

Using Neural Networks for Heart Sounds Classification: A Review

Omer Deperlioglu Department of Computer Programming, University of Afyon Kocatepe, Afyonkarahisar, Turkey E-mail: deperlioglu@aku.edu.tr

Abstract

The high number of sudden deaths from heart diseases today makes early diagnosis of heart diseases very important in order to prevent these deaths. The first method that doctors use to diagnose heart disease is to listen to heart sounds caused by mechanical movements of the heart. This procedure is commonly referred to as auscultation of the heart. Over time, the irregularity of S1 and S2 sounds, which are components of heart sounds, may show some disturbances in heart function. For example, inadequate heart valves or murmurs. Early diagnosis of heart diseases is of great importance. However, not all doctors have the ability to interpret every heart sound and make an early diagnosis. Of course, the fact that their jobs are very busy and too many patients affect this. The status of each patient may be different during diagnosis and treatment. Assistive systems have been developed to enable doctors to diagnose patients with very different symptoms. These systems generally take diagnostics and evaluate all the data and suggest a diagnosis for the patient. In other words, these systems are decision-making mechanisms that classify all data. In the classification studies, researchers used different classification techniques in order to increase the classification success. Because the success of the classification means an increase in the diagnostic ability. For this purpose, many different techniques and mixed methods are used. In this study, a general evaluation of my classification studies of heart sounds with different neural networks such as Artificial Neural Networks (ANN). Convolutional Neural Network (CNN) and

networks such as Artificial Neural Networks (ANN), Convolutional Neural Network (CNN) and Autoencoder Neural Networks (AEN) was performed. PASCAL B-Training and Physiobank-PhysioNet A-Training data sets were used for the classification of heart sounds. The efficiency of the methods were compared by considering the common evaluation criteria in the studies examined.

Keywords: Heart sounds classification, Re-sampled signal energy, Autoencoder neural networks, Artificial Neural Network, Convolutional Neural Network.

DOI: 10.7176/JSTR/5-12-15

1. Introduction

According to the World Health Organization (WHO), approximately 30% of all deaths in the world are caused by heart diseases. Therefore, any method to detect symptoms of heart disease can have a significant impact on world health. Numerous cases of medical data, as well as very different diseases and symptoms, make it difficult to examine and investigate the data. Early diagnosis of heart diseases is of great importance. However, not all doctors have the ability to interpret every heart sound and make an early diagnosis. Of course, the fact that their jobs are very busy and too many patients affect this. The status of each patient may be different during diagnosis and treatment. Assistive systems have been developed to enable doctors to diagnose patients with very different symptoms. These systems generally take diagnostics and evaluate all the data and suggest a diagnosis for the patient. In other words, these systems are decision-making mechanisms that classify all data. In the classification studies, researchers used different classification techniques in order to increase the classification success. Because the success of the classification means an increase in the diagnostic ability. For this purpose, many different techniques and mixed methods are used (WHO, 2017; Deperlioglu, 2019a). The majority of classification procedures are performed to achieve better classification results. Some examples of the work done for the classification of heart sounds other than my own studies are given below.

Shankar and Sangeetha In their study using artificial neural networks for classification, S1 and S2 segmented heart sounds using discrete wavelet transform and Shannon energy envelope. They tried to

138 | P a g e www.iiste.org identify murmur sounds by classifying the S1 vS2 segments of the heart sounds they obtained. According to their results, segmentation contributes to the success of classification. (Shankar and Sangeetha, 2013). In another segmentation study, Bahekar and colleagues used discrete wavelet transform to segment the S1 and S2 sounds. They classified their segments using adaptive neural fuzzy inference system (ANFIS) (Bahekar, 2017). In another study using wavelet transform, Kumar and Alan tried to determine the order of abnormal sounds using a combination of Q-wavelet transform and a second difference signal with a median filter (Kumar and Saha, 2018). Choi and Jiang used Shannon energy envelope information and the Hilbert transform to fully evaluate the components of heart sounds. They stated that they achieved better results than traditional Shannon envelope and Hilbert envelope algorithms (Choi and Jiang, 2008). Saini proposed a method that automatically selects S1 and S2 heart sounds using a three-degree normalized Shannon energy envelope. This method can distinguish the components of S1 and S2 heart sounds by reducing the effect of other sounds. It was said that the noise in the heart sound was easily suppressed because the signal and envelope calculation was preprocessed. (Saini, 2016). In his Sharma study, he proposed a method of analysis using the Shannon energy envelope to study heart sounds and its rhythmic properties. Mean, variance and autocorrelation were used to find the characteristics of heart sounds. (Sharma, 2014). El-Segaier et al. used ECG signals as a reference for segmentation of S1 and S2 sounds. Used QRS complexes and T waves in ECG signals to find S1 and S2 segments (El-Segaier et al., 2005). T waves in some ECG signals may not be clearly selected. For these cases, Carvalho et al. used a new classifier in the selection of S2 sounds for low-quality ECG signals (Carvalho et al., 2005). Many researchers have tried to identify S1 and S2 sounds with several signal processing and statistical methods to reduce overload without using the ECG as a reference.

