement process to eliminate closely neighboring maxima.

(h) **Feature Extraction**: The detected puff events convey significant information for distinguishing smoking instances from non-smoking instances. Four high-level puff-related features have been empirically selected to express this information: (1) **Scale**—the value of the most outstanding puff event; (2) **Number of Puffs**—the total number of puff events detected in the batch; (3) **Fluctuation CV**[1]—this evaluates the dispersion of the values of the puff-associated peaks; (4) **Time-distance CV**—this evaluates the dispersion of the spacing of the puff-associated peaks in time.

## Results

The developed method was applied to all the collected batches of gyroscope data. The visual inspection of the outcome indicates that the proposed signal processing procedure highlights the puff patterns of the smoking instances to amplify the existing differences between smoking instances and controls (see Fig. 4 for an example).

On top of that, the proposed features show great promise in discriminating the two classes. As an example, Fig. 5 visualizes 3 of the features extracted from the data batches of *S2*, demonstrating the clear separation of all but one smoking instances from the controls.

Furthermore, in order to statistically assess the discriminative potential of the proposed features, the Student's *t*-test was applied on the extracted features from the data of each subject under the assumption of unequal variances. The resulting *p*-values are provided in Table 1 and they demonstrate that the inter-class difference of each selected feature is statistically significant in at least one subject.

The proposed features are going to be exploited for developing the classification step of the introduced method (most probably using a machine learning classifier). As an exercise, an SVM classifier using the RBF kernel with default parameters ($C$, $\gamma$) was employed to classify the extracted feature sets extracted from each subject. The cross-validation results that are obtained using the leave-1-instance-out

[1]Coefficient of Variation, i.e., the ratio of the standard deviation to the mean of a given distribution.

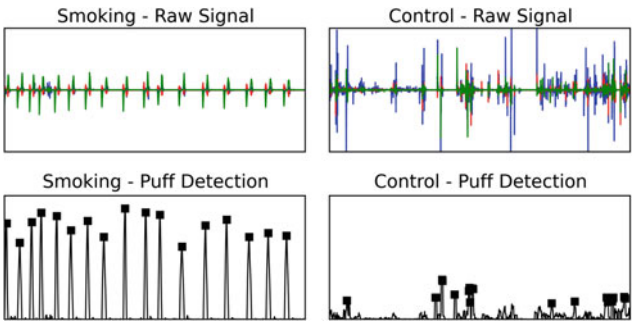



**Fig. 4** Examples of smoking instance (left column) and control (right column) batches

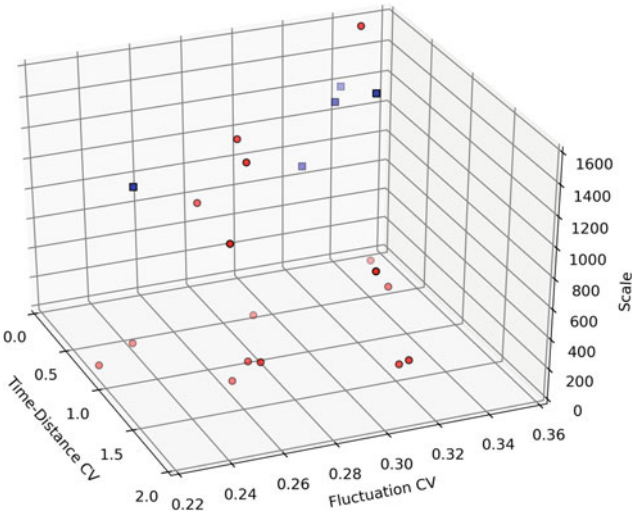



**Fig. 5** Scatter plot of 3 features extracted from the smoking instance (blue square) and control (red circle) batches of *S2*

**Table 1** P-values resulting from the application of t-test on the features extracted for each subject under the unequal variance assumption

| Features | S1 | S2 | S3 | S4 |
|---|---|---|---|---|
| Scale | 0.035 | 0.001 | 0.160 | 0.534 |
| # Puffs | 0.427 | 0.314 | 0.024 | 0.004 |
| Fluctuation CV | 0.070 | 0.366 | 0.723 | 0.006 |
| Time-distance CV | 0.041 | 0.096 | 0.703 | 0.002 |

**Table 2** Leave-1-instance-out cross-validation results per subject using an SVM classifier with the RBF kernel (default $C$, $\gamma$ values)

| Metric | S1 (%) | S2 (%) | S3 (%) | S4 (%) |
|---|---|---|---|---|
| Accuracy | 100 | 85.71 | 76.83 | 100 |
| Sensitivity | 100 | 80 | 77.78 | 100 |
| Specificity | 100 | 87.5 | 76.71 | 100 |

strategy are listed in Table 2 and they provide a strong indication that the extracted features are capable of distinguishing smoking instance from control batches.

## Conclusion

This work presented a sample of 4 smokers who contributed 23 smoking and 107 control instances of gyroscope-measured wrist motion. The proposed smoking detection method achieved out-of-the-box a mean leave-1-instance-out F1-score of 90%. Participants were minimally monitored (Spark Lighter and Android Wear smartwatch) closely resembling *real-life* conditions of just-in-time interventions, and as such presents one of the first studies with increased ecological validity for wearable smoking detection.

**Acknowledgements** This research has been realized in the framework of a State Scholarships Foundation (IKY) Scholarship funded by the Action "Strengthening Post-Doctoral Research"—Co-financed by the ESF and the Greek government.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Wald NJ, Hackshaw AK (1996) Cigarette smoking: an epidemiological overview. Br Med Bull 52:3–11. https://doi.org/10.1093/oxfordjournals.bmb.a011530
2. Spruijt-Metz D, Nilsen W (2014) Dynamic models of behavior for just-in-time adaptive interventions. IEEE Pervasive Comput 13:13–17. https://doi.org/10.1109/MPRV.2014.46
3. Scholl PM, Laerhoven K van (2012) A feasibility study of wrist-worn accelerometer based detection of smoking habits. In: 2012 Sixth international conference on innovative mobile and internet services in ubiquitous computing, pp 886–891
4. Parate A, Chiu M-C, Chadowitz C et al (2014) RisQ: recognizing smoking gestures with inertial sensors on a wristband. In: Proceedings of the 12th annual international conference on mobile systems, applications, and services. ACM, New York, NY, USA, pp 149–161
5. Saleheen N, Ali AA et al (2015) puffMarker: A multi-sensor approach for pinpointing the timing of first lapse in smoking cessation. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing. ACM, New York, NY, USA, pp 999–1010
6. Varkey JP, Pompili D et al (2012) Human motion recognition using a wireless sensor-based wearable system. Pers Ubiquit Comput 16:897–910. https://doi.org/10.1007/s00779-011-0455-4
7. Tang Q, Vidrine DJ et al (2014) Automated detection of puffing and smoking with wrist accelerometers. In: Proceedings of the 8th international conference on pervasive computing technologies for healthcare. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, pp 80–87
8. Scholl PM, Kücükyildiz et al (2013) When do you light a fire?: capturing tobacco use with situated, wearable sensors. In: Proceedings of the 2013 ACM conference on pervasive and ubiquitous computing adjunct publication. ACM, New York, NY, USA, pp 1295–1304

# Towards Value Propositions for Persuasive Health and Wellbeing Applications

M. S. Haque, A. Arman, M. Kangas, T. Jämsä, and M. Isomursu

**Abstract**

Recently, considerable attention has been given to health and wellbeing applications, specifically to persuasive applications. Persuasive applications refer to any interactive computing system designed to transform users' behaviours and attitudes. One of the major challenges of today's world is that health and wellbeing applications are not sustainable and scientifically designed. However, value proposition (VP) as a denominator might enhance the efficacy of the persuasive health and wellbeing applications. Research has shown little evidence on the VPs in health and wellbeing applications. This paper proposes key VPs for the persuasive health and wellbeing applications. A literature review was conducted based on relevant articles on the value within the health domain. Hence, narrative synthesis literature review approach had been used. We proposed and evaluated these VPs into our built persuasive health and wellbeing applications. We found that the VPs works well with our applications which might enhance their efficacy in the long run.

**Keywords**

Value co-creation • Value proposition • Persuasive health and wellbeing applications • Case study

## Introduction

To drive health and wellbeing applications, measurement and improvement of perceived value have been suggested but the value itself remains often misunderstood [1]. If value improves, users (customers, consumers, patients, and actors etc.) can be benefitted as well as the sustainability of the applications [1]. Moreover, VPs define 'why' users will accept the application and the 'offer' [2]. To co-create value, VPs convey users and other stakeholders' solution, i.e. by involving them in one body network [3]. On Service-Dominant Logic (SDL), service or application is considered as the basis of exchange, and actors and other health service providers co-create value. The SDL concept (application for users rather than goods) has moved into the health science literature. In recent times, persuasive applications, aimed at changing behaviours, have become well-known and popular in the health domain [4]. Persuasion is considered as an effective tool to support behaviour change [5]. Persuasive technology-oriented applications have been proposed as an actual method to encourage users' healthy lifestyle i.e. behaviour change and it has a potential for improving the quality of life [6]. Health and wellbeing

M. S. Haque (✉) · M. Kangas · T. Jämsä
Research Unit of Medical Imaging, Physics and Technology (MIPT), Faculty of Medicine, University of Oulu, P.O. Box. 500090014 Oulu, Finland
e-mail: md.haque@oulu.fi

A. Arman
Department of Finance, Oulu Business School, University of Oulu, Oulu, Finland

M. Kangas · T. Jämsä
Medical Research Center Oulu, University of Oulu and Oulu University Hospital, Oulu, Finland

M. S. Haque · M. Isomursu
INTERACT Research Group, Faculty of ITEE, University of Oulu, Oulu, Finland



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_36

applications that are available may not fulfil users' value expectations or preferences. Involving users and communities are necessary to improve health and wellbeing applications [7]. Ryhov Hospital and Jönköping County Council, Sweden initiated idea "Esther" to change health-oriented applications and represented the importance of care redesign that focuses on users' needs and preferences. They established and deployed VPs in improving their health and wellbeing applications [8]. Health application providers are facing difficulties in improving health and wellbeing applications e.g. cost and quality of applications [9]. Collaborating users and relevant stakeholders to create value is a challenge [10]. This leads to the question:

Can value propositions change persuasive health and wellbeing applications?

To answer the research question, we proposed four VPs and validated them through a case study with our built persuasive applications. Users (N = 23 for eating and N = 26 for physical activity) used the applications for one week to measure their eating and regular physical activity behavioural change. The results of this study could support to add an advantage to persuasive health and wellbeing applications, involving the users and their needs.

## Literature Review

### Value Co-creation

Value co-creation is the process to which health and wellbeing application providers, users, and other actors actively work together to create value for the users [11]. Co-creation of value is an interactive process to offer opportunities for the users by bringing a group of actors [12]. Value co-creation has been highlighted by SDL, which is constructed on eleven foundational premises (FP) [13]. Out of these original foundational premises, axiom FP6 (A2) and axiom FP10 (A4), was adopted in our case study since these FPs target dedicated users to determine the value (Fig. 1). Users communicate with health and wellbeing application providers at each stage of the application design and delivery [14].

Users can sense positive experience through the value creation process of health and wellbeing applications. Co-creation of value approach meets the needs of users.

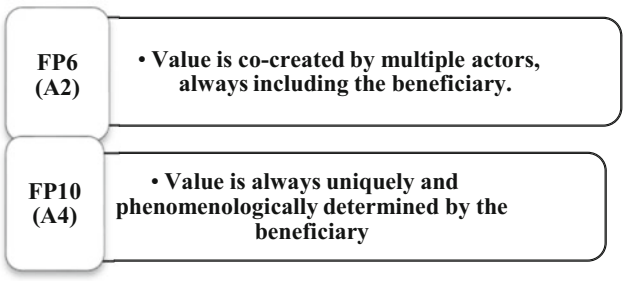



**Fig. 1** SDL FP(A2) and FP10(A4)

### Value Proposition (VP)

The role of value propositions has been explored [15]. VPs can play a vital role in communicating with the users and stakeholders through initiatives and guides [16]. Health application providers offer VPs [11]. Actors cannot deliver value but participate in creation and offering of VPs [14]. This indicates that users in health and wellbeing applications participate in creating and offering VPs.

*Related Work*: Health Technology Assessment (HTA) [17] is widely used in the EU region to seek answers to the key issues: Does technology work? For whom does it work? What is the benefit to the individual? At what cost? How does it compare to alternatives? To address these issues, HTA consults with a range of stakeholders. HTA focuses on nine domains (health problem, technology description, clinical effectiveness, safety, cost-effectiveness, ethical concerns, organizational aspects, social impacts, legal issues) [17].

*Our Approach*: HTA may be a complicated approach to assess health technology. Some of the domains are closely related to our proposed VPs, e.g. technology description, cost-effectiveness and patient satisfaction. HTA does not consider particular engagement of users or scientific and empirical measurement but practiced expert validation, whereas we aimed to use theoretical concept and empirical approach.

## Method

To identify VPs, we conducted literature search through several databases, keywords and eligibility criteria (i.e. inclusion and exclusion criteria), and verified the articles for selection for further analysis. For the inclusion criteria, we looked at the studies published in English language, accessibility with full text, and research examining the influence of values on applications focusing on health and wellbeing, and effect of values in the context of application tools to support health and wellbeing. Hence, narrative literature review approach was used and thirteen unique studies were selected for reviewing, searched for on the online repository systems: ACM Digital Library, Science Direct, Web of Science, Scopus, and EBSCO, using the combination of search terms "value", "value propositions" and "value propositions and health". The second literature search

specifically targeted papers in the context of health-based information systems. From these two types of literature searches, articles were chosen based on whether they imitate value proposition in the persuasive health and wellbeing applications. The approach was conducted through careful reading of each article. Although these articles did not prompt an exact value, we matched and searched for the terms "value" and "value proposition", went through reading articles and identified whether they are useful to users. We selected the values useful to the users. The similar value was coined out from several articles and marked as a sole VP. We marked and categorized four values into four broader classes based on their relativity. Each VP was identified by resembling users' benefit within the persuasive health and wellbeing applications (Fig. 2).

1. *Service Effectiveness and Efficiency Focusing on Increased Health and Wellbeing*: An effective and efficient application includes necessary quality application provided to the users in due time [18] and communication among users and stakeholders [19, 20]. Users are satisfied by perceiving higher value from the application. An effective and efficient application attracts potential users to the existing pool as well as retains existing ones. Stakeholders provide an effective and effective application which focuses on users' health and wellbeing and has the chance of innovation and enriching their experiences [21].
2. *Cost Reasonableness and Smart Resource Management*: Cost reasonableness refer to more benefits at lower cost. However, costs should not reduce to such a low level that affects the quality of the application, leading to ultimate impairment of the value of the application [22]. Costs should be reasonable per the quality of application provided [23]. Costs can be a combination of monetary costs, psychic costs and time costs [24]. Monetary cost is the money spent for the generation of the application. Lower monetary cost helps the user to choose economically beneficial health and wellbeing applications. Psychic cost is related to mental dissatisfaction resulting from extensive and tiresome application procedure [24]. Smart resource management means a scientifically effective and efficient use of available resources to create high-quality applications [18].
3. *Cutting-edge Technology to Furthering Information Accessibility*: Users require proper access to necessary information. Not all relevant stakeholders should have access to all information. Rather specific information should be made available to a specific type of stakeholders, e.g. users, health care professionals and specialists need information access [25, 26] and professionals need timely access to personal health records of their patients/users [20, 27]. Using cutting-edge technology, users and relevant stakeholders get timely access to information with confirmed security and privacy.
4. *Sustainability from Open Innovation Health and Wellbeing Platform*: Shifting from traditionally closed innovation to open innovation, i.e. to open source approach, has been recommended for user-oriented health and wellbeing applications [27]. Open innovation health and wellbeing platform may bring interesting outcomes, e.g. novel concepts and solutions in health and wellbeing applications that are dedicated for the users. Scholars also emphasized on the empathic support and exchange of information as important elements of communication [28]. Open innovation helps the application providers to reduce costs [29] that might encourage users to use the application. Moreover, it helps users to get the most recently innovated application in an easy and convenient way because application providers constantly change and upgrade their applications based on the needs of users [30].


**Fig. 2** VPs for persuasive health and wellbeing applications

## Case Study

To evaluate our proposed VPs, we analyzed two developed persuasive health and wellbeing applications, iCrave to promote healthy eating [31], and iGO to promote physical activity [32]. Both applications are designed with an aim that users are satisfied by perceiving higher value from the application, i.e. value in a sense of their healthy eating and physical activity (PA) promotion. The applications we

designed so that the users can have a positive experience in terms of their eating and PA. This relates to the fulfilment of the first VP. These applications are free to use for the users, i.e. the users can download them from the Android store for a smartphone or they can use the applications on a website. The applications are designed as user-friendly and simple to operate, i.e. maintaining the psychic cost. This is related to fulfilment of the second VP. We applied gamified and persuasion technique within the applications, adding some game elements such as points, badges etc. For the healthy eating application iCrave, we applied a QR code to be scanned by the phone. For the PA promotion app iGO, we applied an algorithm that allows the sensor of the smartphone to track walking steps. Users can insert their information when logging in, and they can monitor their daily activities. The admin can only access users' data once received consents from them. This reflects the fulfilment of the third VP. Lastly, to design the applications we searched the literature for the established theories which can be applied to the application and tested empirically. The healthy eating app iCrave was designed by incorporating the EI intrusion theory and iGO was designed by incorporating the SDT theory, and the design followed the user-centered design (UCD) process. The applications were tested empirically by end users to gather information in relation to the apps upgrade, usability issues etc. Users found the applications to motivate themselves in healthy eating and PA, which implies the fourth VP.

A method is recommended to involve users and relevant stakeholders for a value co-creation model [33]. Our approach in listing the VPs for persuasive health and wellbeing applications might be one successful method since the users and other actors are involved. Our approach to finding the VPs of the health and wellbeing applications is a combination of collecting information on the individual user and other relevant stakeholders' experience, quality and cost of health and wellbeing applications to increase the applications' efficacy and involving users to co-create the value. However, our proposed VPs are limited to literature and expert validation, such as HTA core model does, is not involved. Further study needs more concentration on the extract version of these four VPs to highlight their importance for different types of users. In future, expert validation of VPs should be performed.

