PHONOCARDIOGRAM-BASED DIAGNOSIS USING MACHINE LEARNING: PARAMETRIC ESTIMATION WITH MULTIVARIANT CLASSIFICATION

Shaima Abdelmageed[1], Mohammed Elmusrati[2]

School of Technology and Innovation, University of Vasa, Vasa, Finland

ABSTRACT

The heart sound signal, Phonocardiogram (PCG) is difficult to interpret even for experienced cardiologists. Interpretation are very subjective depending on the hearing ability of the physician. mHealth has been the adopted approach towards quick diagnosis using mobile devices. However, it has been challenging due to the required high quality of data, high computation load, and high-power consumption. The aim of this paper is to diagnose the heart condition based on Phonocardiogram analysis using Machine Learning techniques assuming limited processing power to be encapsulated later in a mobile device. The cardiovascular system is modelled in a transfer function to provide PCG signal recording as it would be recorded at the wrist. The signal is, then, decomposed using filter bank and the analysed using discriminant function. The results showed that PCG with a 19 dB Signal-to-Noise-Ratio can lead to 97.33% successful diagnosis.

KEYWORDS

Analysis, Classification, data quality, diagnosis, filter banks, mHealth, PCG, SNR, transfer function, Wavelet Transform

1. INTRODUCTION

This paper presents the analysis of the heart's acoustic signal at the wrist. To find this signal, a transfer function is proposed to represent the impact of the travelled distance from the heart to the wrist on the characteristics of the PCG. This signal is expected to have low quality and low signal-to-noise ratio. To reduce the required computation load and power consumption and speed up the processing time the signal is downsampled by 100 sample/s before it is decomposed using filter bank into four subbands. Each subband is described using two features; mean and covariance. The system is trained using 300 cases to diagnose the heart condition against six hypothetical diseases. The classification is based on the discriminant function of the unclassified signal. The most probable diagnosis is found by maximising the discriminant function and in other words, minimising the Mahalanobis distance [1]. For the experiment, nonstationary noise is used to simulate nonstationary environment such as the chaotic accident environments.

2. OBTAINING PCG AT THE WRIST

This is achieved by modelling the heart-wrist acoustic wave propagation system using SIMULINK MATLAB, see Figure 1 for the model. A healthy heart acoustic signal with sampling

DOI: 10.5121/bioej.2018.5401

date of 44100 sample/s is used as a reference, it is downsampled by 100 sample/second to reduce the computation load and speed up the processing time. It should be noted that down sampling acts as a filter that removes the high frequencies, which works well in the case of low frequencies like the one at hand. The new resultant signal has slower sample rate than the original signal. This could affect the accuracy of the experiment negatively, since the down sampling of the signal could lose some of the disease's indications that are held in high frequencies. All information and indications held beyond 200 Hz are lost, the down sampled signal shows only 200 frequency components. Not to mention that it would limit the number of levels that can be in the Filter Bank, since some sub bands would show only noise. However, the purpose of this research is to diagnose the phonocardiogram signal (PCG) with machine learning with limited computation load and processing power (mobile devices), down sampling serves this purpose.After which, a random noise is added to account for measurement noise, the total signal is used as an input to the model. The resultant signal at output of the model is a distorted version of the input. Figure 2 shows the original signal after downsampling and the resultant signal at the wrist.