Different examples can be given from other approaches; Ali et al. used three different community techniques to increase classification success. These include: Bagging, AdaboostM1 and Random Subspace group techniques. They said that the community method increased the success of classification (Ali et al., 2017). Gharehbaghi and colleagues have proposed a new method in their work to distinguish between innocent and pathological murmurs using the growing time support vector machine. They compared the results with traditional support vector machines and stated that the growing time support vector machine increased the classification success (Gharehbaghia et al., 2017). Zhang et al. proposed a method based on a scaled spectrogram and tensor decomposition to improve classification accuracy. In this study, support vector machines were used for classification. They showed that the proposed method increases the success of classification (Zhang et al., 2017). Eslamizadeh and Barati used Artificial Neural Network to classify heart sounds as normal and murmur. They used the Modified Neighbor Annealing method for the training of artificial neural networks. They showed that the modified Neighbor Annealing method gives good results for heart sounds classifiers (Eslamizadeh and Barati, 2017).

As given in the examples above, many methods have been tried and applied for the classification of heart sounds. In this review, the results of previous studies on the classification of heart sounds using artificial neural networks and deep neural networks will be discussed.

2. Material and Methods

The block diagram of the processes performed in the investigated studies is given in Figure 1. As can be seen from the figure, prior to the classification process, in general, the heart sound signals are preprocessing and, after this, these signals are filtered with the elliptical filter. Then, if there is a segmentation, the signal energies of all sound files are calculated by using the resampled energy method and segmentation is performed. These operations will be briefly explained in the following sections respectively. International Journal of Scientific and Technological Research ISSN 2422-8702 (Online), DOI: 10.7176/JSTR/5-12-15 Vol.5, No.12, 2019



Figure 1. The processes block diagram of Classification operations.

2.1. Heart Sounds Datasets

PASCAL B-Training and Physiobank-PhysioNet A-Training data sets were used in the classification of heart sounds.

2.1.1. PASCAL B-Training dataset

The sound files in the PASCAL B-Training dataset are in wav format and obtained from a clinical trial using electronic stethoscope DigiScope® during clinical examinations in hospitals (Bentley et al., 2011). From the Btraining dataset, a total of 449 files of 3 different classes, normal, extrasystole and murmur, were selected. Table 1 presents the general characteristics of the heart sound files.

2.1.2. PhysioNet a-training data set

Heart sounds in the PhysioNet a-training dataset are in wav format. These sound files were obtained from clinical trials in hospitals. The dataset contains both healthy and pathological records. Persons with heart sound recordings are children and adults. One to six heart sounds were recorded from each patient. The duration of recordings varies from a few seconds to a hundred seconds. All recordings were set to 2000 Hz and recorded in wav format. [Chengyu et al., 2016; Goldberger et al., 2016]. The dataset contains 409 files belonging to 2 different classes, normal and abnormal. Table 2 presents the general characteristics of the heart sound files.

Sound File Type	Duration Sampling Frequency		Number of Files	
Normal	8 second	4000 Hz.	196	
Noisy Normal	8 second	4000 Hz.	118	
Extrasystole	8 second	4000 Hz.	47	
Murmur	8 second	4000 Hz.	61	
Noisy Murmur	8 second	4000 Hz.	27	
Total			449	

Table 1. Characteristics of sound files in the PASCAL B-training dataset.