## Conclusions

We integrated the theoretical concept of value co-creation to gauge the VPs. The VPs derived from the literature are described. We combined VPs derived from literature and integrated them by finalizing four key VPs. We performed a case study to evaluate the proposed VPs, showing VPs being a successful approach. Health application providers and other relevant stakeholders should work together to provide high quality and efficient health and wellbeing applications to users.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References


1. Porter ME (2010) What is value in healthcare? New Engl J Med 363:2477–2481
2. Lindic J, Silva CM (2011) Value proposition as a catalyst for a customer focused innovation. Manag Decis 49:1694–1708
3. Gummesson E (2008) Customer centricity: reality or a wild goose chase? Eur Bus Rev 20(4):315–330
4. Chatterjee S, Price A (2009) Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inf Assoc pp 171–178
5. O'Keefe DJ (2002) Persuasion theory and research, 2nd (edn). Sage Publications
6. Fogg BJ (2003) Persuasive technology: using computers to change what we think and do. Morgan Kaufm, San Fransisco
7. KPMG (2014) Creating new value with parents. www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/what-works/creating-new-value-with-patients 04 Nov 2016
8. Schnarr S, Alessi C, Schnarr K (2014) It's all about me: the personalization of health systems. Ivey Business School
9. Vargo SL, Maglio P, Akaka M (2008) On value and value co-creation: a service systems and service logic perspective. Eur Manag J 26(3):145–152
10. Grönroos C (2008) Service logic revisited: who creates value? and who co-creates? Eur Bus Rev 20(4)
11. Gronroos C, Vioma P (2013) Critical service logic: making sense of value creation and co-creation. Acad Mark Sci 41:133–150
12. Jacob C (1992) A power primer. Psychol Bull 112(1)
13. Vargo SL, Lusch RF (2016) Institutions and axioms: an extension and update of service-dominant logic. J Acad Mark Sci
14. Ballantyne D (2004) Dialogue and its role in the development of relationship-specific knowledge. J Bus Ind Mark 19(2):114–123
15. Frow P, Payne A (2011) A stakeholder perspective of the value proposition concept. Eur J Mark 45:233–240
16. Ballantyne D, Frow P, Varey R, Payne A (2011) Value propositions as communication practice: taking a wider view. Ind Mark Manage 40(2):202–210
17. HTA Core-Model (2014). www.hiqa.ie/HTA-GuidelinesStakeholderManagement 07 Nov 2016
18. Huerta TR, Ford EW, Peterson LT, Brigham KH (2008) Testing the hospital value proposition: an empirical analysis of efficiency and quality. Health Care Manag Rev 33(4):341–349
19. Pilon B, Crutcher TD, Lemmin-Lee S, Watters R, Wolgast KA, Arnow D (2014) The value proposition for graduate education of emerging nurse leaders: immediate benefit to organizations. Nurse Leader 12(3):81–85
20. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW (2008) A research agenda for personal health records (PHRs). J Am Med Inf 15(16)
21. Abidi SSR (2007) Healthcare knowledge sharing: purpose, practices, and prospects. Healthc Knowl Manag 67–86. Springer
22. Brown S, Saint M (2013) Value proposition for mHealth monitoring solution of diabetes. In: IST-Africa 2013 proceedings, IIMC international information management corporation

23. Krey M, Bettina H, Matthias K, Steven F (2010) IT governance, risk management and compliance in Swiss healthcare. In: Proceedings of the UKSim 12th international conference on computer modelling and simulation, pp 340–345
24. Narayan G, Nerurkar A (2006) Value-proposition of e-governance services: bridging rural-urban digital divide in developing countries. Int J Educ Dev ICT
25. Helman D, Addeo E, Walters D (2011) Ubiquity and integration in m-Health: implications for brand management. In: Sarnoff symposium
26. Manary M, Boulding W, Staelin RW, Glickman SW (2013) The patient experience and health outcomes. N Engl J Med 368:201–203
27. Lundberg N, Koch S, Hägglund M, Bolin P, Davoody N, Eltes J, Jarlman O, Perlich A, Vimarlund V, Winsnes C (2013) Mycare pathways: creating open innovation in healthcare. In: MEDINFO, pp 687–691
28. Bullinger AC, Rass M, Adamczyk S, MOeslein KM, Sohn S (2012) Open innovation in health care: analysis of an open health platform. Health Policy 105(2–3):165–175
29. Huizingh EK (2011) Open innovation: State of the art and future perspectives. Technovation 31(1):2–9
30. Reinhardt R, Bullinger AC, Gurtner S (2015) Open innovation in health care. In: Challenges and opportunities in health care management, pp 237–246
31. Hsu A, Yang J, Yilmaz Y, Haque Md, Cengiz C, Blandford A (2014) Persuasive technology for overcoming food cravings and improving snack choices. In: 2014 Proceedings of ACM CHI, Ontario, Canada
32. Haque Md, Abdullah W, Rahman S, Kangas M, Jämsä T (2016) Persuasive health and wellbeing application: a theory-driven design in promoting physical activity. In: 2016 Proceedings of IEEE MediTec, Dhaka, Bangladesh
33. Demiris G, Kneale L (2015) Informatics systems and tools to facilitate patient-centered care-coordination. Yearb Med Inf 10(1)

# Parkinson's Disease Patients Classification Based on a Motion Tracking Methodology

Eleftheria Polychronidou, Sofia Segkouli, Elias Kalamaras, Stavros Papadopoulos, Anastasios Drosou, Konstantinos Votis, Sevasti Bostantjopoulou, Zoe Katsarou, Charalambos Papaxanthis, Vassilia Hatzitaki, Panagiotis Moschonas, and Dimitrios Tzovaras

**Abstract**

This study demonstrates how a computer based methodology for tracking motor abilities of Parkinson's disease can be utilized for patient classification and assessment of the Parkinson's disease severity. The Line Test methodology evaluates the impaired voluntary movement and generates a set of features that describe the motion. A total cohort of 6 control subjects and 37 Parkinson's disease subjects were recruited and assessed for the test. During the test, a vertical line appears on the screen and the device evaluates patient's performance by producing features that correlate the motion to the last medication dosage, the line-test position, the line-test reaction time and the line-test total error. A common cohort of 24 Parkinson's disease subjects (patients that carried out the Line Test more than once) was formed to track the features alterations between repetitions in time. Results evaluation was performed in both cohorts based on information visualization methodology, optimized for the multi-objective dataset. The line-test position and the time from the last medication dosage features were proved to present the major relation to patients' group formation. Additionally, line-test reaction time and the line-test total error features proved significant between patients' performance in the common cohort. Study limitations are correlated to the size of the cohort and the time frame of the study. In general, the current practice supports further investigation into using Line Test methodology for addressing Parkinson's disease severity.

**Keywords**

Parkinson's disease • Patients classification • Motor abilities • Line test methodology • Information visualization

E. Polychronidou (✉) · S. Segkouli · E. Kalamaras · S. Papadopoulos · A. Drosou · K. Votis · P. Moschonas · D. Tzovaras
Information Technologies Institute, Centre for Research and Technology Hellas, 6th km Harilaou – Therm, Thessaloniki, Greece
e-mail: epolyc@iti.gr

S. Bostantjopoulou · Z. Katsarou
Department of Neurology, Aristotle University of Thessaloniki, Thessaloniki, Greece

C. Papaxanthis
UFR STAPS, Universit de Bourgogne, Campus Universitaire, Dijon, France

V. Hatzitaki
Department of Physical Education and Sports Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

## Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative diseases. PD affects the central nervous system and causes disorders to the motor system including bradykinesia, rest tremor, rigidity, and postural and gait impairment [1]. The disorder type characterizes the disease stage as the symptoms do not present simultaneously in patients. For example the manifestation of bradykinesia and rigidity is often in the early stages of the disease [2]. Hence diagnosis methodologies were developed based on symptom classification.



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_37

To diagnose PD a pathological analysis at autopsy is necessary. In any other case, PD diagnosis and symptoms demand a regular monitoring, and the probability of an inaccurate diagnosis is approximately 25%. The aforementioned fact underlines that there is not an objective method for diagnosis PD [3]. To this end, decision support tools developed on electronic devices. They provide the ability to monitor patient's symptoms through sensors and contribute to the early diagnosis and to the development of treatment strategies. Biomarker identification is accomplished through data collection and analysis, realized by the decision support systems.

One typical hallmark of PD is disruption in the execution of practiced skills such as handwriting. People with PD frequently have severe difficulties in coordinating of the components of a motor sequence movement. Tremor disorder affects upper extremity function and fine motor skills [4]. Previous studies have investigated "irregular paths" and movement slowness due a disruption of control processes [5]. Additionally, PD patients face difficulties in repetitive hand movements (opening/closing) or finger tapping [6]. Usually these impairments are manifested during rapid alternating movements of fingers, hand due to reduced speed and motion amplitude [7]. Repetitive rapid movements are commonly tested in upper extremities control.

The tracking of motor disruption is a promising method for identifying the level of PD severity. Currently, the evaluation of the PD patient status is based on qualitative feedback (questionnaires). As a consequent, many issues arise e.g. cumbersome test approach, sparse repetition of the test, necessity of doctor's presentation and subjectivity in evaluation. A combined method that utilizes automated motor monitoring and completion of the well-established PD classification protocols could be proved effective for optimal clinical decision. Towards this, the main objective of the current study was formed.

The aim of this work is to demonstrate how a computer based motion monitoring method can be utilized for patient classification and assess the severity of PD. The real-time and reliable analysis of patient's motor abilities was established in our previous work [8]. The electronic device translates patient's move into a set of features that depicts the movement's main characteristics. The analysis output consists of multi-parametric features that can be annotated towards PD biomarker discovery.

## Classification Protocols

Motor symptoms constitutes one of the most common clinical features of PD and is mainly characterized as parkinsonism [1]. Qualitative methods of PD severity assessment were established through the years. The Unified Parkinson's Disease Rating Scale (UPDRS) was originally developed in the 1980s and has become the most widely used clinical rating scale for Parkinson's disease [9]. In 2001, The Movement Disorder Society (MDS) revised the existing scale and proposed the MDS-sponsored UPDRS revision (MDS-UPDRS) which presents higher internal consistency in validation studies [10]. According to the rating scale there are four main categories: (a) nonmotor experiences of daily living, (b) motor experiences of daily living, (c) motor examination and (d) motor complications. Based on the final score the patients' stage is classified as normal, slight, mild, moderate or severe.

Hoehn and Yahr scale [11] and modified Hoehn and Yahr scale were initially developed to address the clinical function in PD, combining functional deficits (disability) and objective signs (impairment). The modified version was created to address the intermediate stages that were missing from the other rating systems. More explicitly, the latter system classifies the patients into seven categories: (a) Unilateral involvement only, (b) Unilateral and axial involvement, (c) Bilateral involvement without impairment of balance, (d) Mild bilateral disease with recovery on pull test, (e) Mild to moderate bilateral disease; some postural instability; physically independent, (f) Severe disability; still able to walk or stand unassisted, (g) Wheelchair bound or bedridden unless aided. These clinical scales are used as gold standard evaluation tools to characterize motor dysfunction in PD, disadvantages regarding the subjectivity in implementation derive [12].

To narrow the gap, electronic equipment in the form of mobile and tablet devices and health-monitoring devices in the form of wearable pads, wrist-bands and straps provide recordings of motion activity, electrophysiological activity and acute physiological responses that have significantly improved the understanding of Parkinson's disease [13]. Most of the existing application focusing on tremor records coming from the stability in patient's arm Few of the developed software are able to study factors such as velocity, target deviation, reaction time and minimum jerk [14]. The comparison of the measurements in time can reveal the PD progress.

## Materials and Methods

The current methodology relies on the results of the previous study of [8]. The Line Test methodology was implemented in a group of participants. Despite the previous study that focused in features related to patients motor abilities, the current Line Test resulted in a set of features that correlates patient's motion abilities to medication and comprehensive performance evaluation. Table 1 presents the differences of the feature extraction processes between the two studies.

**Table 1** Feature extraction methods

| Line test features in [8] | Current line test features |
|---|---|
| Position in pixels | Last medication dosage |
| Velocity | Line-test position |
| Acceleration | Line-test reaction time |
| Simulated muscle activation [15] | Line-test total error |

## Dataset Analysis

The total number of subjects measured for this study were 43. From the total subject number the former 37 were PD patients of various ages (mean: 62, stdev: 9.1) while the latter 6 were healthy subjects that were utilized as control cohort. PD patients were diagnosed with Idiopathic Parkinson's disease which is the most common type of PD. The age of the participants had a range from 33 to 77 years. Their PD diagnosis varies from 8 months to 18 years. A "common cohort" formed by 24 out of the 37 patients. This cohort was selected to be monitored for a longer period, in order to study their performance correlated to the disease progression. All of the participants in the common cohort remain in stable Hoehn and Yahr class between the two studies. A summary of the demographics and clinical characteristics of the study cohort is presented in Table 2. Additionally, in Table 3 the patient classification according to the established protocols is presented. The experimental procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000 and 2008.

## Multi-objective Visualization Methodology

To analyze the complicated results produced by the Line Test, the multi-objective visualization of [8] was modified based on the current features. The results of each patient were used in order to select the most significant through feature extraction methods. A feature vector that characterized each patient produced. A comprehensive method of information visualization presented the relation between patients in an interactive approach, with an ability to feature weight modifications.

## Results

Line Test resulted in a set of data that describe the performance of the 43 subjects. A more comprehensive analysis in the Line Test methodology is described in our previous work in [8]. Based on the output patients were categorized into 4 main categories. Each category corresponds to a PD condition. Specifically, green is related to healthy subjects, yellow represents patients with mild PD condition, orange represents patients with medium PD condition and red represents patients with severe PD condition. The subjects performed the Line Test with both hands and the extracted features were averaged, resulting into a set of features for each participant.

Line-test position and the time from the last medication dosage features proved to be the most significant in the classification of the patients. The validation of the aforementioned fact is obvious through the visualization tool as, by selecting to increase the weight in the line-test position the healthy patients create a discrete group related to the other subjects (1). Additionally, patients of medium and severe condition present a relation as they form a group

**Table 2** Demographics and clinical information of participants

| | Control group | PD group |
|---|---|---|
| Number of participants | 6 | 37 |
| Average age | 57.5 years | 61.9 years |
| Gender | 4 F, 2 M | 12 F, 25 M |
| Mean disease duration | 0 | 7.6 years |
| Mean PDStage | 0 | 2 |
| Dominant hand | Right | Right |
| Severed hand—left (%) | None | 37.8 |
| Severed hand—right (%) | None | 24.3 |
| Severed hand—both (%) | None | 21.6 |

**Table 3** Groups based on severity levels

| No. of patients | Hoehn and Yahr | UPDRS |
|---|---|---|
| 5 | 1 | N/A |
| 26 | 2 | N/A |
| 6 | 3 | N/A |

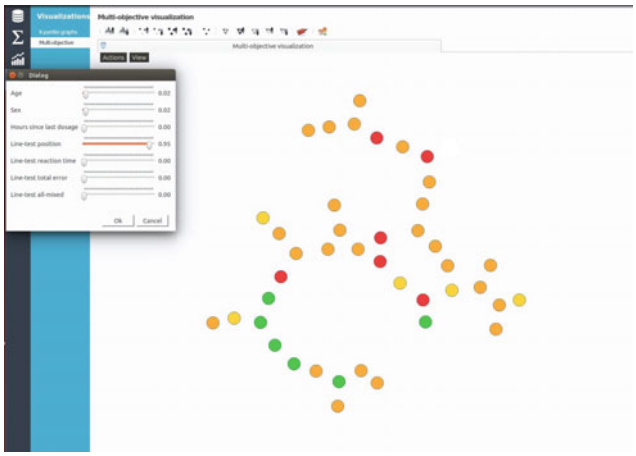


**Fig. 1** Significance of feature weight increase. By increasing the line-test position feature, separated groups of patients were formed

in space. Other features have been tested, but the discrepancy was not significant (Fig. 1).

Regarding the "common cohort", data comparison revealed that approximately 20% presented similar reaction time between the two tests and approximately 10% of the participants presented similar test total error. Additionally, similar performance in line-test position between the two measurements of the same patient was revealed by the visualization as is presented in Fig. 2.

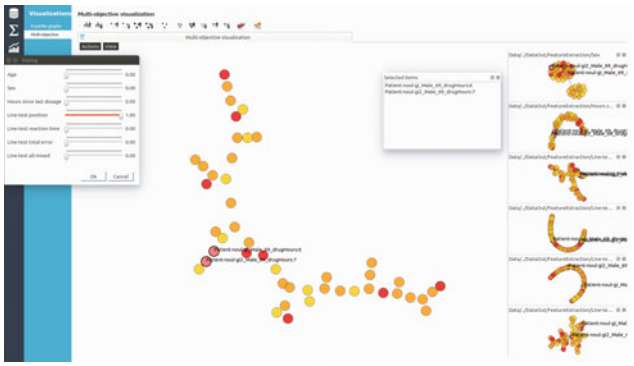


**Fig. 2** Visualization of the common cohort

## Conclusion

The current study reveal the relation of the last medication dosage, the line-test position, the line-test reaction time and the line-test total error to the severity of PD. Based on the multi-objective visualization these features affects significantly the severity of the disease. The major features that present the higher correlation to the PD progression is last medication dosage and the line-test position in the complete cohort and the line-test position in the common cohort. Hence the study proposes that the aforementioned features could reveal significant information for monitoring PD patients and translate the severity of their disease into groups of patients with similar characteristics.

Study limitations are correlated to the size of the cohort and the time frame of the study. The former limitation arises from the fact that the patients PD cohort and the control cohort is small to establish a disease rule. Based on that there a need to extend the cohort size in order to validate the study results. The later suggests that the patients should be monitored for a longer period of time in order to confirm the study suitability. This parameters are related to the common cohort of study, as a longer time frame could provide alterations in the PD stage and more comprehensive visualization results.