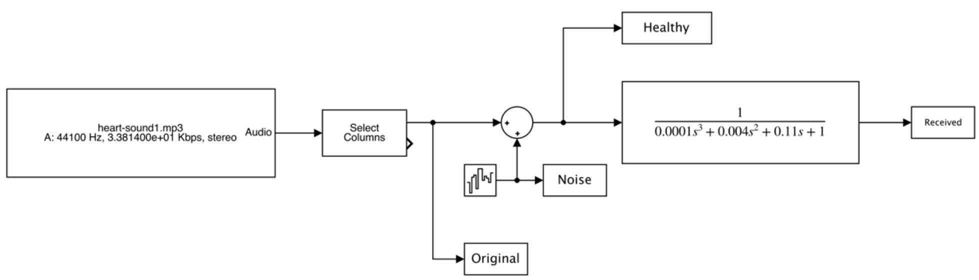

Figure 1. Heart-wrist acoustic propagation model

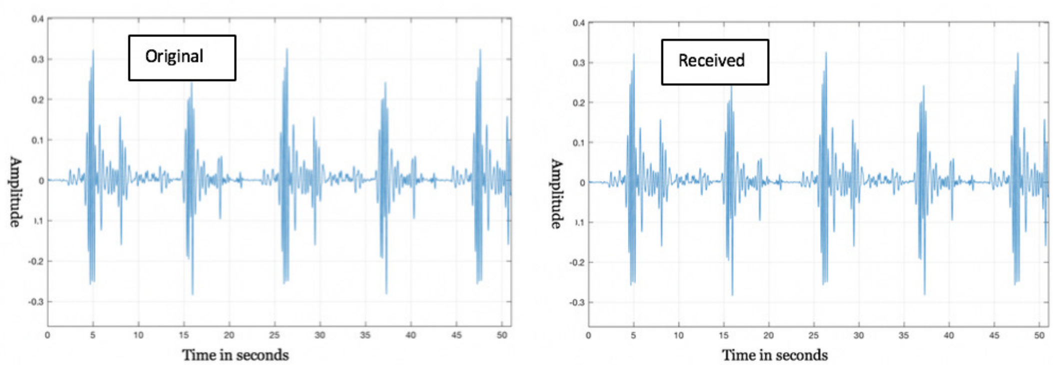

Figure 2. Original and received signal at the wrist

2.1. Generating the Hypotheses

The experiment starts by defining five hypotheses, each representing a heart disease, these hypotheses are used as reference classes for the diagnostic system. This is done by introducing a unique transfer function that alters the healthy heart acoustic signal to form a hypothetical disease, the resultant is then sent through the heart-wrist model (acoustic propagation transfer function) to be received as hypothesis x. And here is what makes these hypotheses valid - In simple words, an unhealthy heart is a heart that produces faulty PCG. When the cardiovascular system has a problem, the acoustic signal of the heart should reflect that in time/frequency space. This appears as corruption in the heart sound. Such change in the acoustic signals might not be audible or sensed by human ears (not even experienced cardiologist). However, it could be sensed and potentially classified with sensitive sensors and proper machine learning algorithms. A heart disease could be modelled as a linear corruption of signal of the healthy sound. This corruption has been performed by applying the healthy heart sound signal to different linear filters with different characteristics. There is no medical basis for selecting these filters during the simulation. However, this could be another research for modelling different heart problems with linear/nonlinear transfer functions. The diagnostic system's job is to identify this noise and consequently conclude the most probable condition.

3. THE EXPERIMENT

The experiment is to apply the proposed solution to classify 300 cases. All cases are generated from the hypotheses discussed in 2.1, below are the steps of the experiment.

3.1. Decomposing with filter bank

The signal is decomposed using Wavelet Transform and specifically Filter Banks [2] with different number of levels for every cycle of this experiment. The training and test sets were decomposed into a number of bands equivalent to 2x Filter Levels. For each band, the mean and covariance were calculated, so that each case is described by a number of features equivalent to 2x the number of bands. These features are concatenated in a 1 x the number of features matrix named "descriptive matrix". The number of levels was selected using trial and error as the method of optimisation. Starting the 1 level (2 bands), the experiment trials and compare results from 2 levels (4 bands), 3 levels (8 bands), 4 levels (16 bands), 5 levels (32 bands), and finally 6 levels (64 bands). Increasing the levels beyond 6 did not add any value to the classification nor did it improve the classification result, for that this experiment was stopped at 6 levels.