Sound File Type	Duration	Sampling Frequency	Number of Files
Normal	15 second	2000 Hz.	294
Abnormal	15 second	2000 Hz.	115
Total			409

Table 2. Characteristics of sound files in the PhysioNet A-training data set

2.2. Signal Pre-Processing

The preprocessing step of heart sounds usually involves three procedures, such as resampling, normalization and filtering. At the stage of classification, all heart sounds should have the same sampling frequency. Also, all sounds files must have the same length. Thus, the matrix dimensions of learning dataset are equal. For this purpose, the lengths of the sound files are set to 8 seconds for the Pascal dataset and 15 seconds for the PhysioNet dataset. Since the records for each datasets have the same sampling frequency, no action was taken in the resampling stage. Then, normalization was applied to all sound files and elliptical filtering was performed.

2.3. Segmentation

In segmentation classification studies, after filtering, segmentation was performed using the resampled energy method (Deperlioglu, 2018a). Sounds S1 and S2 were selected using the intervals generated by the signal energy. A short-term instant energy calculation is used to calculate the energy the sound signal has at a given time. In these studies, first the energy calculation was made and then the segmentation of S1 and S2 sounds were performed. All applications were performed in MATLAB environment.

2.4. Classification

Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Auto-encoder Neural Network (AEN) were used for the classifications of heart sounds.

The Artificial Neural Network is an effective data processing system that basically emerges from the simulation of biological neural networks. ANN creates a large array of networks connected in a specific way to communicate between units. These units, also called nodes or neurons, are simple parallel processors that operate in parallel. A neural network is a connected combination of simple processing elements, units, or nodes that loosely operate based on an animal's neuron. The ability to process the network is recorded with a number of training examples in the combination of weight obtained with the learning process. A feedforward network is a non-repeating network with processing units or nodes on the layer, and all nodes on the layer are linked to the nodes of the previous layers. The connection has different weights. There is no feedback loop, the input signal can only be transferred in one direction. The multilayer feedforward network is a feedforward ANN with multiple weight layers. There are one or more layers between the input and output layers, which are called hidden layers (Gurney, 2014; Point, 2017).

After ANN, the opinion of creating ideas about the task in the minds of researchers before uncontrolled training has been used in different ways and in different ways since the mid 1980s. However, Hinton et al. have systematically used this idea to demonstrate that deep neural networks can be successfully trained (Hinton and Salakhutdinov, 2006). They used the limited productive model, formerly called the Boltzmann machine. Later, a simple structure called Autoencoders, which resembles feedforward neural networks, was introduced (Choi, 2019). Recently, it has been further developed with a supervised learning phase with a deep architectural approach. Thus, it has been shown that it can produce very easy solutions to difficult classification and regression problems in state-of-the-art products requiring artificial intelligence (Baldi, 2012). (Baldi, 2012). Autoencoder neural networks can be thought of as feed-forward neural networks are trained as reliable copies of inputs. However, in network architecture, the input layers are called encoders and the output layers are called decoders. Intermediate layers that contain fewer neurons are called hidden layers. Thus, AEN may generally be considered as a pair of encoders and decoders associated with at least a number of neurons, including the coded representation of network entries. Since each layer depends only on the values of the previous layer, it is sufficient to calculate the reconstruction corresponding to the coding information (Valentine and Trampert, 2012).

A CNN generally has a multi-layered structure, such as feedforward networks. Unlike feedforward networks, it can consist of several convolute layers with a sub-sampling section. In this multi-layered structure, the layers are completely interconnected. CNN is designed to capture two-dimensional images and handle them easily. Provides local links and variable weights for easy processing of images. Thus,

CNNs may have the ability to produce less educated parameters (ufldl.stanford, 2018). A Convolutional Neural Networks generally consist of one or more convolute layers, which are a sub-sampling step. It is then composed of one or more fully connected layers, such as a standard multilayer neural network. A CNN architecture is designed to take advantage of the 2D structure of two-dimensional (2D) image inputs. This is summed up by local connections and associated weights, leading to later-changing properties. Another benefit of CNNs is that they have fewer training and easier parameters than networks that are fully connected to the same number of hidden nodes. In addition, CNNs use spatial-local correlation by applying a local binding pattern between neurons of adjacent layers (LeNet, 2018).