**Acknowledgements** This work was supported in part by the EU co-funded project "NoTremor" (EC FP7 Grant agreement no 610391).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Massano J, Bhatia KP (2012) Clinical approach to Parkinson's disease: features, diagnosis, and principles of management. Cold Spring Harbor Perspect Med 2:a008870
2. Zham P, Kumar DK, Dabnichki P, Arjunan S, Raghav S (2017) Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral. Front Neurol 8:435
3. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M (2016) Evaluation of handwriting kinematics

and pressure for differential diagnosis of Parkinson's disease. Artif Intell Med 67:39–46
4. Dibble LE, Cavanaugh JT, Earhart GM, Ellis TD, Ford MP, Foreman KB (2010) Charting the progression of disability in Parkinson disease: study protocol for a prospective longitudinal cohort study. BMC Neurol 10:110
5. Mazzoni P, Shabbott B, Cortés JC (2012) Motor control abnormalities in Parkinson's disease. Cold Spring Harbor Perspect Med 2:a009282
6. Redgrave P, Rodriguez M, Smith Y et al (2010) Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease nature reviews. Neuroscience 11:760
7. Berardelli A, Conte A, Fabbrini G et al (2012) Pathophysiology of pain and fatigue in Parkinson's disease. Parkinson Rel Disord 18: S226–S228
8. Moschonas P, Kalamaras E, Papadopoulos S et al (2016) Discovering the discriminating power in patient test features using visual analytics: a case study in Parkinson's disease. In: IFIP international conference on artificial intelligence applications and innovations. Springer, pp 600–610
9. Martínez-Martín P, Gil-Nagel A, Gracia LM, Gómez JB, Martínez-Sarriés J, Bermejo F (1994) Unified Parkinson's disease rating scale characteristics and structure. Mov Disord 9:76–83
10. Goetz CG, Tilley BC, Shaftman SR et al (2008) Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 23:2129–2170
11. Ramaker C, Marinus J, Stiggelbout AM, Van Hilten Bob J (2002) Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov Disord 17:867–876
12. Dubayova T, Krokavcova M, Nagyova I et al (2013) Type D, anxiety and depression in association with quality of life in patients with Parkinson's disease and patients with multiple sclerosis. Qual Life Res 22:1353–1360
13. Son D, Lee J, Qiao S et al (2014) Multifunctional wearable devices for diagnosis and therapy of movement disorders. Nat Nanotechnol 9:397–404
14. Farley BG, Koshland GF (2005) Training BIG to move faster: the application of the speed amplitude relation as a rehabilitation strategy for people with Parkinson's disease. Exp Brain Res 167:462–467
15. Stanev D, Moschonas P, Votis K, Tzovaras D, Moustakas K 2015 Simulation and visual analysis of neuromusculoskeletal models and data. In: IFIP international conference on artificial intelligence applications and innovations. Springer, pp 411–420

# Patient Empowerment Through Summarization of Discussion Threads on Treatments in a Patient Self-help Forum

Sourabh Dandage, Johannes Huber, Atin Janki, Uli Niemann, Ruediger Pryss, Manfred Reichert, Steve Harrison, Markku Vessala, Winfried Schlee, Thomas Probst, and Myra Spiliopoulou

**Abstract**

Self-help patient fora are widely used for information acquisition and exchange of experiences, e.g., on the effects of medical treatments for a disease. However, a new patient may have difficulties in getting a fast overview of the information inside a large forum. We propose TinnitusTreatmentMonitor, a prototype tool for the summarization and sentiment characterization of postings on medical treatments. We report on applying TinnitusTreatmentMonitor on the platform TinnitusTalk, a self-help platform for tinnitus patients.

**Keywords**

Self-help patient fora • Opinions on treatments • Discussion threads • Sentiment analysis • Medical mining

## Introduction and Related Work

Self-help internet fora allow patients to share experiences on their disease. However, a forum may contain a huge number of discussion postings and new users may have difficulties in acquiring a fast overview of the discussed contents. We propose TinnitusTreatmentMonitor, a framework that gives users a fast overview of discussions on tinnitus treatments.

Tinnitus is defined as the condition of hearing sounds without external stimulus. According to [1], tinnitus prevalence is 10-15%, while 1–2% of the patients experience a deterioration of quality of life. Insights on potential therapies are intensively discussed in social platforms like TinnitusTalk. This platform was established in March 2011 and supports discussions of treatments, exchange of experiences and support. Its two subfora on treatments contained (in July 2017) more than 35,000 postings by approximately 1100 authors [2].

The analysis of discussions in online patient platforms is intermittently done manually. For example, [3] focuses on information correctness. Most frequently, machine learning is used though, as in [4–7]. Relevant to our approach are the tasks of sentiment analysis (e.g., [5]), and opinion target extraction (e.g., [6, 7]). Our approach is partially inspired by [7], which identifies drugs discussed in opinionated postings and also detects subjectivity. Apart from these methods, TinnitusTreatmentMonitor also investigates polarity evolution for the studied treatments.

Our contribution is a proof-of-concept framework that gives an overview of discussions on medical treatments. It encompasses components for the recognition of treatments and of the dominant polarity associated with each treatment

Sourabh Dandage, Johannes Huber with have equal contribution. TinnitusTalk.com, operated by TinnitusHub.com.

S. Dandage · J. Huber (✉) · A. Janki · U. Niemann · M. Spiliopoulou (✉)
Otto-von-Guericke University, Magdeburg, Germany
e-mail: myra.spiliopouiou@ovgu.de

R. Pryss · M. Reichert
University of Ulm, Ulm, Germany

S. Harrison · M. Vessala
TinnitusHub, England, UK

W. Schlee
University Hospital Regensburg, Regensburg, Germany

T. Probst
Donau University Krems, Krems an der Donau, Austria



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_38

at each timepoint. We present the framework and used materials in the next sections. Our results and discussion are presented thereafter. We close the paper with a summary and outlook.

## Materials

For our analysis, we used the 9 TinnitusTalk subfora listed in Table 1. Thereby, the Postings column counts postings referring to treatments, the Mentions column counts sentences mentioning treatments.

## TinnitusTreatmentMonitor

In Fig. 1, we show the workflow of TinnitusTreatmentMonitor. Tasks are in dark blue and outputs in light blue. The back-end consists of the components for the tasks of data collection, identification of sentences mentioning treatments, sentence labeling, aggregation and scoring. They were implemented in Python, using the external libraries *Scrapy* for crawling, *NLTK* [8] for text processing *scikit-learn* [9] for classification and *Pandas* [10] for data aggregation. The front-end task of visualization was implemented in Javascript, using the *Aurelia* framework: it acquired inputs from an API built on top of the *Tornado* server and used *Bokeh* [11] for graph rendering. We describe all tasks hereafter.

### Subforum Crawler

In each subforum, the crawler extracts the ID, timestamp and author ID of each posting, the thread containing the posting and the number of users who clicked the button `agree' for it. It also stores the text (no images) after removing double newlines and quotations of earlier postings.

**Table 1** Materials from 9 subfora of TinnitusTalk (collected in July 2017)

| # | Subforum | Postings | Mentions | Authors |
|---|---|---|---|---|
| 1 | Alternative treatments | 15990 | 8188 | 710 |
| 2 | Collaboration space | 126 | 87 | 21 |
| 3 | Introduce yourself | 27595 | 11365 | 1854 |
| 4 | Research news | 24722 | 6776 | 640 |
| 5 | Success stories | 8110 | 2199 | 522 |
| 6 | Support (Tinnitus) | 121320 | 28893 | 2062 |
| 7 | Support (Pulsatile T.) | 2889 | 457 | 134 |
| 8 | Support (Hyperacusis) | 6965 | 2328 | 254 |
| 9 | Treatments | 19657 | 15523 | 1108 |
| | Total | 227374 | 75816 | 3950 |

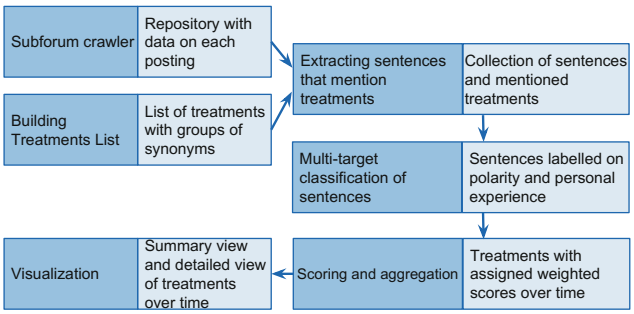



**Fig. 1** The workflow and tasks of TinnitusTreatmentMonitor

### Building a List of Treatments

This component takes a handcrafted list of treatment names as input: we extracted them from the titles of posting threads in the subfora "Treatments" and "Alternative Treatments and Research". A treatment is mentioned with multiple names: we traced those synonyms and grouped them together. During the task of labeling a sample of sentences manually (see section "TinnitusTreatmentMonitor"), we identified further names and added them to the list, completing with 149 names for 48 different treatments in total.

### Extracting Sentences that Mention Treatments

This component first splits each posting into sentences, using the "Punkt" sentence tokenizer of [12]. If a sentence *s* mentions a treatment, it is stored together with the treatment(s) it refers to. The two subsequent sentences are also inspected; if they do not mention a different treatment, they are stored with *s* and jointly considered during classification.

### Multi-target Classification of Sentences

For this task, we train and apply a *multi-target classifier* on two target variables, namely "polarity" and "personal experience". To this purpose, we model the sentences as vectors of derived features, which we compute using natural language processing tools and lexical resources, including [13–16]. For the target variable "polarity", we define the labels "positive", "negative" and "neutral`. The target variable "personal experience" specifies whether the author of the posting discusses the treatment on the basis of the own personal experience, hence we define that this target variable has the values YES and NO.

We created a random sample of 600 sentences and labeled them, splitting into a training set of 400 sentences and using the rest for testing. We use a multi-target random forest classification algorithm, the Python scikit-learn [9] implementation of [17]. Once the multi-target classifier is learned, we apply it to all sentences extracted in the previous task.

## Scoring and Aggregation

Building upon the labels assigned to each sentence, this component computes scores and aggregates them for each treatment and timeframe. In particular, a weighted score is computed for each sentence, by mapping the polarity labels "positive", "negative" and "neutral" to the *scores* 1, −1 and 0 respectively, and the personal experience labels to the *weights*1.0 (label YES) and 0.2 (label NO). Then, for each posting $x$ and treatment $y$, where $S(x,y)$ is the set of sentences in $x$ that refer to $y$, we compute:

$$pScore(x,y) = \frac{\sum_{s \in S(x,y)} score(s) \cdot weight(s)}{\sum_{s \in S(x,y)} weight(s)} \quad (1)$$

Each posting acquires a weight depending on the maximum sentence weight in the posting and the number of users clicking the "agree"-button for it:

$$pWeight(x,y) = \max_{s \in S(x,y)} weight(s) \cdot (1 + 0.5 \cdot agrees(x)) \quad (2)$$

For each time period and treatment, the weighted average of the associated posting scores is calculated, stored and presented as "treatment score" for this period. Currently, we support two time granularities, month and year.

## Visualization

The last component of TinnitusTreatmentMonitor is an interactive web application that shows how each treatment is mentioned and perceived in the forum. It consists of a "summary view" over all treatments and a "detailed view" for each treatment chosen by the user.

## Results

We run TinnitusTreatmentMonitor on TinnitusTalk. Of the postings recorded till July 2017, 41,193 (written by 3,950 users) mentioned treatments. These mentions were in 75,816 sentences, 12,979 written in the last year. We identified a negative tendency in the users' opinions: only 9 of the 48 treatments had treatment scores with a positive average.

On the testing subsample of our manually annotated sample of statements, our multi-target classifier achieved an accuracy of 60% for the polarity target, and 68% for the personal experience target. We compared to a baseline that assigns to each statement the label of the majority class: its accuracy was 57% for polarity, and 52% for personal experience, hence our model improved the baseline.

Figure 2 depicts the "summary view", which is also the start page of the front end of TinnitusTreatmentMonitor. It contains one row per treatment, consisting of three tiles: the names associated with the treatment (leftmost tile), the number of mentions in the last period (middle tile) and the treatment score (rightmost tile). This view also allows that the user filters out treatments or sorts them.

By clicking on the tiles at the right side of the ``summary view'' for some treatment, the human expert comes to the ``detailed view'' of a treatment. Figure 3 depicts one treatment. Above the graph, we see its names. To the right of the

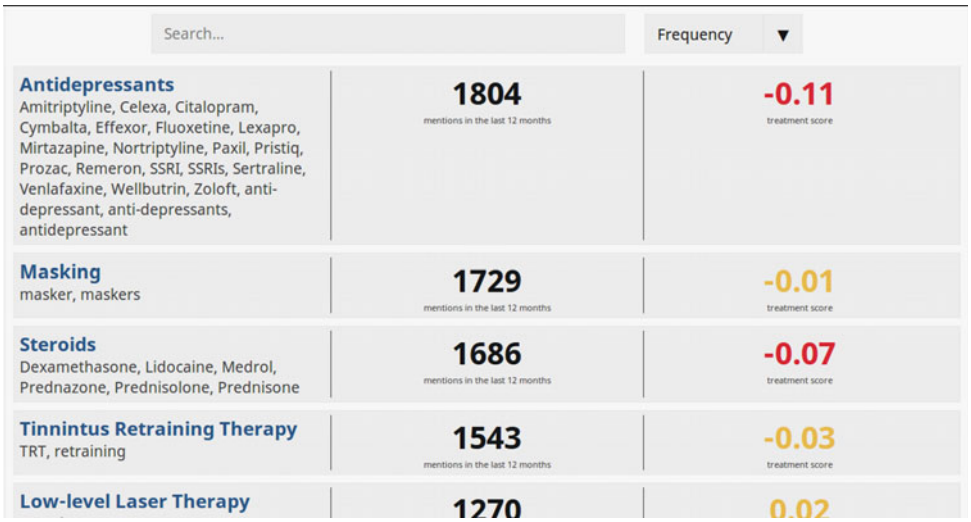



**Fig. 2** Summary view

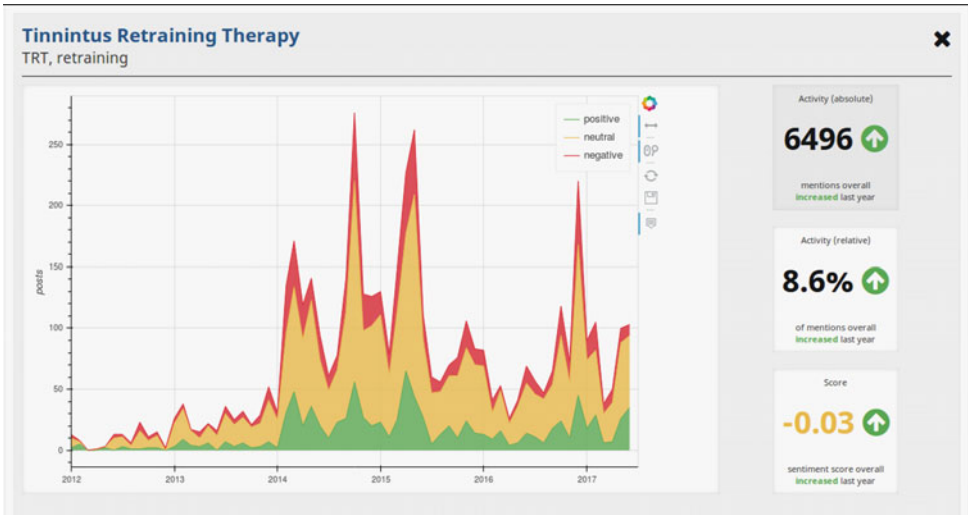



**Fig. 3** Detailed view

graph, the tiles from top to bottom display the absolute number of mentions, the percentage of mentions and the treatment score. The arrow indicates the respective trend over the last year.

The main part of the Fig. 3 offers three interactive graphs: each can be chosen by clicking at the corresponding tile. The graph chosen on Fig. 3 shows the absolute number of mentions over time, with positive, neutral or negative sentiment indicated by green, yellow, resp. red color. The ``detailed view'' also includes a choice of threads and sentences per treatment (not shown in Fig. 3). In particular, for each treatment, the posting threads are ranked on the number of the treatment's mentions in them and the top-5 threads are shown. Within each of these threads, the sentences are ranked on the target variable of personal experience (YES preferred over NO), on the number of agrees for the posting and the timestamp of the posting. The top-5 sentences are presented, with links to the original postings, so that the user can follow the links as entry points into the corresponding subfora.

## Discussion

TinnitusTreatmentMonitor aims at providing users with a comprehensive treatment overview and how these treatments are perceived over time. The presented back-end components contribute to this goal by identifying the postings mentioning treatments, classifying them and eventually visualizing them. The front-end visualization is based on the back-end and complements TinnitusTreatmentMonitor. The summary view of the visualization component (cf. Fig. 2) assists users in obtaining a first impression of all treatments. Using the leftmost column, users can figure out which names are used for the treatment. In the middle column, users see whether a treatment is subject of vivid discussions (large number of recent mentions) or it is stalled (small number). From the rightmost column, users can conclude whether the treatment is perceived positively or negatively. Hence, for new users, the summary view provides a first and compact impression of all discussed treatments. For users with a specific treatment in mind, the summary view allows them to compare its statistics with those of other treatments.

The detailed view builds upon the first overview to assist a user in understanding how treatments are perceived. By juxtaposing the area under each color in the graph of Fig. 3, the user can decide easily on whether the discussion on the specific treatment is mostly neutral (when the yellow area is predominant) or mostly sentimental (when the red or green areas are predominant). Hence, the user gets a first impression on the opinion of other patients on the treatment over time, without needing to read the postings. Hence, our proof-of-concept prototype can assist new users in acquiring insights on the discussion intensity and sentiment for all treatment categories and for each treatment separately. In a next step users will get the opportunity to evaluate the prototype in our lab and then in a sandbox of TinnitusTalk. Note that our approach revealed also drawbacks that must be further addressed. First, treatments are collected manually and, hence, we plan to use methods for Opinion Target Extraction [18, 19]. Next, the label assignments done by the classifier after training are not verified by a human expert and therefore we intend to address this task with active learning methods (see [20] on a stream of postings). Finally, the approach does not take the user's interests into account. A first step constitutes a personalized search by acquiring, e.g., keywords from the user or identifying treatments where only these keywords show up.

## Summary and Outlook

We proposed a method for the compact representation of postings in a social platform for patient self-help. Thereby, we focused on treatments and developed mechanisms, which assess the polarity of postings referring to a treatment, visualizing popularity and polarity of each treatment across the time axis, and also providing aggregated information over all treatment categories. Our next steps include the automation of treatment recognition from texts, and semi-supervised or active learning for the acquisition of human-verified polarity labels. Regarding user acceptance, we anticipate a first study with users interacting with our environment for different search tasks. The ultimate goal of our approach is to assist users in finding information. Hence, we plan to extend the approach with a keyword-based querying mechanism allowing users to learn about popularity and polarity trends for specific treatments. The information returned can be used in different ways. Of particular interest is offering it to a patient via a mobile self-help app.