3.2. Training the system

Training the system in this approach is about using the training set to construct the discriminant function. By calculating the mean of the descriptive matrix to get a 1x128 vector and calculate the covariance of the descriptive matrix, which is a square matrix of 128x128. These values are used to construct the Discriminant Function (DF) [1] that is given by Equation 1

$$DF=-0.5 \log(\det(s))-0.5 (x-m)^{\prime}*\mathrm{inverse}(s)*(x-m) \tag{1}$$

Where s is the covariance matrix, m is the mean vector, and x is the training case. It certainly helped to add a confirmation step here that tests the DF using the training set, by simply calculating the DF for the training set and maximising the result, in order to confirm the validity of the training. This is a simple test; because if the training is valid the classification must be 100% correct. During the confirmation step, it became evident that the determinant of the covariance matrix is zero in many cases, which made the first component of the DF function (-infinity). For that, the discriminant function equation was rewritten as follow

$$DF=-0.5 (x-m)' * \mathrm{inverse}(s) * (x-m) \tag{2}$$

This formula was used when the filter levels reached 4 (16 bands, 32 features).

3.3. Testing the system

To test the system, the discriminant function calculated in step (2) is used to classify the test set. The classification was 97.33% successful using 8 features (2 levels), with only 4 false classifications. With 128 features (6 levels), the classification was 79.33% successful. Total of 31 cases out of the 150 cases were falsely classified, while the rest 119 were correctly classified.

4. DISCUSSION

After repeating the above steps for every all Filter Levels from 1 to 6, it was concluded that the best possible configuration for this scenario using this approach is to use a Filter Bank with 2 levels. This decomposes the signal into total of 4 bands and allows the signal to be described by 8 features. This is an interesting result because it defies the proposition behind this approach, that is more features will improve the success rate of the classification. Despite defying the proposition of the approach, this result is eye-opening. Decomposing signals is considered insightful, because it provides more descriptive details about the original signal. For example, a signal that is decomposed to 4 bands is more descriptive than one decomposed to 2 bands. Since the number of features is twice as much with 4 bands. All this makes sense and is well-known from the theory and concept of sub-band decomposition. What makes this result eye-opening, is that it argues against that. When studying the correlation coefficients between the cases with different decomposition levels, it was evident that the first band of every hypothesis correlate with the others, these bands are from a 6-level filter bank (64 bands and 128 features). This is visible in Figure 3-a, where the scatter plot shows that apart from some outliners, there is a positive linear correlation with moderate-strength across the first band of every hypotheses for majority of the values within these bands.As opposed to what can be seen in Figure 3-b, where the scatter plot shows little to no correlation across the first band of every hypotheses from a 2-level filter bank (4 bands and 8 features).

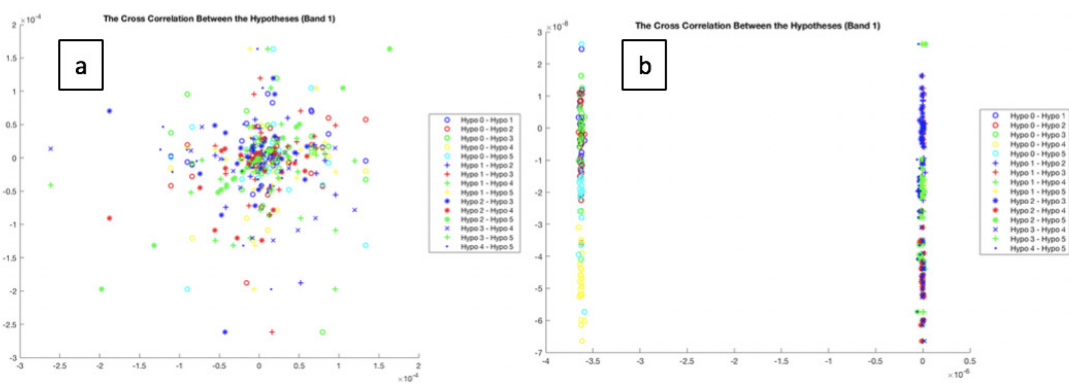

Figure 3. Correlation between subbands in different decomposition levels.