2.5. Performance Evoluation

Accuracy, sensitivity and specificity are commonly used performance criteria in medical classification studies. These criterias have been used to assess the accuracy of the proposed method. These are calculated as follows:

$$Accuracy = \frac{T_P}{T_P + F_P} \tag{1}$$

$$Sensitivity = \frac{T_P}{T_P + F_N}$$
(2)

$$Specificity = \frac{T_N}{F_{PTN} + T_N}$$
(3)

In the equations, T_P and F_P represent the number of true positive and false positive, respectively. T_N and F_N are true negative and false negative numbers, respectively. F_{PTN} also represents the number of false positives and is calculated from negative samples in the classification results.

The accuracy of the classifier's correct diagnostic capability is determined by its accuracy. The degree to which the model accurately determines the formation of the target class is defined by the sensitivity ratio. The scope of the model's ability to target class is defined by the specificity ratio (Deperlinglu, 2018b).

3. Classifications of Heart Sounds

In the study conducted with Artificial Neural Networks (ANN), S1-S2 sounds were segmented using the resampled energy method and the contribution of segmentation to classification performance was examined. PASCAL B-training dataset, which is widely used in cardiac diseases, was used in the application. Feedforward Artificial Neural Network was used to classify these sounds. Bayes regularization algorithm was used to train artificial neural networks from previous experiences (Deperlioğlu, 2018d). In order to compare the obtained results, two different classifications were made for segmented and non-segmented sounds (Deperlioğlu, 2018c).

In another study conducted with ANN, the usage of signal energy used as a sub-method of segmentation in direct classification was investigated. For this purpose, energies of heart sounds were calculated with the resampled energy method and the data were classified with artificial neural networks. In this study, 192 heart sounds selected from PASCAL B-training heart sounds dataset were used. In particular, the effect of sampling time on classification success was investigated. The obtained results were compared in itself and S1-S2 sounds with segmented classification. In the direct classification with the resampled energy method, the accuracy of the classification increased as the sampling rate decreased. In addition, it was found that the classification accuracy rate was higher than the classification with the segmented dataset (Deperlioğlu, 2018, May).

Another one is the classification of phonocardiograms of heart sound with convolutional neural networks which is one of the deep learning methods. In this study, three different heart sound data such as normal, murmur and extrasystole were classified in PASCAL B-training dataset. Phonocardiograms (PCG) obtained from heart sounds in the dataset were used for classification. In this study, first 8-second heart sounds samples were taken. The sound signals are normalized during the preprocessing phase. After preprocessing, the signals was filtered with elliptical filter. Then the PCG of each signal was generated. The image size of the PCGs was resized to 100x75 pixels to shorten the learning phase in the classification. At the end of the study, in order to compare the success of classification, it was classified with both ANN and CNN for the same data set (Deperlioglu, 2018b).

In another study with CNN, heart sounds in the PhysioNet A-training dataset, which had two different heart sounds data such as normal and abnormal, were classified by CNN. For this purpose, the S1 and S2 parts of the heart sounds were segmented by the resampled energy method. Phonocardiogram images

obtained from S1 and S2 sections in heart sounds were used for classification. Small resized phonocardiogram images were classified with CNN (Deperlioglu, 2019b).

In another study, the classification of segmented heart sounds by an autoencoder neural network (AEN) is described. Three different heart sounds data such as normal, murmur, and extrasystole in the PASCAL B-training dataset were classified. For this purpose, first of all, normalization of heart sounds was performed. After this, filtration was carried out with an elliptical filter. Subsequently, segmentation was performed with the resampled energy method. Finally, the classification was made with AEN. As a result, segmentation of the main components of heart sounds S1-S2 sounds and classification of AEN have been proposed for a high-performance classification (Deperlioglu, 2019c).

The results obtained from all these studies are collected in Table 3 for easier evaluation.