**Acknowledgements** Partly, the work done by U. Niemann and M. Spiliopoulou was within the German Research Foundation project OSCAR ``Opinion Stream Classification with Ensembles and Active Learners'': U. Niemann is partially funded by OSCAR, whereas M. Spiliopoulou is project investigator.

**Compliance with Ethical Standards** The authors declare that they have no conflict of interest and no conflict with ethical standards. The social platform is public domain.

## References

1. Baguley D, McFerran D (2013) Tinnitus. Lancet 382:1600–1607
2. Thomas Probst et al (2017) Outpatient tinnitus clinic, self-help web platform, or mobile application to recruit tinnitus study samples? Front Aging Neurosci 9:113
3. Jens Türp, Harald Ohla (2012) Temporomandibular joint pain: analyzing discussions in online forums. Zeitschrift für Kraniomandibuläre Funktion 4:227–244
4. Liu X, Chen H (2013) AZDrugMiner: an information extraction system for mining patient-reported adverse drug events in online patient forums. Springer, Berlin, pp 134–150
5. Korkontzelos I et al (2016) Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. J Biomed Inf 62:148–158
6. Lorraine G et al (2012) Sentiment lexicons for health-related opinion mining. In: Proceedings of the 2nd ACM SIGHIT Int'l health informatics symposium. ACM, pp 219–226
7. Asghar DM et al (2013) Health miner: opinion extraction from user generated health reviews, vol 5, pp 279–284
8. Bird S et al (2009) Natural language processing with python. O'Reilly, 1st edn 9
9. Pedregosa F et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830
10. McKinney W (2012) Python for data analysis. O'Reilly, 1st edn
11. Bokeh Development Team (2014) Bokeh: Python library for interactive visualization
12. Kiss T, Strunk J (2006) Unsupervised multilingual sentence boundary detection. Comput Linguist 32:485–525
13. Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM 2004
14. Motoda H et al (2013) Advanced data mining and applications. Lect Notes Artif Intell 1:XXII, 588
15. Hutto CJ, Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: 8th International AAAI Conference on Weblogs and Social Media
16. Keyuan Jiang et al (2016) Construction of a personal experience tweet corpus for health surveillance. ACL 2016 2016:128
17. Louppe G (2014) Accelerating random forests in scikit-learn
18. Deng L, Wiebe J (2015) Joint prediction for entity/event-level sentiment analysis using probabilistic soft logic models. In: 2015 Conference on empirical methods in natural language processing. Association for Computational Linguistics
19. Niklas J, Gurevych I (2010) Extracting opinion targets in a single- and cross-domain setting with conditional random fields. In: Conference on empirical methods in natural language processing. Association for Computational Linguistics, pp 1035–1045
20. Zimmermann M et al (2015) Incremental active opinion learning over a stream of opinionated documents. In: WS on issues of sentiment discovery and opinion mining at KDD

# Part VIII

# e-Coaching and Patient Support for Physical Activity Promotion

# A Computer-Assisted System with Kinect Sensors and Wristband Heart Rate Monitors for Group Classes of Exercise-Based Rehabilitation

A. Triantafyllidis, D. Filos, R. Buys, J. Claes, V. Cornelissen, E. Kouidi, A. Chatzitofis, D. Zarpalas, P. Daras, I. Chouvarda, and N. Maglaveras

**Abstract**

Exercise-based rehabilitation for chronic conditions such as cardiovascular disease, diabetes, and chronic obstructive pulmonary disease, constitutes a key element in reducing patient symptoms and improving health status and quality of life. However, group exercise in rehabilitation programmes faces several challenges imposed by the diversified needs of their participants. In this direction, we propose a novel computer-assisted system enhanced with sensors such as Kinect cameras and wristband heart rate monitors, aiming to support the trainer in adapting the exercise programme on-the-fly, according to identified requirements. The proposed system design facilitates maximal tailoring of the exercise programme towards the most beneficial and enjoyable execution of exercises for patient groups. This work contributes in the design of the next-generation of computerised systems in exercise-based rehabilitation.

**Keywords**

Exercise • Rehabilitation • Computer-assisted systems • Sensors • Kinect

## Introduction

Chronic conditions such as cardiovascular disease (CVD), diabetes, and chronic obstructive pulmonary disease, require daily self-management and optimal adherence to treatment plans in order to reduce symptoms and improve health status. Exercise-based rehabilitation has been proven to be a core component of effective chronic disease management, resulting in reduced morbidity and mortality and improved cardiovascular and respiratory function as well as quality of life [1]. Group-based exercise programmes offered in the community, constitute a common treatment approach [2].

A shortcoming of the currently provided group-based exercise programmes is the lack of proper support to the trainer/instructor to allow optimal tailoring of the exercise programme. The trainer instructs exercises to a group of patients based on his/her expertise along with received visual or verbal feedback, and adapts the exercise session to the average patients' performance. However, the trainer has no access to objective quantitative information regarding the performance of individual participants in a class, thereby

A. Triantafyllidis · D. Filos · I. Chouvarda · N. Maglaveras
Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece

A. Triantafyllidis (✉) · D. Filos · I. Chouvarda · N. Maglaveras
Lab of Computing, Medical Informatics and Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: atriand@auth.gr

R. Buys · J. Claes
Department of Cardiovascular Sciences, KU Leuven, Leuven, Belgium

V. Cornelissen
Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium

E. Kouidi
Lab of Sports Medicine, Department of Physical Education and Sport Science, Aristotle University of Thessaloniki, Thessaloniki, Greece

A. Chatzitofis · D. Zarpalas · P. Daras
Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_39

limiting the potential of fine-tuning the programme to maximise its effect on the individual patients' physical status. Consequently, adherence of patients to group-based exercise programmes might be limited due to the lack of tailoring to the patients' capabilities and needs. When a patient stops exercising on a regular basis, he/she is driven to the loss of gained cardiovascular benefits and the reduction of his/her functional capacity [3]. In this context, new computer-assisted services in exercise-based rehabilitation may attract patient interest and gain wide uptake.

In this paper, we demonstrate an approach to optimal exercise programme adaptation in patient groups by means of a computerised and sensor-enhanced system. In this regard the aim of this paper is to: a) demonstrate the structure and most important parameters of current exercise programmes according to current guidelines, along with associated challenges, and b) propose a design of a novel computer-assisted and sensor-enabled system, also highlighting directions for future research.

## Exercise Programmes

### Goals and Structure

Exercise programmes including aerobic exercises supplemented with resistance exercises, are recommended by international guidelines [4], to maximise health benefits for chronic patients. More specifically, these programmes are typically performed 3 times per week in a controlled environment (e.g., specialized rehabilitation centres, gyms, etc.), under the supervision of experts in exercise-based rehabilitation (trainers). The programme typically targets a specific patient group (e.g., patients with CVD), lasts at least 30 min, and consists of a series of exercises of different intensity and difficulty.

The effectiveness of exercise-based rehabilitation programmes largely depends on four exercise characteristics, namely Frequency, Intensity, Type, and Time (FITT) [5]. As such, a main goal of the programme is to guide patients to exercise within their target heart rate zones ("Intensity" characteristic in FITT), which play a key role in effective exercise performance [4]. Beneficial heart rate zones can be formulated based on the results of a cardiopulmonary exercise test (CPET), by which both resting heart rate and peak heart rate are identified. Perceived exertion and enjoyment should also be considered when structuring exercise programmes [6, 7].

An exercise session (Fig. 1) is divided into 3 separate phases: (a) the warm-up phase which targets a gradual heart rate increase, preparing the body for the execution of more intense exercises, (b) the main phase which corresponds to the main components of an exercise session, and (c) the cool-down phase which targets at a gradual transition from exercise to rest.

The different exercises of the programme ("Type" characteristic in FITT) are chosen by the trainer according to his expertise, the health status of the average participants and received visual or verbal cues. Overall, the trainer decides on the intensity of the exercise class based on the history of workouts ("Frequency" and "Time" characteristics in FITT) by the group and the condition and ability of the participants. Optimally, the trainer proposes a variety of exercises involving all body parts (legs, arms, trunk, etc.), with different levels of intensity and difficulty. For example, an exercise with low intensity (e.g., "walking on the spot") could be chosen during warm-up/cool-down, or after a series of exercises with high intensity (e.g., "jumping"). The trainer can also change the level of intensity or difficulty (e.g., exercises requiring balance or coordination of different body parts), if he/she notices discomfort, exhaustion, difficulties or dislike to be present in one or more participants.

**Fig. 1** Patient group performing exercises instructed by a trainer in an exercise-based rehabilitation programme

### Challenges

In current exercise-based rehabilitation in a group setting, a number of barriers preclude its effectiveness in optimally improving patient physical status and health condition. In brief some of these barriers are:

(a) *Lack of personalization*: Due to the fact that the programme is targeted at a patient group, there is an inherent problem of programme individualisation according to one's specific needs. In this respect, the trainer may instruct exercises which are not performed in the correct form by all participants due to fitness level, motor impairments, health problems, etc. Following the same rationale, an instructed exercise might be disliked or might be causing overexertion in some participants, depending on personal characteristics.

(b) *Lack of quantification*: The trainer mainly relies on his expertise and experience to adapt an exercise programme without support from important quantified feedback during an exercise class. In this respect, the trainer has no sign of whether a patient's optimal heart rate zone is reached, or whether he/she is adhering to the exercise form, and enjoys the training session.

(c) *Group statistics*: The trainer has no feedback on the overall group's performance (e.g., heart rate dynamics, motion accuracy, perceived exertion, etc.) and progress during an exercise session. This limits the possibility of taking the proper actions in terms of beneficial adaptations of an exercise class according to the FITT principles.

## Proposed System Design

We propose a novel system (Fig. 2), based on the experience gained from the PATHway system for exercise-based cardiac rehabilitation at home [8, 9], which can help to overcome current barriers in exercise-based rehabilitation. More specifically, we propose a system consisting of the following components:

(a) *Motion tracking*: Kinect cameras target at capturing motion during exercise, which has been proven to be effective in the context of rehabilitation [10]. In this way motion accuracy can be evaluated via appropriately applied algorithms [11], which is important from kinesiology viewpoint for the safety and health of the participants. A Kinect sensor points to the trainer to recognise the selected exercise (based on a pool of available exercises in the system). Following exercise recognition, motion accuracy is detected for all group participants with separate Kinect sensors. Then, an overall average motion accuracy score for the group can be calculated for each individual exercise and displayed to a projector screen. The trainer can use this information to evaluate whether or not it is needed to emphasize on exercise execution by providing key instruction points or adapting the speed of the execution, especially for those who experience difficulties.

(b) *Heart rate monitoring*: Heart rate is monitored through Bluetooth-enabled wristband heart rate monitors worn by the participants. The accuracy of such devices has been reported in previous work [12]. The target heart rate zones for each individual are identified during system initialization before an exercise session starts, and based on CPET. During exercise, the trainer is enabled to view in the projector screen whether participants have reached their target heart rate as an overall percentage. This allows the trainer for example to change the exercise intensity when the percentage is low (e.g., less than 50%).

(c) *Exercise type monitoring*: The trainer is enabled to view the duration of the type of exercises already performed (aerobic, resistance, exercises for different muscle groups), and can decide on the type of exercise which patients should focus on.

(d) *Exertion and enjoyment monitoring*: The addition of push buttons to the wristband heart rate monitor can allow patients to easily indicate levels of exertion and/or enjoyment. The participant feedback in terms of perceived exertion and enjoyment can subsequently be viewed in the projector screen and the trainer can adapt the exercise session accordingly.

The proposed system is easy to set up. Each Kinect sensor is connected (via USB) with a laptop computer to enable motion recognition and heart rate monitoring in each

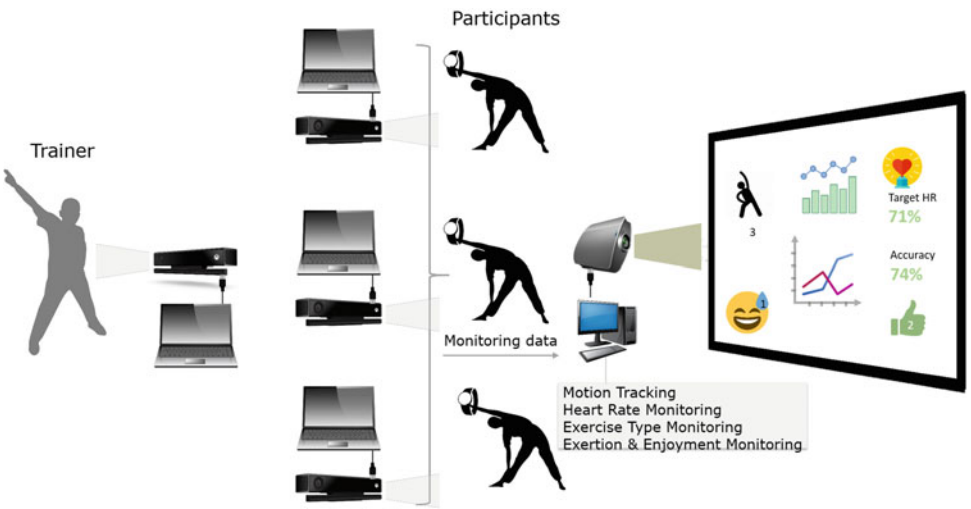



**Fig. 2** System architecture utilizing Kinect sensors, wristband heart rate monitors, laptops, a projector screen and a central computer to process monitored data

participant. The Kinect should be placed in a distance 2 m - 3 m from the patient. The heart rate monitors connect via Bluetooth communication technology with the laptop computers. Motion and heart rate data from each patient are analyzed in terms of accuracy and intensity. Finally, all monitored data (heart rate, motion accuracy, self-reports of enjoyment and exertion) are transmitted via a LAN or Internet connection to a central computer which is responsible for data aggregation, processing, and display to a projector screen.

## Discussion

Performance monitoring during exercise training with computer-assisted systems has been the focus of research presented e.g., in [13, 14], but only motion accuracy or steps were received as objective indicators of performance. In other systems, performance indicators included target heart rate, e.g., in [15], but only a single type of activity (i.e., cycling) was monitored. In [16], both accelerometry and heart rate were used to quantify personalised performance during exercise, but the focus was on outdoor activities and motion capturing was not applied. Therefore, to the authors' knowledge, the presented system which combines heart rate monitoring with motion accuracy tracking for group exercise-based rehabilitation is the first of its kind.

Future work could focus on the development of computer-assisted systems which are not only able to quantify the actual performance, but also able to predict the performance of the patients in an exercise session, in order to suggest adaptations to the exercise class and assist the trainer with decision making for programme optimisation. The usability of the proposed system for both patients and trainers needs to be investigated in future studies [17]. The feasibility and effectiveness of the described system could be explored in a clinical trial, in comparison with "traditional" group exercise-based rehabilitation.

## Conclusions

We presented the design of a computer-assisted system for optimization of group exercise-based rehabilitation classes in a controlled environment (e.g., a gym) under supervision of a trainer. The system will empower trainers to provide more effective exercise programmes according to patient needs, based on quantified feedback on exercise performance and progress. As such, this work contributes to the development of the next-generation of computer-assisted systems targeting at better health management in the community.

**Acknowledgements** Author AT was supported by the "IKY fellowships of excellence for postgraduate studies in Greece—SIEMENS program". Authors DF, JC, RB, VC, AC, DZ, PD, IC, and NM were supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation Action under Grant Agreement no. 643491, 'PATHway: Technology enabled behavioural change as a pathway towards better self-management of CVD'.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Sagar VA, Davies EJ, Briscoe S et al (2015) Exercise-based rehabilitation for heart failure: systematic review and meta-analysis. Open Hear 2:e000163. https://doi.org/10.1136/openhrt-2014-000163
2. Kouidi E, Karagiannis V, Grekas D et al (2010) Depression, heart rate variability, and exercise training in dialysis patients. Eur J Cardiovasc Prev Rehabil 17:160–167. https://doi.org/10.1097/HJR.0b013e32833188c4
3. Nocon M, Hiemann T, Müller-Riemenschneider F et al (2008) Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. Eur J Cardiovasc Prev Rehabil 15:239–246. https://doi.org/10.1097/HJR.0b013e3282f55e09
4. Vanhees L, Geladas N, Hansen D et al (2012) Importance of characteristics and modalities of physical activity and exercise in the management of cardiovascular health in individuals with cardiovascular risk factors: recommendations from the EACPR. Part II. Eur J Prev Cardiol 19:1005–33. https://doi.org/10.1177/1741826711430926
5. Thompson PD, Arena R, Riebe D, Pescatello LS (2013) ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription, 9th (edn). Curr Sports Med Rep 12:215–217. https://doi.org/10.1249/JSR.0b013e31829a68cf
6. Raedeke TD (2007) The relationship between enjoyment and affective responses to exercise. J Appl Sport Psychol 19:105–115. https://doi.org/10.1080/10413200601113638
7. Borg G, Hassmén P, Lagerström M (1987) Perceived exertion related to heart rate and blood lactate during arm and leg exercise. Eur J Appl Physiol Occup Physiol 56:679–685. https://doi.org/10.1007/BF00424810
8. Claes J, Buys R, Woods C et al (2017) PATHway I: Design and rationale for the investigation of the feasibility, clinical effectiveness and cost-effectiveness of a technology-enabled cardiac rehabilitation platform. BMJ Open. https://doi.org/10.1136/bmjopen-2017-016781
9. Filos D, Triantafyllidis A, Chouvarda I et al (2016) PATHway: decision support in exercise programmes for cardiac rehabilitation. Stud Health Technol Inform 224:40–45
10. Chang K-M, Liu S-H (2011) Wireless portable electrocardiogram and a tri-axis accelerometer implementation and application on sleep activity monitoring. Telemed J E Health 17:177–184. https://doi.org/10.1089/tmj.2010.0078
11. Chatzitofis A, Zarpalas D, Filos D et al (2017) Technological module for unsupervised, personalized cardiac rehabilitation exercising. In: 2017 IEEE 41st annual computer software and applications conference (COMPSAC). https://doi.org/10.1109/COMPSAC.2017.230

12. Claes J, Buys R, Avila A et al (2017) Validity of heart rate measurements by the Garmin Forerunner 225 at different walking intensities. J Med Eng Technol 41:480–485. https://doi.org/10.1080/03091902.2017.1333166
13. Kranz M, Möller A, Hammerla N et al (2013) The mobile fitness coach: towards individualized skill assessment using personalized mobile devices. Pervasive Mob Comput 9:203–215. https://doi.org/10.1016/j.pmcj.2012.06.002
14. Compernolle S, Vandelanotte C, Cardon G et al (2015) Effectiveness of a web-based, computer-tailored, pedometer-based physical activity intervention for adults: a cluster randomized controlled trial. J Med Internet Res 17:e38. https://doi.org/10.2196/jmir.3402
15. Peng H-T, Song C-Y (2015) Predictors of treatment response to strengthening and stretching exercises for patellofemoral pain: An examination of patellar alignment. Knee 22:494–498. https://doi.org/10.1016/j.knee.2014.10.012
16. Buttussi F, Chittaro L (2008) MOPET: A context-aware and user-adaptive wearable system for fitness training. Artif Intell Med 42:153–163. https://doi.org/10.1016/j.artmed.2007.11.004
17. Triantafyllidis AK, Koutkias VG, Chouvarda I, Maglaveras N (2014) Development and usability of a personalized sensor-based system for pervasive healthcare. In: 2014 36th Annual international conference of the IEEE engineering in medicine and biology society EMBC. https://doi.org/10.1109/EMBC.2014.6945146

# A Computerized System for Real-Time Exercise Performance Monitoring and e-Coaching Using Motion Capture Data

Anargyros Chatzitofis, Dimitris Zarpalas, and Petros Daras

**Abstract**

A lack of exercise and physical activity is one of the main health-risk behaviors, causing chronic diseases. This paper proposes a computerized system for monitoring exercise performance and e-coaching. The aim of the system is to increase users' physical activity and fitness levels, improve the effectiveness of exercise-based rehabilitation and training, and subsequently motivate people to become more active. Capable of acquiring and fusing motion capture data from different modalities (Kinect and IMUs), depending on the physical exercise intricacy to be captured and evaluated, the proposed system performs real-time anthropometric measurements and analysis. The main challenges of physical exercise performance monitoring are also addressed. Lying-down physical exercises or exercises with caregiver or trainer support can be captured, monitored and evaluated. Exercise repetition detection and evaluation, and e-coaching of the subjects through comprehensive semantic feedback are performed in real-time.