4. RELATED RESEARCH

The heart sound signal, Phonocardiogram (PCG) is difficult to interpret even for experienced cardiologists. Interpretation are very subjective depending on the hearing ability of the physician. Therefore, many researchers attempted to analyse and classify heart acoustic signals in order to diagnose the heart condition, especially using Wavelet Transform and many cases combined with Neural Networks. In 2016, team of two attempted building heart sound monitor using wearable wrist sensor [3]. Although they focused on the hardware, the mathematical model of the heart acoustic system was based on a similar model to what is used in this paper. they worked on finding the inverse function for the pulse wave between two locations along the same artery; chest and wrist and used that to estimate the recorded pulse wave at the chest from the one recorded at the wrist. Moreover, in 2017, wavelet transform, and a neural network were used to process and identify the heart condition using the heart sound signal obtained from a novel digital stethoscope [4].Digital stethoscope is an advancement in the field of heart sound diagnosis, because it overcomes the limitation of the acoustic stethoscope, such as; its reliance on the physician's hearing sensitivity and the impossibility of saving its sound into the patients' record. The novel stethoscope converts the analogue audio into digital signal, amplify, and low-pass filter the signal to produce an audible digital signal. The signal was decomposed using filter bank with 10-decomposition level, then a simple neural network with two layers and 75 neurons each, was used to identify the heart condition in question, the accuracy level varied 70% - 100% depending on the heart condition at hand, the research considered six conditions.And in 2018, CWT was used to classify heart sound recordings [5]. They built an automatic detecting system of the anomalies in the heart sounds, to get objective classification away from the subjectivity of the physician's hearing sensitivity. The phonocardiogram (discussed more in next chapter) obtained from Physionet database [6] was processed to extract features using adoptive segmentation, the features are then used along with k-nearest neighbour method to classify the heart sounds as normal and abnormal. Their results had high sensitivity, specificity, and accuracy.

5. CONCLUSION

The cardiovascular system is modelled in a transfer function to provide PCG signal recording as it would be recorded at the wrist. The signal is, then, decomposed using filter bank and the analysed using discriminant function. The results showed that PCG with a 19 dB Signal-to-Noise-Ratio can lead to 97.33% successful diagnosis. From the above discussion, it can be concluded that smaller bands tend to have stronger correlation across the hypotheses, which causes the classification to be more error prone. Such bands result from decomposing the signal using filter banks with large number of levels. This could have been caused by the down sampling, since it removed information held in high frequencies and left some bands with just noise. However, the trade-off was worthy, because the down sampling that was performed at the beginning of the experiment reduced the required computational load and processing power. Not to mention that it served the purpose of this research; limited energy and processing power (the use of mobile devices). Similar result had been noted in image texture recognition that filter banks with smaller number of levels performed better than larger ones [7].

REFERENCES

[1] Alpaydin, E. (2016). Introduction to Machine Learning. 3rd ed. ed. Cambridge: The MIT Press.

[2] Mertins, A. (1999). Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transforms and Applications. Chichester [u.a.]: Wiley.

[3] Shi, W.Y. and Chiao, J.-. (2016). Neural Network Based Real-Time Heart Sound Monitor using a Wireless Wearable Wrist Sensor. IEEE Available from: http://ieeexplore.ieee.org/document/7791150 Available from: http://dx.doi.org/ 10.1109/DCAS.2016.7791150.

[4] Suseno, J.E. and Burhanudin, M. (2017). The Signal Processing of Heart Sound from Digital Stethoscope for Identification of Heart Condition using Wavelet Transform and Neural Network. IEEE Available from: https://ieeexplore.ieee.org/document/8276354 Available from: http://dx.doi.org/ 10.1109/ICICOS.2017.8276354.

[5] BüşraKübraKaraca, et al. (July 2018). Classification of Heart Sound Recordings with Continuous Wavelet Transform Based Algorithm. 2018 26th Signal Processing and Communications Applications Conference (SIU) 2-5 May 2018. Turkey: IEEE Xplore.

[6] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/content/101/23/e215.full]; 2000 (June 13). PMID: 10851218; doi: 10.1161/01.CIR.101.23.e215

[7] Randen, T. (1997). Filter and Filter Bank Design for Image Texture Recognition. Norwegian University of Science and Technology Stavanger College.

AUTHORS

Shaima Tajalsir Abdelmageed Abdel Rahman is a Doctoral student at the University of Vasa, Finland. She has been a researcher in the field of mHealth since 2011.

Prof. Mohammed Elmusrati is full professor and head of communications and systems engineering group at the University of Vasa. He is also a visiting professor at Aalto Univeristy. Prof. Mohammed is also supervising Shaima’s Ph.D.