**Keywords**

Exercise-based rehabilitation • Exercise performance monitoring • e-Coaching • Motion capture • Exercise evaluation

## Introduction

Nowadays, chronic diseases have become the center of attention in public health worldwide, causing more than 36 million deaths a year and consuming great amounts of resources in healthcare costs [1]. From 2012 to 2014, 117 million people in the United States (US)—about half of all adults—had one or more chronic health conditions [2]. One of the main healthy risk behaviors that cause chronic diseases is the lack of exercise or physical activity. In 2015, in the US, 50% of adults did not meet recommendations for aerobic physical activity, while 79% of them did not meet recommendations for both aerobic and muscle-strengthening physical activity [3]. However, increasing physical activity can decrease the risk of disease, while physical exercise-based rehabilitation can be effective for many chronic diseases such as cardiac diseases (e.g. Coronary Heart Disease) or movement disorders (e.g. Parkinson's Disease [4]).

Group-based, supervised exercise programmes in gyms or rehabilitation centers constitute the main approach for physical exercise. In these programmes, the trainers instruct the subjects on how to exercise correctly, adapting the exercise sessions based on their experience and the feedback they retrieve. On the other hand, patients with movement difficulties (e.g. with sub-acute stroke, Parkinson's disease, etc.) often seek private rehabilitation programmes with caregiver support [5]. However, in both cases, the main disadvantage is the lack of information regarding the quality of exercise execution and the quantitative measurements and reports regarding the training or rehabilitation progress

A. Chatzitofis (✉) · D. Zarpalas · P. Daras
Centre for Research and Technology Hellas, Information Technologies Institute, 6th km Charilaou-Thermi, Thessaloniki, Greece
e-mail: tofis@iti.gr



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_40

(e.g. anthropometric body part movements, number of exercise repetitions, exercise execution accuracy, etc.).

In this context, also exploiting the recent advances in visual computing and wearable, wireless sensor technologies, the development of more sophisticated, technology-enabled fitness and rehabilitation solutions has been enforced. Ofli et al. [6] propose an interactive exercise coaching system for older adults, guiding them through a series of video exercises, tracking their movements and providing them with real-time feedback. In [7], ExerciseCheck, a remote monitoring and evaluation platform for home-based physical rehabilitation is proposed by Saraee et al., aiming to give the patients feedback on how-to adjust their movements in order to achieve higher accuracy of exercise performance.

In this paper, a computerized system for exercise performance monitoring and e-coaching is proposed, using motion capture data from several modalities. The proposed system addresses some of the main challenges in monitoring, evaluating and coaching physical exercises. In particular:

- Lying-down exercises can be captured, monitored and evaluated.
- Exercises where one or more caregivers or trainers are present can be captured, monitored and evaluated.
- Repetition detection, accurate evaluation and semantic feedback are provided.

The paper is organized as follows. Section "Proposed System" describes the architecture, the implemented and integrated technological modules and the related functionalities. In section "Discussion", the effectiveness and the novelty of the proposed system are discussed. Conclusion and future work are finally given in Sect. IV.

## Proposed System

### Architecture

A computerized system for exercise performance monitoring and e-coaching is proposed. The system, which supports multimodal real-time motion capture, instructs the users how to appropriately execute their physical exercise programme with or without supervision. The system is available for Windows, while the required equipment consists of a MS Kinect v2 sensor (Kinect), a set of inertial measurement units (IMUs) (optional) and a Windows PC/laptop, meeting the minimum system requirements for using Kinect.

The system consists of two main technological modules which interact with each other; the motion capture and the exercise performance evaluation (Fig. 1). The motion capture module is capable of motion capture data acquisition using either solely Kinect or Kinect and IMUs. These data

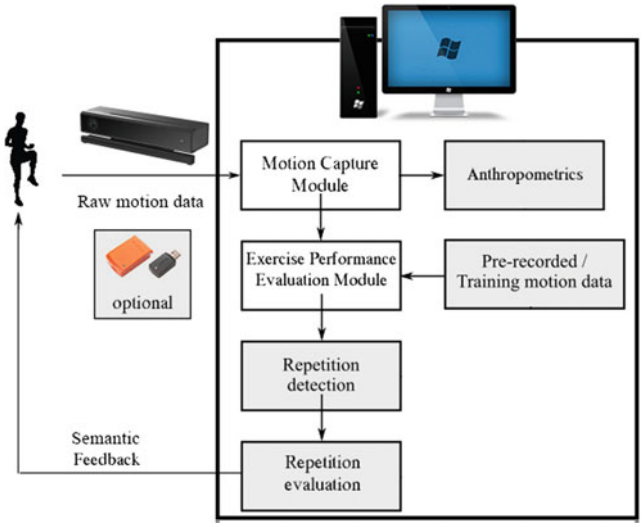



**Fig. 1** System architecture and pipeline. White containers indicate the modules of the system

(a) are used as input for the exercise performance evaluation module and (b) are processed and analyzed for anthropometric measurements. Regarding the exercise performance evaluation module, pre-recorded training motion data have been considered as reference and used to train the module. Analyzing then the captured motion data, the time interval of the performed repetition is detected. Finally, the detected and the reference exercise repetitions are compared, resulting in semantic feedback.

### Motion Capture Module

Several training and rehabilitation monitoring systems integrate real-time motion capture solutions [8–10]. In the proposed system, a novel motion capture module has been developed and integrated, enabling motion capture and antropometric measurements.

#### Motion Capture

As mentioned above, the proposed system has been designed to function with the aid of Kinect and an arbitrary number of IMUs to minimize the associated equipment cost and complexity, depending on the exercise intricacy to be captured. Except from the number of the inertial sensors, their placement and usage can also vary, depending on the use case and the type of physical exercise. For instance, in physical exercises that target only the upper body, the inertial sensors can be placed just on the trunk and the arms.

User's raw 3D body information is provided per frame using Kinect, including global 3D positions and orientations of 25 body joints. If a higher level of motion capture accuracy is required, 1 to $N$ inertial sensors can be placed on

the subject's body, where $N$ represents the desired number of sensors. Fusing raw motion capture data with inertial data (acceleration, angular velocity, magnetic field data), the motion capture accuracy for the end-joints of the bones on which the inertial sensors have been placed is improved. The Kinect-IMU fusion method has been described in [11].

## Anthropometric Measurements

A significant component for efficient exercise performance monitoring is the real-time anthropometric measurement and analysis [12]. Analyzing the provided motion capture data allows for robust extraction of anthropometrics. Given the 3D positions of the joints, the planes of human body motion (i.e., saggital, frontal and transverse) are defined [13], calculating the orientation of the body. In particular, the upper and lower body orientations can be separately estimated in the proposed system, enabling holistic and separate or partial motion analysis. To this end, for the upper body, taking into account the 3D positions of the middle of the spine $\mathbf{p}_s$, the neck $\mathbf{p}_n$ and the left shoulder $\mathbf{p}_{ls}$, the orientation $\mathbf{R}_{ub}$ is calculated. Correspondingly, the positions of the spine base $\mathbf{p}_{sb}$, the middle of the spine $\mathbf{p}_s$ and the left hip $\mathbf{p}_{lh}$, give the orientation of the lower body, $\mathbf{R}_{lb}$. More specifically, let $\mathbf{v}'_{x1}$ be the normalized vector from the neck to the left shoulder position ($\mathbf{v}'_{x1} = \mathbf{p}_{ls} - \mathbf{p}_n/|\mathbf{p}_{ls} - \mathbf{p}_n|$) and $\mathbf{v}_{y1}$ be the normalized vector from the neck to the middle of spine position ($\mathbf{v}_{y1} = \mathbf{p}_s - \mathbf{p}_n/|\mathbf{p}_s - \mathbf{p}_n|$). The cross product between $\mathbf{v}_{y1}$ and $\mathbf{v}'_{x1}$ gives $\mathbf{v}_{z1}$ ($\mathbf{v}_{z1} = \mathbf{v}_{y1} \times \mathbf{v}'_{x1}$), while, then, the cross product between $\mathbf{v}_{z1}$ and $\mathbf{v}_{y1}$ gives $\mathbf{v}_{x1}$ ($\mathbf{v}_{x1} = \mathbf{v}_{z1} \times \mathbf{v}_{y1}$). As a result, the rotation matrix $\mathbf{R}_{ub}$ of the upper body is defined by the estimated normalized vectors ($\mathbf{R}_{ub} = [\mathbf{v}_{x1}\mathbf{v}_{y1}\mathbf{v}_{z1}]^T$). In holistic body analysis, where the body is not separately analyzed, $\mathbf{R}_{ub}$ is considered the orientation of the whole body. The rotation matrix $\mathbf{R}_{lb}$ the lower body is correspondingly defined using $\mathbf{p}_{sb}$, $\mathbf{p}_s$ and $\mathbf{p}_{lh}$.

Using the motion capture data and the estimated planes of motion, anthropometric angles (e.g., flexion, extension, abduction, etc.) and other anthropometric parameters (e.g., body height, trunk length and width, leg and arm length, etc.) can be given.

## Exercise Performance Evaluation Module

This module is based on the recognition and detection of a priori known exercise instances within a sequence of motion data. Subsequently, these instances can be evaluated, providing feedback of whether the right exercise has been performed, how accurate the exercise performance was and how to improve the performance accuracy.

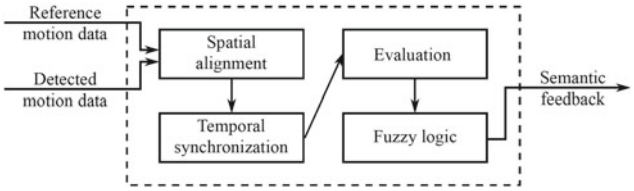



**Fig. 2** Exercise repetition evaluation pipeline

## Exercise Repetition Detection

The exercise repetition detection has been developed for (a) recognizing that a repetition of a specific exercise was performed and (b) detecting its time interval. To efficiently serve repetition detection, even in complex physical exercises (e.g. bilateral lunge stretch with trunk rotation), gesture recognition algorithms have been extended, recognizing sequences of key-frames. Details regarding the repetition detection method have been described in [11].

## Exercise Repetition Evaluation

After the detection of a performed exercise repetition, the motion capture data get segmented and labeled, allowing for exercise evaluation to be applied. To this end, a motion data processing pipeline is proposed, utilizing the segmented motion capture data and a set of sequences of this specific exercise. This set is obtained by the training set and is used as reference, resulting in semantic feedback (Fig. 2).

It is worth noting that the motion data, before being analyzed by the evaluation components, are pre-processed. At first, applying body bone normalization prevents errors due to body structure differences. Subsequently, spatio-temporal alignment is performed. Spatial alignment is achieved by estimating the rotational offset between the body rotations. Then, by applying multivariate Dynamic Time Warping (m-DTW) on the joint 3D positions, the sequences are temporally aligned. The 3D joint position and linear velocity errors are calculated and normalized, enabling statistical joint error analysis. The DTW and error statistical analysis are given as input to a fuzzy logic engine, resulting in semantic feedback.

# Discussion

The main challenges addressed by the proposed system, as described in section "Introduction", are discussed in this section. The achievements of the motion capture module are primarily based on the capability of the system to fuse Kinect and IMUs, addressing Kinect low frame rate and one-side view limitations as well as the difficulty to efficiently estimate 3D positions using inertial data.

## Lying Down Exercises

A variety of exercise groups, including abdominal, stretching and many rehabilitation exercises, are executed while the subject lies down. In this case, Kinect fails to capture the human body movements, due to self-occlusions and the fact that Kinect functions properly when the subject is standing or seated. The proposed system goes beyond Kinect tracking by placing inertial sensors on the bones of the subject and applying Kinect-IMU fusion. In Fig. 3a, where the subject lies on a table, 3 IMUs are enough to accurately capture the arm stretching.

## Exercises with Caregiver or Trainer Support

In several exercise-based rehabilitation or training programmes, especially for patients with movement difficulties, exercises with caregiver support are included. In Fig. 3b, the subject is supported by a caregiver in order to perform a specific rehabilitation exercise. In such cases, Kinect cannot accurately capture the subject's motion due to occlusion from the caregiver's body. For this purpose, 9 inertial sensors have been placed on the subject's body to capture the movements.

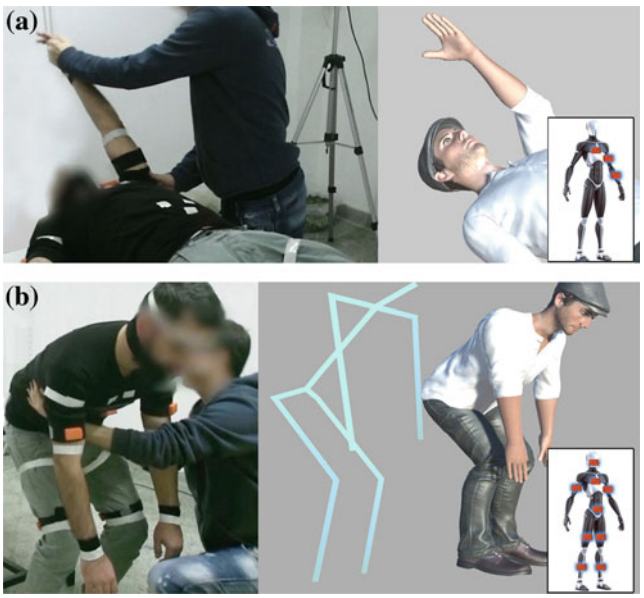



**Fig. 3** **a** Lying down physical exercise. Kinect-IMU fusion using 3 sensors; 1 on the left upper arm, 1 on the left forearm and 1 on the trunk. **b** Exercise with caregiver support using 9 sensors; 2 on the upper arms, 2 on the trunk, 1 on the head, 2 on the thighs and 2 on the calves

## Exercise Evaluation and Semantic Feedback

With the aim to include challenging physical exercises, further techniques were implemented during the system development. For some continuous exercises, for instance, two repetitions, each performed by different limbs, are combined and considered as one (e.g. a walking on the spot repetition is considered completed when the subject performs two steps).

Finally, the semantic feedback given by the proposed system uses pre-defined templates in order to inform the users regarding their performance (e.g. "HIGH score! The highest POSITION error is detected for the LEFT WRIST… move your LEFT WRIST…", here, the words with capital letters are being selected by the outcome of the fuzzy logic).

## Conclusion and Future Work

Such systems show considerable promise, advancing the effectiveness of exercise-based training and rehabilitation, motivating people to become more active and improving the quality of their lives. The next step for the system is to overcome current difficulties relating to the monitoring of subjects exercising with a trainer or caregiver, due to the possibility of a IMU being moved, which increases the risk for erroneous motion capture. For this purpose, research is being conducted using reflective materials and ToF depth-sensing devices in order to increase the motion capture stability and quality.

**Acknowledgements** The authors were supported by the EU Horizon 2020 Framework Programme, for Research and Innovation Action under Grant Agreement no. 643491, 'PATHway H2020'.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Durstine JL, Gordon B, Wang Z, Luo X (2013) Chronic disease and the link to physical activity. J Sport Health Sci 2:3–11
2. Ward BW, Schiller JS, Goodman RA, (2014) Peer reviewed: multiple chronic conditions among us adults: a (2012) update. Prev Chronic Dis 11
3. Disease Prevention Office, Promotion Health, others (2017) US Department of Health and, Human Services: Healthy people 2020, Office of Disease Prevention and Health Promotion, US Department of Health and Human Services
4. Paolo B, Eliana T, Pierpaolo S, Attilio P (2014) Effects of physical activity in Parkinson's disease: a new tool for rehabilitation. World J Methodol 4:133

5. Harris JE, Eng JJ, Miller WC, Dawson AS (2016) The role of caregiver involvement in upper-limb treatment in individuals with subacute stroke. Phys Ther 90:1302–1310
6. Ofli F, Kurillo G, Obdržálek Š, Bajcsy R, Jimison HB, Pavel M (2016) Design and evaluation of an interactive exercise coaching system for older adults: lessons learned. IEEE J Biomed Health Inf 20:201–212
7. Saraee E, Singh S, Hendron K et al (2017) ExerciseCheck: remote monitoring and evaluation platform for home based physical therapy. In: Proceedings of the 10th international conference on PErvasive technologies related to assistive environments, ACM
8. Chatzitofis A, Vretos N, Zarpalas D, Daras P (2013) Three-dimensional monitoring of weightlifting for computer assisted training. In: Proceedings of the virtual reality international conference: Laval virtual: 3, ACM
9. Cannell J, Jovic E, Rathjen A et al (2017) The efficacy of interactive, motion capture-based rehabilitation on functional outcomes in an inpatient stroke population: a randomized controlled trial. Clin Rehabil. https://doi.org/10.1177/0269215517720790
10. Knippenberg E, Verbrugghe J, Lamers I, Palmaers S, Timmermans A, Spooren A (2017) Markerless motion capture systems as training device in neurological rehabilitation: a systematic review of their use, application, target population and efficacy. J Neuroeng Rehabil 14:61
11. Chatzitofis A, Zarpalas D, Filos D et al (2017) Technological module for unsupervised, personalized cardiac rehabilitation exercising. In: 2017 IEEE 41st annual computer software and applications conference (COMPSAC), vol 2. IEEE, pp 125–130
12. Norton K, Olds T (1996) Anthropometrica: a textbook of body measurement for sports and health courses. UNSW Press
13. Day BL, Steiger MJ, Thompson PD, Marsden CD (1993) Effect of vision and stance width on human body motion when standing: implications for afferent control of lateral sway. J Physiol

# Design of a Fully Automated Service to Generate an Individualized Exercise Rehabilitation Program for Adults with Congenital Heart Disease

R. Buys and V. A. Cornelissen

**Abstract**

Physical activity is key in the prevention of cardiovascular diseases and as such should also be targeted in the long-term care of patients with congenital heart disease [CHD]. Mounting evidence, derived from small proof of concept studies, shows that patients with CHD can safely exercise and are able to increase their exercise capacity. Yet, the implementation of physical activity and exercise programs is challenging in patients with CHD given the need for a highly individualized approach. Therefore, efforts are needed to further tailor the current exercise prescription recommendations in order to obtain a truly personalized and most likely more effective exercise rehabilitation program in terms of adherence and health benefits. Here we propose the design of a fully automated service for the generation of an individualized, patient tailored exercise rehabilitation program for adults with CHD. This computer based information system will consider individual patient data to ensure safety, effectiveness as well as attractiveness in order to automatically propose an exercise prescription. This prescription will consist of the short and mid-term exercise goal for the patient to target. Moreover, the system will propose a structured exercise program in order to gradually work toward achieving the goals. The system will support CHD specialists in prescribing detailed, understandable and effective exercise and physical activity programs according to patient needs and wants, based on quantified information on cardiac status, exercise performance, physical activity level and patient preferences. As such, this work contributes to the development of computer-assisted systems targeting at personalized medicine and ultimately better health outcomes in CHD.

**Keywords**

Exercise program • Rehabilitation • Congenital heart disease • Computer based information system

## Introduction

Approximately one percent of all newborn baby's has a congenital heart defect [CHD] [1]. Nowadays, almost 90% of these patients survives into adulthood [2], owing to the developments in medicine the last decades. Consequently, the number of adults with CHD is estimated to have risen to at least 1.2 million in Europe [3].

According to current guidelines, patients with CHD need regular cardiac follow-up, as long-term morbidity is high [4]. Moreover, the risk of developing ischemic heart disease is increased in patients with CHD in comparison to age-matched peers. In the context of this growing patient population with high healthcare needs, preventive strategies are an important public health challenge, especially since an increasing prevalence of obesity and diabetes can also be

R. Buys (✉)
Department of Cardiovascular Sciences, KU Leuven, Herestraat 49, Bus 1501, Leuven, Belgium
e-mail: roselien.buys@kuleuven.be

R. Buys · V. A. Cornelissen
Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_41

noted here [5]. Physical activity is key in the prevention of cardiovascular diseases and as such should also be targeted in the long-term care of patients with CHD.

Mounting evidence, derived from small proof of concept studies, shows that patients with CHD can safely exercise and are able to increase their exercise capacity [6]. Yet, the implementation of physical activity and exercise programs is challenging in patients with CHD. That is, the wide spectrum of CHD precludes the prescription of a more or less generalized exercise program and warrants a highly individualized approach. Moreover, adult patients with CHD report that they are concerned about physical activity in their daily life [7] and that are willing to engage in exercise programs [8, 9]. At the same time they express their fears and uncertainties with regard to which type and intensity of exercise/sports is safe and consequently often remain inactive. Given the importance of a physically active lifestyle, the difficulties associated with its implementation and the fact that the majority of patients remains inactive, it was suggested that exercise-based interventions should become part of routine follow-up [6, 10]. To this aim, Budts et al. developed an individualized exercise prescription algorithm based on the absence or presence of certain key hemodynamic and electrophysiological elements rather than on the different underlying heart defects as was common in previous guidelines [11]. This new approach was a first step towards patient tailoring of an exercise program. Despite this, the outcome still remains a rather general advice regarding the intensity of static and dynamic exercise and both the patient and the healthcare provider remain with questions regarding the optimal training program. In the era of integrated care and personalized medicine, more efforts are needed to further tailor the current exercise prescription algorithm in order to obtain a truly personalized and most likely more effective exercise rehabilitation program in terms of adherence and health benefits.

Here we propose the design of a fully automated service for the generation of an individualized, patient tailored exercise rehabilitation program for adults with CHD. In section "Content" of this paper, we summarize the exercise content of the system, as well as the data that needs to be fed into the system to allow personalization of the exercise program. In section "Output", we summarize the output that will be generated by the system, through the processing of the data. Finally, in section "Discussion and Future Challenges", we speculate on the potential of the proposed system along with directions for future research.

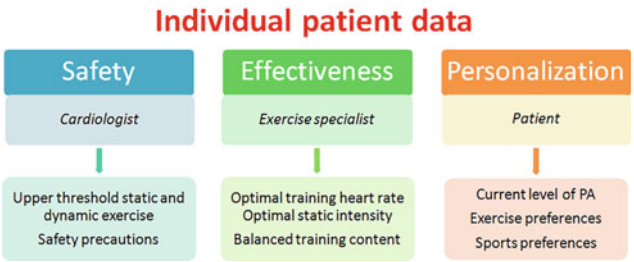



**Fig. 1** Overview of the different modules on individual patient data

## Content

The system needs to encompass (a) patient safety during exercise, (b) effectiveness in terms of maximal health benefit from the training program and (c) attractiveness and acceptability ensuring high adherence of the patient. Therefore, the service will consist of a computer based information system capable of combining data from 3 different modules filled in by the cardiologist specialized in CHD, the exercise specialist and the patient respectively, with the exercise content of the system. A schematic representation of the different modules regarding individual patient data is represented in Fig. 1.

### Safety

By fully adhering to the prevailing recommendations [11, 12], safety is ensured to the highest possible level. Following the most recent recommendation, the system will require data input regarding the five key hemodynamic parameters to be evaluated in the context of exercise prescription, on top of the first questions to the cardiologist, namely whether the patients' cardiac status is stable and whether the patient is not subject to non-cardiological conditions needing restrictions for exercise. As such, the CHD specialist enters data regarding the patients ventricular function, pulmonary artery pressure and aortic diameter, as well as regarding the presence of arrhythmias and desaturation at rest and/or during cardiopulmonary exercise testing [CPET] [11].

Using this information, along with information on peak heart rate reached during CPET, the system implements the related algorithm in order to obtain an upper heart rate threshold as well as an upper rate of perceived exertion [11].

### Effectiveness

In order for an exercise training program to be effective in terms of gains in physical fitness as defined by peak oxygen uptake, as well as in terms of benefits in health and quality of life, an aerobic exercise program can be optimized by specifying Frequency, Intensity, Type, and Time [FITT] of the exercise sessions [13]. Especially the intensity is key when improvement in fitness level is aimed for. The longer-term goal of the program will be to gradually build up an exercise volume with optimal training intensity. To define the optimal intensity, the system will use information from CPET, as this objective measurements provides data that take into account age, gender, current fitness level as well as cardiac status. Based on ventilatory thresholds and cardiac response during CPET, the system will be able to set the heart rate zone for optimal aerobic exercise training.

Further, next to gains in aerobic capacity, the exercise program will also target peripheral and respiratory muscle strength and endurance as well as flexibility and coordination if needed. Therefore, data derived from standardized fitness tests will be provided to the system by the exercise specialist [14].

### Personalization

The ultimate aim of any exercise program is the uptake of a life-long physically active lifestyle. Indeed, implementation of an exercise routine in daily life is necessary to obtain lasting health benefits [15]. Therefore, an exercise program must be clear, approachable, attractive and encouraging. To achieve this, information on current exercise and physical activity habits of the patients, as well as patient preferences towards exercise, physical activity and sports participation needs to be taken into account. The system will collect this information by implementing online questionnaires of which the scores are automatically calculated in order to feed the systems' rules and algorithms.

### Exercise and Sports Database

The exercise and sports database of the system will encompass an extensive set of exercises allowing for generating balanced exercise sessions, with every exercise labeled for intensity of the static and dynamic component, for difficulty and for suitability for different exercise phases such as warming up and endurance training. Furthermore, the sports database will consist of a large set of sports, as listed in the Compendium of Ainsworth [16] and with these sports classified as adapted from Mitchell et al. [17] to again be able to take into account static as well as dynamic intensity.

## Output

The proposed system will automatically propose an exercise prescription based on the variables provided (see Fig. 2). This prescription will consist of the short and mid-term exercise goal for the patient to target. Moreover, the system will propose a structured exercise program in order to gradually work toward achieving the goals.

### Exercise Prescription and Target

The exercise prescription will consist of the frequency, intensity, type and time of exercise to ultimately adhere to lifelong, along with a list of suggested sports and physical activities. Moreover safety advices related to the underlying heart defect and cardiac status are provided.

### How to Gradually Build Up the Exercise Program

On top of the exercise prescription, the system will take into account the current level of physical activity and physical fitness of the patient to provide a gradual exercise program aiming at reaching the prescribed exercise volume.

## Discussion and Future Challenges

The proposed system is rather easy to set up and given the implementation of current guidelines and recommendations along with expertise from exercise scientists in the field we believe that results of using the generated exercise programs will be most likely beneficial. Therefore, this system could become a useful tool for health professionals in CHD when prescribing exercise programs. Moreover, this exercise program generator can easily be build into technology enabled rehabilitation systems such as PATHway, such that the patient does not only receive an exercise program, but is guided towards following the program using an integrated

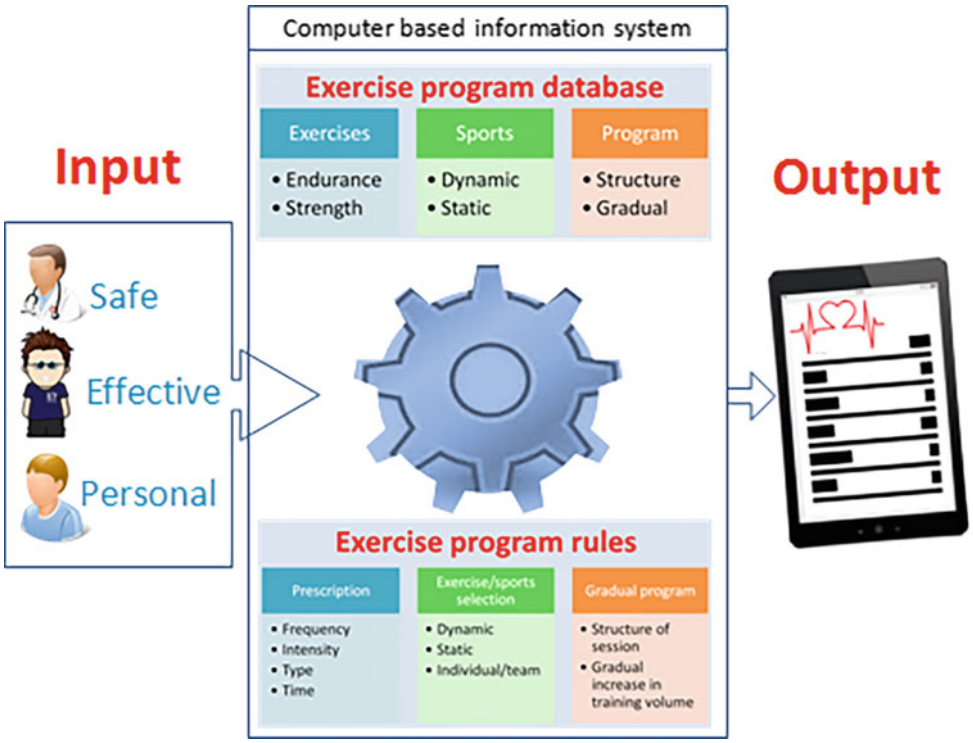



**Fig. 2** General structure of the proposed system

system that encompasses heart rate monitoring and motion capturing for providing exercise training sessions [18].

Nevertheless, testing of the system by means of a clinical trial will be needed to document the effectiveness as well as to unveil shortcomings to be taken into account when improving the system.

## Conclusion

In this paper, we presented the design of a fully automated service to generate an individualized exercise rehabilitation program for adults with CHD. The system will support CHD specialists in prescribing detailed, understandable and effective exercise and physical activity programs according to patient needs and wants, based on quantified information on cardiac status, exercise performance, physical activity level and patient preferences. As such, this work contributes to the development of computer-assisted systems targeting at personalized medicine and ultimately better health outcomes in CHD.

**Acknowledgements** This work has been partially funded by a research grant from the Research Foundation Flanders [FWO], and the European Union's Horizon 2020 Framework Programme for Research and Innovation Action under Grant Agreement no. 643491, 'PATHway: Technology enabled behavioural change as a pathway towards better self-management of CVD'.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Moons P, Sluysmans T, De Wolf D et al (2009) Congenital heart disease in 111 225 births in Belgium: birth prevalence, treatment and survival in the 21st century. Acta Paediatr 98(3):472–477
2. Moons P, Bovijn L, Budts W et al (2010) Temporal trends in survival to adulthood among patients born with congenital heart disease from 1970 to 1992 in Belgium. Circulation 122(22):2264–2272
3. Deanfield J, Thaulow E, Warnes C et al (2003) Management of grown up congenital heart disease. Eur Heart J 24(11):1035–1084
4. Verheugt CL, Uiterwaal CS, Grobbee DE et al (2008) Long-term prognosis of congenital heart defects: a systematic review. Int J Cardiol 131(1):25–32
5. Moons P, Van Deyk K, Dedroog D et al (2006) Prevalence of cardiovascular risk factors in adults with congenital heart disease. Eur J Cardiovasc Prev Rehabil 13(4):612–616
6. Duppen N, Takken T, Hopman MT et al (2013) Systematic review of the effects of physical exercise training programmes in children and young adults with congenital heart disease. Int J Cardiol
7. Harrison JL, Silversides CK, Oechslin EN et al (2011) Healthcare needs of adults with congenital heart disease: study of the patient perspective. J Cardiovasc Nurs 26(6):497–503

8. Dontje ML, Feenstra M, de Greef MH et al 2013 Are grown-ups with congenital heart disease willing to participate in an exercise program? Congenit Heart Dis
9. Gierat-Haponiuk K, Haponiuk I, Chojnicki M et al (2011) Exercise capacity and the quality of life late after surgical correction of congenital heart defects. Kardiol Pol 69(8):810–815
10. Rhodes J, Ubeda TA, Jenkins KJ (2010) Exercise testing and training in children with congenital heart disease. Circulation 122 (19):1957–1967
11. Budts W, Borjesson M, Chessa M et al (2013) Physical activity in adolescents and adults with congenital heart defects: individualized exercise prescription. Eur Heart J 34(47):3669–3674
12. Takken T, Giardini A, Reybrouck T et al (2012) Recommendations for physical activity, recreation sport, and exercise training in paediatric patients with congenital heart disease: a report from the Exercise, Basic & Translational Research Section of the European Association of Cardiovascular Prevention and Rehabilitation, the European Congenital Heart and Lung Exercise Group, and the Association for European Paediatric Cardiology. Eur J Prev Cardiol 19(5):1034–1065
13. Vanhees L, Geladas N, Hansen D et al (2012) Importance of characteristics and modalities of physical activity and exercise in the management of cardiovascular health in individuals with cardiovascular risk factors: recommendations from the EACPR. Part II. Eur J Prev Cardiol 19(5):1005–1033
14. Council of Europe (1995) Eurofit for adults. Assessment of health-related fitness
15. Claes J, Buys R, Budts W et al (2017) Longer-term effects of home-based exercise interventions on exercise capacity and physical activity in coronary artery disease patients: A systematic review and meta-analysis. Eur J Prev Cardiol 24(3):244–256
16. Ainsworth BE, Haskell WL, Whitt MC et al (2000) Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 32(9 Suppl):S498–S504
17. Mitchell J, Haskell W, Snell P et al (2005) Task force 8: classification of sports. J Am Coll Cardiol 45(8):1364–1367
18. Claes J, Buys R, Woods C, Briggs A et al (2017) PATHway I: design and rationale for the investigation of the feasibility, clinical effectiveness and cost-effectiveness of a technology-enabled cardiac rehabilitation platform. BMJ Open 7(6):e016781

# Adherence to Physical Activity in Patients with Heart Disease: Types, Settings and Evaluation Instruments

K. Livitckaia, V. Koutkias, N. Maglaveras, E. Kouidi, M. van Gils, and I. Chouvarda

**Abstract**

Physical fitness is one of the main therapeutic recommendations for patients with heart disease. Yet, adherence to physical activity regimen in both daily life and rehabilitation programs remains to be low. Elaborating on the analysis of factors associated with physical activity behavior, we noticed that the concept of adherence is complex, and its evaluation depends on a specific behavior and its settings. Evaluation of adherence to exercise in leisure time and exercise in the settings of cardiac rehabilitation program requires an application of different instruments. In this paper, we present a summary of findings from the literature analysis regarding types, settings and evaluation instruments of physical activity adherence in daily life and cardiac rehabilitation settings.

**Keywords**

Patient adherence • Physical activity • Heart disease • Health behavior • Translational research

## Introduction

Poor patient adherence to therapies and lifestyle recommendations is becoming one of the most alarming public health-related issues globally. Studies show that the low level of adherence is widely spread among patients suffering from various diseases, yet patients with chronic diseases are in the pressing position due to the need to follow medication and lifestyle change recommendation for a long period of time, and, often, lifelong. Regardless of the importance of making healthy choices daily, an overall patient adherence to such changes is poor [1].

Particularly acute adherence problem is in patients with cardiovascular disease (CVD). Physical activity and fitness are two of the main components in therapeutic recommendations for patients with heart disease. However, adherence to physical activity regimen in both, lifestyle choices and rehabilitation programs remains to be critically low [1]. According to the results of the studies, the percentage of adherence to cardiac rehabilitation exercise programs varies from 39 to 46% [2, 3].

CVD patients do not follow given recommendations for a variety of reasons and influencing factors, including lack of motivation, having high levels of anxiety, being busy with family responsibilities, etc. Studying the factors associated with physical activity behavior in patients with heart disease, we noticed that the concept of adherence is complex from the

K. Livitckaia (✉) · N. Maglaveras · I. Chouvarda
Lab of Computing, Medical Informatics & Biomedical Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Aristotle University Campus, 54124 Thessaloniki, Greece
e-mail: kristinaliv@auth.gr

V. Koutkias
Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thermi, Thessaloniki, Greece

E. Kouidi
Laboratory of Sports Medicine, School of Physical Education and Sports Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

M. van Gils
VTT Technical Research Centre of Finland Ltd., Tampere, Finland



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*, IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_42

perspective of types and settings of the specific activity behavior. Evaluation of adherence differs depending on the context of physical activity and exercise as well. For example, exercise in leisure time and exercise in the setting of a cardiac rehabilitation program (CRP) are not the same behavior, and adherence should be evaluated using different instruments. Moreover, the same patient could be adherent to the exercise training in CRP, but does not exercise on a daily basis after finishing the program. We found it relevant to continue with aforesaid analysis. Therefore, the aim of this study is to summarize and present findings from the literature analysis regarding types, settings and evaluation instruments of physical activity adherence in daily life and CRP settings.

## Material and Methods

The study is organized and reported as theoretical analysis, as it relies on published material that is relevant for the analysis. For the review, we selected forty-one scientific articles retrieved from MEDLINE and the Cochrane Library between June and August 2017 [4–44]. All the articles were published in English, available in full text for the authors of this study, focusing on CVD patients and investigating physical activity adherence.

In the thematic analysis, we addressed subjects such as (a) settings of CRPs, (b) interpretation and calculation of adherence in cardiac rehabilitation (CR), (c) settings and types of independent physical activity and exercise, and (d) instruments to evaluate adherence to independent physical activity behavior. The selection and analysis process was iterated several times.

## Results

Patient adherence reflects a performed level of physical activity and fitness in everyday life, as well as in CRPs. From the studied articles, we noticed that physical activity and exercise performed outside the hospital (or other healthcare facility), and when the patient has no real-time supervision from the healthcare professional, is called independent physical activity and exercise. In opposite, an exercise performed under the supervision in an organized facility (e.g., hospital, rehabilitation center), it is specified as controlled exercise. Thus, we will refer to physical activity behavior performed in everyday life as an independent physical activity and exercise, while controlled exercise will imply exercise in the form of CRPs.

### Settings of Cardiac Rehabilitation Programs

The majority of CRPs in the selected studies are similar in the design. The main part is represented by a supervised physical exercise training with two to three sessions per week, lasting for an hour on average, and supported by additional health-related educational sessions, and occasionally psychological or emotional support.

Aiming to understand specific settings of supervised physical exercise, we found that programs' durations varied from two weeks to four years. Larger part of programs was organized for a short-term period (up to three months), less programs lasted up to a year, and the minority was organized for periods of up to three years and longer. Our analysis showed that as the duration of CRPs gets longer, less studies are organized to investigate adherence.

Continuing with the analysis of the programs durations, we outlined CRPs with respect to the phase of cardiac rehabilitation. Most of the programs were designed as outpatient exercise training, representing second and third rehabilitation phase. Two programs were designed for in-patient first phase training. In one paper the key interest was in adherence to any of participated CRPs, rather than in a specific phase.

### Defining and Quantifying Adherence in Cardiac Rehabilitation

Adherence to CR physical exercise in the selected studies is defined through the program attendance. Commonly, researchers quantify adherence to cardiac rehabilitation programs as a number of exercise sessions attended divided by the number of sessions scheduled. However, there are differences in values when defining adherence and nonadherence. Patients who attended less than 50% of or less than 66% of the scheduled study sessions are categorized as being non-adherent.

Another concept related to the patients not following sufficient number of prescribed physical exercises in the rehabilitation program is dropout. Among the chosen publications, nine studies are focusing on investigations related to a patient dropout. The dropout was recorded when a patient attended 75% or less or two third or less of a scheduled program, or three or more consecutive activity sessions were missed. If there was a confirmation by a patient that he or she no longer wished to or was unable to participate, or if the patient was unreachable by telephone, the program participation was considered as a dropout.

## Settings, Types and Dimensions of Independent Physical Activity and Exercise

The analysis of studies focusing on independent physical activity behavior, allows to outline two groups of behaviors —an independent physical activity and independent exercise. In most studies, physical activity behavior was investigated without a particular context. We also noticed that independent physical activity and exercise were studied under specific settings related to heart disease, such as during and after cardiac rehabilitation program, and after index hospitalization. Further, we identified behavior dimensions such as level, regularity, frequency, persistence, maintenance, intensity, and the total amount.

## Instruments to Evaluate Adherence to Independent Physical Activity Behavior

In majority of the selected studies related to independent physical activity and exercise, adherence was evaluated through scores of subjective instruments such as assessments and self-reported diaries. Larger number of the assessments were organized in a free form as questionnaires, open-ended questions, and interviews. Generally, patient adherence was evaluated based on the answers regarding minutes and regularity of performed physical activity; however, we noted several assessments regarding weekly, monthly, and yearly adherence evaluation. Below we detail these instruments.

### Adherence Evaluation for the Previous Week

*Godin Leisure-Time Exercise Questionnaire (LTEQ)*: In this questionnaire, patients are asked about performed physical activity during a typical seven-day period; how many times on average do they perform a light, moderate, and strenuous exercise for more than 15 min. The LTEQ score is calculated by adding the frequency of exercise within the specified physical activity.

*Physical Activity Scale for the Elderly (PASE)*: In this ten-item tool, patients are asked for how many days per week, and how much time they spend for sitting, walking, and moderate or vigorous activity over the past week. A higher PASE score indicates a greater level of everyday physical activity and independent exercise.

*International Physical Activity Questionnaire (IPAQ, s-IPAQ)*: The instrument is targeting frequency and duration of physical activity (vigorous and moderate activity, walking) and inactivity (sitting) during the past week. The total activity score is transformed into the metabolic equivalent (MET), and the patients can be assigned to low, moderate, or high physical activity group. The difference between full IPAQ and s-IPAQ, is that in the full version patients are asked regarding (in)activity behaviors within job-, transport-, home-, and sport and leisure-related activities.

*7-Day Physical Activity Recall Questionnaire (PAR)*: In PAR, patients are assessed about the time spent for sleeping and being engaged in moderate, hard and very hard physical activity within the previous week. From hours spent in physical activities, total kilocalories/day can be estimated and transformed into MET.

### Adherence Evaluation for the Previous Month

*The Self-Care of Heart Failure Index (SCHFI)*: Using this instrument, patients are asked about the regularity of performing physical activity over the past month by two questions, addressing regularity and duration of exercise.

### Adherence Evaluation for the Previous Year

*Minnesota Leisure-Time Physical Activity Questionnaire (LTPA)*: In the questionnaire (modified version), physical activities are grouped (i.e., walking and miscellaneous, and conditioning exercise) and assessed through the interview. Patients should answer about the activities they performed and did not perform within the last 12 months.

## Discussion

In this section, we discuss study limitations and our thoughts on the summary of the presented results.

*Limitations*: The limitation of our study is the number of papers selected for the analysis. We believe that forty-one articles might not be enough to create a strong evidence base for all physical activity adherence aspects; however, we consider our selection adequate for outlining that such aspects should be addressed and specified in adherence studies.

*Not certain understanding of (non)adherence*: From the analysis, most of the CRPs were relatively short (i.e., lasting up to three months). We consider an evaluation of short-term adherence based on the attendance rates as to the purpose. However, taking into account forth phase of CRP, when physical fitness is performed not only under supervision, but as an independent activity as well, tools for the adherence evaluation might not be limited to the list of participation. Moreover, we found that the level of adherence that is required to be obtained for the therapeutic benefit, is not clearly defined, or differs in the studied materials.

*The subjectivity of adherence evaluation instruments*: An attention should be given to the evaluation instruments that were applied to an independent physical activity and exercise. Mostly, patient adherence was evaluated by the scores from verbal or written assessments. However, the fact that most of the questions were very limited in terms of the content (e.g., including one or two questions about the

activity), the quality of such data and further adherence evaluation should be considered. The use of technology-based resources (e.g., wearables) might provide an objective perspective for patient adherence evaluation through gathered data on the number of daily steps, activities performed, calories burned and heart rate variations; however, in the selected materials, only in one study such instrument was employed.

## Conclusions

The results of the analysis summarized in the paper highlight essential aspects of heart disease patient adherence, including types, settings and evaluation instruments regarding specific physical activity behavior in daily life, as well as in CRPs. Such aspects should be considered when developing eHealth and employed to usual healthcare practice interventions, when investigating patient adherence for further knowledge reusability.

**Acknowledgements** This overview has been conducted within the Connected Health Early Stage Researcher Support System (CHESS) project; under the framework of Marie Skłodowska-Curie grant agreement No. 676201.

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. World Health Organization (2003) Adherence to long-term therapies: evidence for action at http://www.who.int/chp/knowledge/publications/adherence_full_report.pdf
2. Conraads VM, Deaton C et al (2012) Adherence of heart failure patients to exercise: barriers and possible solutions. Eur J Heart Fail 14(7):802
3. van der Wal MH, Jaarsma T et al (2006) Compliance in heart failure patients: the importance of knowledge and beliefs. Eur Heart J 27(4):434–440
4. Blanchard CM, Rodgers WM et al (2002) Self-efficacy and mood in cardiac rehabilitation: should gender be considered? Behav Med 27(4):149–160
5. Schwarzer R, Lippke S (2008) Social-cognitive predictors of physical exercise adherence: three longitudinal studies in rehabilitation. Health Psychol 27(1S):S54–S63
6. Mak Y, Chan W et al (2005) Barriers to participation in a phase II cardiac rehabilitation programme. Hong Kong Med J 11(6):472–475
7. Bock BC, Albrecht AE et al (1997) Predictors of exercise adherence following participation in a cardiac rehabilitation program. Int J Behav Med 4(1):60–75
8. Laustsen S, Hjortdal VE et al (2013) Predictors for not completing exercise-based rehabilitation following cardiac surgery following cardiac surgery. Scand Cardiovasc J 47(6):344–351
9. Wittmer M, Volpatti M et al (2011) Expectation, satisfaction, and predictors of dropout in cardiac rehabilitation. Eur J Prev Cardiol 19(5):1082–1088
10. Anderson DR, Emery CF (2014) Irrational health beliefs predict adherence to cardiac rehabilitation: a pilot study. Health Psychol 33(12):1614–1617
11. Gallagher R, McKinley S et al (2003) Predictors of women's attendance at cardiac rehabilitation programs. Prog Cardiovasc Nurs 18(3):121–126
12. Forhan M, Zagorski BM et al (2013) Predicting exercise adherence for patients with obesity and diabetes referred to a cardiac rehabilitation and secondary prevention program. Can J Diab 37 (3):189–194
13. Banerjee AT, Gupta M et al (2007) Patient characteristics, compliance, and exercise outcomes of South Asians enrolled in cardiac rehabilitation. J Cardiopulm Rehabil Prev 27(4):212–218
14. Glazer KM, Emery C et al (2002) Psychological predictors of adherence and outcomes among patients in cardiac rehabilitation. J Cardiopulm Rehabil 22(1):40–46
15. Marzolini S, Grace SL et al (2015) Time-to-referral, use, and efficacy of cardiac rehabilitation after heart transplantation. Transplantation 99(3):594–601
16. Marzolini S, Danells C et al (2016) Feasibility and effects of cardiac rehabilitation for individuals after transient ischemic attack. J Stroke Cerebrovasc Dis 25(10):2453–2463
17. Oldridge NB, Donner AP et al (1983) Predictors of dropout from cardiac exercise rehabilitation. Am J Cardiol 51(1):70–74
18. Dorn J, Naughton J et al (2001) Correlates of compliance in a randomized exercise trial in myocardial infarction patients. Med Sci Sports Exerc 33(7):1081–1089
19. Cannistra LB, Balady GJ et al (1992) Comparison of the clinical profile and outcome of women and men in cardiac rehabilitation. Am J Cardiol 69(16):1274–1279
20. Shanmugasegaram S, Oh P et al (2015) A comparison of barriers to use of home versus site-based cardiac rehabilitation. J Cardiopulm Rehabil Prev 33(5):297–302
21. Reges O, Vilchinsky N et al (2013) Illness cognition as a predictor of exercise habits and participation in cardiac prevention and rehabilitation programs after acute coronary syndrome. BMC Pub Health 13:956
22. Zhang KM, Dindoff K et al (2015) What matters to patients with heart failure? The influence of non-health-related goals on patient adherence to self-care management. Patient Educ Couns 98 (8):927–934
23. Bentley D, Khan S et al (2013) Physical activity behavior two to six years following cardiac rehabilitation: a socioecological analysis. Clin Cardiol 36(2):96–102
24. Tierney S, Elwers H et al (2011) What influences physical activity in people with heart failure? A qualitative study. Int J Nurs Stud 48 (10):1234–1243
25. Albert NM, Forney J et al (2015) Understanding physical activity and exercise behaviors in patients with heart failure. Heart Lung 44 (1):2–8
26. Klompstra L, Jaarsma T et al (2015) Physical activity in patients with heart failure: barriers and motivations with special focus on sex differences. Patient Prefer Adherence 9:1603–1610
27. Lima G, Ghisi DM et al (2015) Disease-related knowledge in cardiac rehabilitation enrollees: correlates and changes. Patient Educ Couns 98(4):533–539
28. Rogerson MC, Murphy BM et al (2012) "I don't have the heart": a qualitative study of barriers to and facilitators of physical activity for people with coronary heart disease and depressive symptoms. Int J Behav Nutr Phys Act 9:140

29. Heydari A et al (2015) Relationship between awareness of disease and adherence to therapeutic regimen among cardiac patients. Int J Community Based Nurs Midwifery 3(1):23–30
30. Heydari A, Ahrari S et al (2011) The relationship between self-concept and adherence to therapeutic regimens in patients with heart failure. J Cardiovasc Nurs 26(6):475–478
31. Mosleh SM, Darawad M (2015) Patients' adherence to healthy behavior in coronary heart disease: risk factor management among jordanian patients. J Cardiovasc Nurs 30(6):471–478
32. Hardcastle SJ, Mcnamara K et al (2015) Using visual methods to understand physical activity maintenance following cardiac rehabilitation. PLoS ONE 10(9):e0138218
33. Moore SM, Dolansky MA et al (2003) Predictors of women's exercise maintenance after cardiac rehabilitation. J Cardiopulm Rehabil 23(1):40–49
34. Evangelista LS, Berg J et al (2001) Relationship between psychosocial variables and compliance in patients with heart failure. Heart Lung 30(6):476–478
35. Gallagher R, Luttik M et al (2011) Social support and self-care in heart failure. J Cardiovasc Nurs 26(6):439–445
36. Hellman EA (1997) Use of the stages of change in exercise adherence model among older adults with a cardiac diagnosis. J Cardiopulm Rehabil 17(3):145–155
37. Cooper LB, Mentz RJ et al (2015) Psychosocial factors, exercise adherence, and outcomes in heart failure patients: insights from heart failure: a controlled trial investigating outcomes of exercise training (HF-ACTION). Circ Heart Fail 8(6):1044–1051
38. Kuhl EA, Fauerbach JA et al (2009) Relation of anxiety and adherence to risk-reducing recommendations following myocardial infarction. Am J Cardiol 103(12):1629–1634
39. Subramanian U, Hopp F et al (2008) Impact of provider self-management education, patient self-efficacy, and health status on patient adherence in heart failure in a Veterans Administration population. Congest Heart Fail 14(1):6–11
40. Murray T, Rodgers W (2012) The role of socioeconomic status and control beliefs on frequency of exercise during and after. Appl Psychol Health Well Being 4(1):49–66
41. Urbinati S, Olivari Z et al (2015) Secondary prevention after acute myocardial infarction: drug adherence, treatment goals, and predictors of health lifestyle habits. The BLITZ-4 registry. Eur J Prev Cardiol 22(12):1548–1556
42. Mosleh SM, Almalik MM (2014) Illness perception and adherence to healthy behaviour in Jordanian coronary heart disease patients. Eur J Cardiovasc Nurs 15(4):223–230
43. Russell KL, Bray SR (2009) Self-determined motivation predicts independent, home-based exercise following cardiac rehabilitation. Rehabil Psychol 54(2):150–156
44. Chan RH, Gordon NF et al (2008) Influence of socioeconomic status on lifestyle behavior modifications among survivors of acute myocardial infarction. Am J Cardiol 102(12):1583–1588

# Training System Methodology Using ECG Signal

E. Butkeviciute, L. Bikulciene, and K. Poderiene

**Abstract**

The personalized smart training systems is a new technology that supports the personal activity of professional or amateur sport. It is important to avoid overtraining and reach expected results of physical activity during exercises and competitions. One of the most severe health disorders caused by overtraining is the heart disfunction or even heart failure. The electrocardiogram (ECG) signal conveys several different parameters that allow characterizing health conditions and identifying several pathologies. That is why the ECG signal analysis is so important for training. However, during exercise the registration of the signal could be complicated because various noises might appear due to: electrode contact instability, power line interference, baseline wanderings caused by muscular movements and etc. For the analysis the ECG was registered while person was doing squats. This paper contains the description of few methods for ECG data filtering which is done separately: trend (low frequency noise) removal and high frequency noise reduction. This provide more accurate estimation of the signal parameters and it allows to create the training system methodology for decision making. The main task of this research was to describe a training system methodology which includes ECG parameters estimation and training intensity regulations. All these algorithms could be used in future researches for signal compression and improvement of various personalized training systems.

**Keywords**

ECG signal analysis • Personalized training system • Rule based decision making

## Introduction

These days more and more people practice sports and do exercises, often without properly evaluating their abilities and health conditions. Health evaluation is usually carried out stationary. However, this technique is inconvenient for professionals or ordinary people during physical activity. Furthermore, there are some difficulties in evaluating health changes during physical activity. There are many ECG signal registration devices which detect heart problems accurately. However, these devices are used in stationary conditions and that does not allow to detect some health disorders especially during exercises. Portable devices usually show only the heart rate and find out the running distance or the number of calories which were burnt during exercises.

It is important to follow the sportsman health state condition during whole training session. Everyone would like to know if their training session is too intense and injuries might appear. Meanwhile if the athlete is preparing for the competition, he seeks for the best shape and recommendations

E. Butkeviciute (✉) · L. Bikulciene
Department of Applied Mathematics, Kaunas University of Technology, Studentu str. 50, Kaunas, Lithuania
e-mail: egle.butkeviciute@ktu.edu

K. Poderiene
Institute of Sport Science and Innovations, Lithuanian Sport University, Sporto str. 6, Kaunas, Lithuania



N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6_43

about training session are necessary. He would like to know if the exercises are sufficiently intensive and allows to reach the best results. Also, the dynamics and individual activity in daily-life activities are important to be considered for exercise intensity control.

The training system should contain primary ECG signal analysis, parameter estimation and feedback generation about training intensity. This part is described in this paper. The proposed training system is applicable to healthy people who are physically active. However, the whole training methodology should contain personalized training algorithms that will be done in future researches with gathered data about each person. The final decision-making algorithm should also show if participant is doing some progress or not.

## ECG Signal Analysis

During exercise, the whole body is moving, and it causes high and low frequency noise. The electrodes from the T-shirt send the data to the Cardioscout [1] recorder and then to a mobile device (smart phone or smart watch). In this paper the ECG signal was recorded while person was doing squats. Each value of the signal was taken every 2 ms at 500 Hz frequency.

Even in stationary conditions the ECG small parameters (that gain the small amplitude or duration values, like AST or DJT) estimation is a complicated task. Of course, the beat-to-beat (RR) interval [2] evaluation is usually possible in any case: while person is moving or not. However, there are many other parameters like amplitude of ST (AST) [3] or duration of JT interval (DJT) (see Fig. 1) that are almost impossible to estimate if ECG signal contains much of noise. The main task of this ECG signal analysis is to eliminate data trend and reduce high frequency noise. The trend appears mainly because of the motion artifacts during exercises (like doing squads), muscular movements and conductance between the electrodes and the human body. Also, the electrode contact noise, power line interference and other disturbance causes high frequency noise.

There are many algorithms for ECG data low frequency noise reduction like Butterworth algorithm [4] or fast Fourier transform (FFT) [5]. Even though they are very effective for stationary signals, they fail in volatility in time. After comparison of different algorithms, the baseline estimation and denoising sparsity (BEADS) was selected in respect of root mean square errors (RMSE). The BEADS [5] algorithm was used for chromatogram baseline estimation [6] and it seemed to be suitable for trend removal (low frequency noise reduction) for ECG signals. In this algorithm, the "baseline" refers to the smoothest part of the trend or bias and the "noise" is more stochastic part. This algorithm is based on

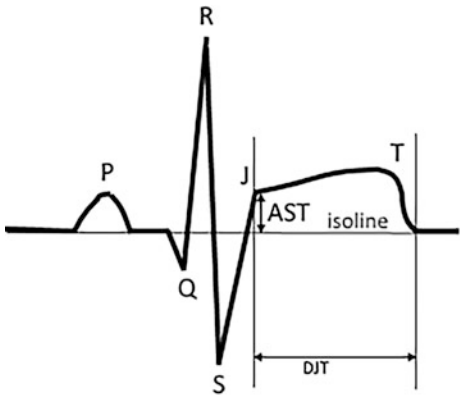



**Fig. 1** ECG parameters

ECG data baseline (or low-frequency noise wave) detection and its removal.

Furthermore, the discrete wavelet transform (DWT) was chosen for high frequency noise removal. This algorithm is appropriate to be used for non-stationary signals (like ECG) because it decomposes the signal into a frequency time scale. In this paper, the Daubechies wavelets [7] are chosen for ECG data filtering. The DWT algorithm contains three steps: (1) the DWT is applied on a noisy signal (on $x_{input}$); (2) the thresholding process is applied. The wavelet coefficients (noise is generated by small value coefficients) are filtered by throwing them away (resetting to zero) during this process; (3) remaining coefficients are back-converted in time domain (Inverse Discrete Wavelet Transform (IDWT)) [7]. The Butterworth filter was compared with DWT algorithm and the DWT gave better RMSE values: $RMSE_{Butterworth} = 0.1206$, $RMSE_{DWT} = 0.0559$. The scheme of DWT signal processing is presented in Fig. 2. In this scheme, the filtered signal (output) is marked as $x_{output}$.

The real data from the T shirt was taken for the ECG data filtering. In Fig. 3 the data of participant who was doing squads is shown. The original data taken from the T-shirt are shown in part (a). In the ECG signal the trend jumps at each

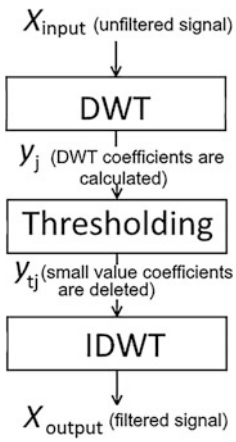



**Fig. 2** A schematic block of DWT algorithm

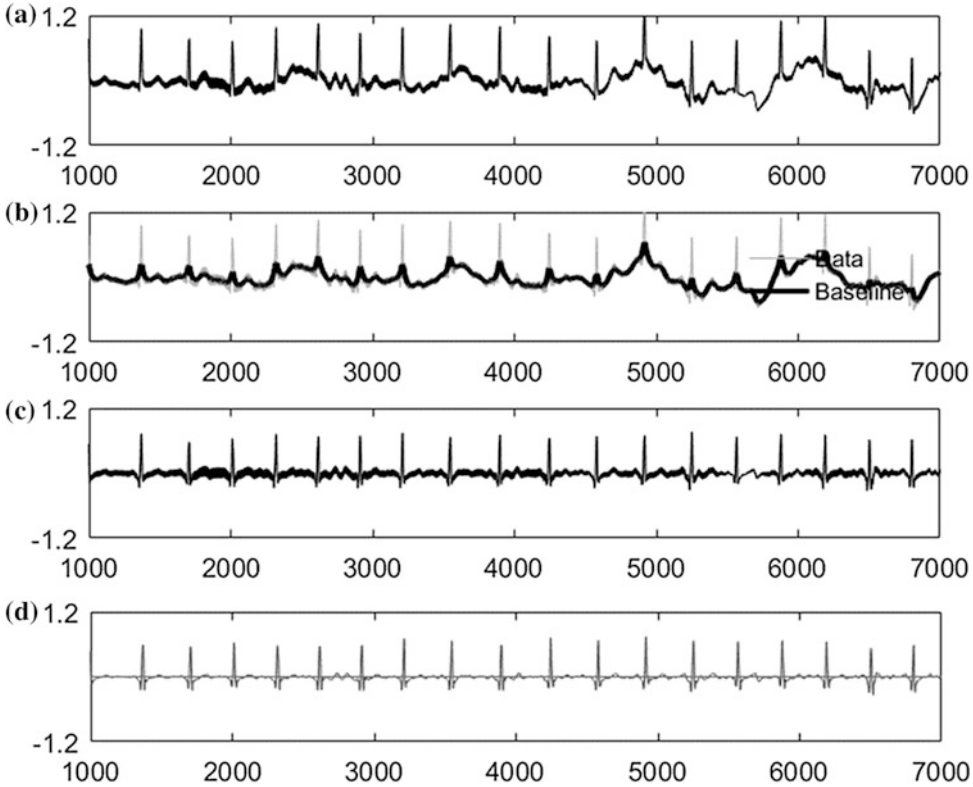



**Fig. 3** **a** Original data from T-shirt while participant is doing exercises; **b** Base line detection; **c** Detrended ECG data; **d** Denoised and detrended ECG signal

squad. Also, the signal contains some additional high frequency noise which creates small peaks marked as thicker line in part (b). The high and low frequency noise filtering results are presented in parts (c) and (d).

## ECG Parameter Estimation

The ECG parameter estimation is quite complicated task if the data contains too much noise. After de-noising the parameter search algorithms could be applied. There are three steps required in parameter evaluation:

- Beat-to-beat (RR) peaks detection, which should be the highest value according to isoline. This interval should not reach the 180 ms duration.
- Using R value, the Q and S peaks are found. Each peak is detected when the first derivative changes its sign. Also, P and T peaks are the highest points between S and Q waves. The S wave ends after reaching the isoline. The T wave beginning point is a point where the signal starts to increase again with respect to isoline.
- The J point is the inflection point between the end of S wave and the beginning of T wave. This parameter is important in estimation of JT duration (DJT).

In some cases, P or T values are too small to detect because it depends on the location from which the recordings take place, on the body mass and even on each person heart pathology. If the T or P wave peak was not detected during calculations the previous value is selected. However, after all applied filters, the ECG data still contains some noise and it causes some difficulties in DJT parameter estimation. Inflection point is very sensitive to all small peaks.

Furthermore, the ECG discriminant is the characteristic of concatenation between RR and QRS intervals which is very important in making training system more personalized [10]. The formula of discriminant (defined as D) is written below:

$$D(RR-QRS)_n = (RR_n - QRS_n)^2 + 4(RR_{n-1} - QRS_{n-1})(RR_{n+1} - QRS_{n+1}) \quad (1)$$

Here RR and QRS are durations of RR and QRS complexes, $n$ is the number of cardio cycle [8]. This parameter is very sensitive for all small ECG signal changes, but is widely used in health status estimation during stress tests [9]. In some cases, the parameters like QRS complex or T, P waves does not show the changes in health condition while the determinant D is very sensitive in all ECG signal changes [10].

## Simulation of the Training System

All people who are exercising would like more easily monitor their training intensity and follow their training plan. Together with scientists from Lithuanian Sports University (LSU) the training system methodology was prepared.

In this paper the real-time data is defined as 10 s time duration which contains 10 s ECG signal. Each 10 s the QRS interval and discriminant (D) are measured and their values are averaged. The training session can be modeled as it is written bellow:

*Duration*. Duration of the training session depends on the RR interval quantity. The recommended duration of the exercise should not exceed 1500 RR intervals [11].

*Heart rate*. The lower and upper bounds of heart rate have been defined as [12]:

$$\mathrm{HR_{low}} = ((220 - \mathrm{A}) - \mathrm{HR_b}) \cdot 0.5 + \mathrm{HR_b}, \quad (2)$$

$$\mathrm{HR_{up}} = ((220 - \mathrm{A}) - \mathrm{HR_b}) \cdot 0.85 + \mathrm{HR_b}, \quad (3)$$

where A is the age of the person and $\mathrm{HR_b}$ is the heart rate before exercises (see [12]). The heart rate estimation in decision making algorithms is widely used in various training sessions, like group training [13]. However, these systems are not personalized and should be improved.

If heart rate during training session is lower than $\mathrm{HR_{low}}$, the training intensity should be increased. Also, if heart rate is higher than $\mathrm{HR_{up}}$, the exercises should become less intensive.

*QRS interval*. If whole 30 s QRS increases by 5%, the training intensity should be reduced. Then for the next 60 s the QRS estimation is suspended. After that time, if QRS interval still increases in 30 s, the training session should be stopped.

*DJT interval*. If DJT becomes less than 190 ms, the intensity should be reduced. Then this parameter is not being estimated for 60 s. And if the DJT is still less than 190 ms, the training session must be stopped.

*Discriminant*. Discriminant evaluation is similar to QRS interval. If it increases for 30 s by 10%, the training intensity should be reduced. After 60 s, it is measured again. And if it is increasing for the 30 s the training session should be stopped.

If ECG parameters are stable (QRS and D do not increase) or do not reach critical values, the intensity should remain the same. Otherwise, the training intensity should be increased, decreased or training session must be stopped [14].

The whole rule based decision making algorithm is presented in Fig. 4. A variable "index" in this scheme is used for assignment of the first or second ECG critical parameters value changes. At the start of the algorithm this variable is equal to zero and before intensity reduction it gains the value of 1. When corresponding parameter is not evaluated for 60 s, the training session can be stopped (if parameter still indicates critical health condition). Otherwise, if all parameters show stable health condition, the value of "index" variable is zero.

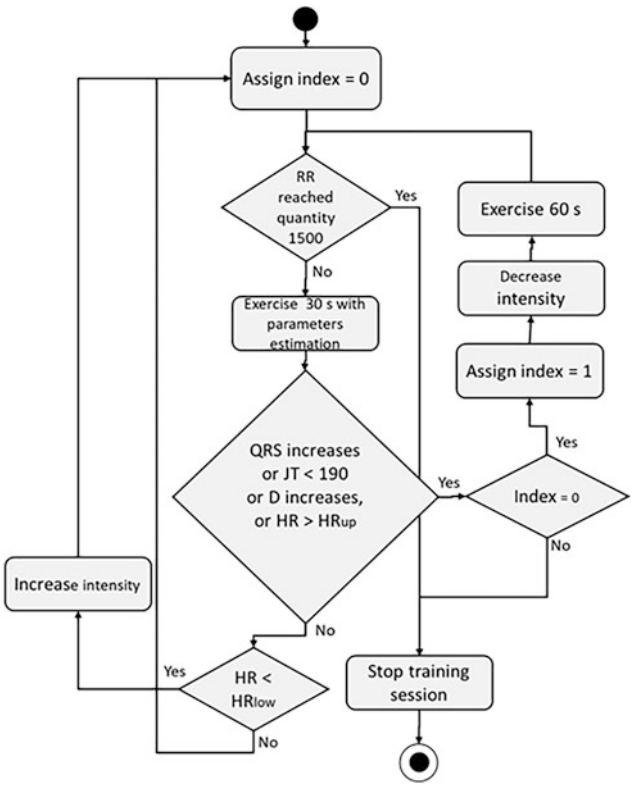



**Fig. 4** Visualization of decision making algorithm

The purpose of this algorithm is to detect physiological state changes and optimal exercise intensity according to internal and external factors (as adaptation to different geographical conditions, remained fatigue, etc.). In future researches it will be used for exercise stress tests with ECG monitoring methods that are usually used to detect individual physiological limits for optimal exercise intensity.

## Conclusions

In this paper the personalizes training system was considered and decision-making algorithm for training session was proposed. This algorithm contains ECG parameters (QRS, DJT, D and HR) estimation and training intensity regulation. For estimation of parameters various signal processing algorithms were used.

The BEADS algorithm was used to remove trend (low frequency noise) from ECG signal. This method concentrated the data on the isoline (see Fig. 3). Also, the DWT

algorithm was found to be suitable for ECG data high frequency noise reduction. Finally, ECG parameter estimation algorithms included the peak search, evaluation of the first derivative changes and detection of inflection point algorithms.

The personalized training session is using practical experience and recommendations from the scientists of LSU. Future work could contain additional validation of personalized training system in various exercises and creation of the smart technology with the integrated algorithms.

**Acknowledgements** This work was supported by a grant from the TD1405 TD COST Action TD1405 European Network for the Joint Evaluation of Connected Health Technologies (ENJECT).

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Cardioscout multi ECG. http://www.cardioscout.net/team.html
2. Basco J, Chen TM, Tapiador J, Peris-Lopez P (2016) A survey of wearable biometric recognition systems. ACM Comp Surv Article 43
3. Elbuni A, Kanoun S, Elbuni M, Ali N (2009) ECG parameter extraction algorithm using (DWTAE) algorithm. In: International conference on computer technologies and development, pp 57–62
4. Jagtap SK, Uplane MD (2012) The impact of digital filtering to ECG analysis: Butterworth filter application. In: International conference on information and communication technology (ICCICT), pp 1–6
5. Selesnick IW, Graber HL, Pfeil DS, Barbour RL (2014) Simultaneous low-pass filtering and total variation denoising. IEEE Trans Signal Process 1109–1124
6. Ning X, Selesnick I W, Duval L (2014) Chromatogram baseline estimation and denoising using sparsity (BEADS). Polytechnic School of Engineering, pp 1–23
7. Garg G, Singh V, Gupta JRP, Mittal AP (2010) Optimal algorithm for ECG denoising using discrete wavelet transforms. IEEE
8. Akshay N, Jonabhotla NAV, Sadam N, Yeddanapudi ND (2010) ECG noise removal and QRS complex detection using UWT. ICEIE V2:438–442
9. Clifford G, Sameni R, Ward J, Robinson J, Wolberg AJ (2011) Clinically accurate fetal ECG parameters acquired from maternal abdominal sensors. Am J Obstet Gynecol 205:47e1-5
10. Venskaityte E, Poderys J, Balagué N, Bikulciene L (2009) Assessment of dynamics of inter-parameter concatenation during exercise tests. Electron Electr Eng 6(94):89–92
11. Pratt WM1, Siconolfi SF, Webster L, Hayes JC, Mazzocca AD, Harris BA Jr (1991) A comparison between computer-controlled and set work rate exercise based on target heart rate. Aviat Space Environ Med 899–902
12. Poderys J, Venskaitytė E, Poderienė K, Buliuolis A, Vainoras A (2010) Functional state assessment on the dynamics of interparametric concatenations during exercise tests. Medicina 429–434
13. Vehkaoja A, Verho J, Comert A, Aydogan B, Perhonen M, Lekkala J, Halttunen J (2008) System for ECG and heart rate monitoring during group training. IEEE EMBS Conf, Canada
14. James DA, Petrone N (2016) Sensors and wearable technologies in sport technologies, trends and approaches for implementation. Springer briefs in applied science and technology

# Author Index




N. Maglaveras et al. (eds.), *Precision Medicine Powered by pHealth and Connected Health*,
IFMBE Proceedings 66, https://doi.org/10.1007/978-981-10-7419-6