Order No	Paper	Method and Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	Deperlioğlu, 2018c Segmented	ANN PASCAL	84.0	86.5	86.9
2	Deperlioglu [20] Non- segmented	ANN PASCAL	88.60	89.25	88.00
3	Deperlioğlu, 2018, May Segmented	ANN PASCAL	90,6	93,5	95,7
4	Deperlioğlu, 2018b	ANN PASCAL	82.80	92.40	88.82
5	Deperlioğlu, 2018b	CNN PASCAL	97.90	99.47	98.42
6	Deperlioglu, 2019b Non-segmented	CNN PhysioNet	96.28	93.91	98.65
7	Deperlioglu, 2019b Segmented	CNN PhysioNet	97.21	94.78	99.65
8	Deperlioglu, 2019c Segmented	AEN PASCAL	99.93	99.77	99.77

Table 3. Comparison of heart sounds classification data

In the studies performed with Artificial Neural Networks, the highest accuracy rate was obtained by classifying the energies of the heart sounds with 90.6. In the studies conducted with deep learning methods, the highest accuracy rates were obtained in the study with AEN. These values are shown in bold in the table.

4. Conclusion

Many methods have been tried and applied for the classification of heart sounds. In this study, the results of previous studies on the classification of heart sounds using artificial neural networks and deep neural networks were discussed. In the studies examined, heart sound signals were preprocessing before the classification operation and after this process was completed, elliptical filter was applied. Then, if there is segmentation, the instant signal energies of all sound files are calculated by using the resampled energy method and segmentation is performed. PASCAL B-Training and Physiobank-PhysioNet A-Training data sets were used in the classification of heart sounds. Heart sounds in these data sets were classified in different ways by ANN, CNN and AEN.

In the studies conducted with ANN, Bayesian regulation education algorithm gave better results than other learning algorithms for medical data. In addition, another result obtained from these studies is that segmentation contributes to the success of classification. As seen in Deperlioğlu, 2018c, when the difference between the classification with non-segmented data and the classification with segmented data is compared, it is seen that segmentation contributes 4% to the success of classification.

On the other hand, in the classifications made with CNN and AEN which are deep learning methods, it is seen that AEN is more successful in the classification of heart sounds. Moreover, it is seen that segmentation does not contribute much when deep learning methods are used. It is thought that the same success can be achieved by making different arrangements in classifier instead of wasting time with segmentation.

Another prominent detail is that deep learning methods process faster than normal ANNs and save time.

143 | P a g e www.iiste.org Even though the learning process takes some time in CNNs because they process images, AEN processes faster than both ANN and CNN.

References

- Ali, S., Adnan, S. M., Nawaz, T., Obaid Ullah M., Aziz, S. (2017), Human Heart Sounds Classification using Ensemble Methods, Technical Journal, University of Engineering and Technology (UET), 22(1), 113-120.
- Bahekar, L., Mishal, A., Bisen, M., Koche, D., & Alone, S. (2017). Heart Valve Diseases Detection Using Anfis and Wavelet Transform. International Journal of Research In Science & Engineering, 3(2), 279-291.
- Baldi, P. (2012, June). Autoencoders, unsupervised learning, and deep architectures. In Proceedings of ICML workshop on unsupervised and transfer learning (pp. 37-49).
- Bentley P., Nordehn G., Coimbra M. and Mannor S. The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results, http://www.peterjbentley.com/heartchallenge/index.html. [Last access: 30 July 2019]
- Carvalho, P., Gilt, P., Henriques, J., Eugénio, L., & Antunes, M. (2005, September). Low complexity algorithm for heart sound segmentation using the variance fractal dimension. In IEEE International Workshop on Intelligent Signal Processing, 2005. (pp. 194-199). IEEE.
- Chengyu L., David S., Qiao L., Benjamin M., Ricardo Abad J., Francisco JC et al., An open access database for the evaluation of heart sound algorithms, Physiological Measurement, (2016) 37: 21-81.
- Convolutional Neural Networks (LeNet) http://deeplearning.net/tutorial/lenet.html. [Last access: 10 January 2018]
- Deperlioglu, O. (2018a). Segmentation of heart sounds by re-sampled signal energy method. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 9(1), 17-28.
- Deperlioglu, O. (2018b). Classification of phonocardiograms with convolutional neural networks. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 9(2), 22-33.
- Deperlioğlu, Ö. (2018c). Classification of segmented heart sounds with Artificial Neural Networks. International Journal of Applied Mathematics, Electronics and Computers, 6(4), 39-44.
- Deperlioglu, O. (2018d). The Effects of Different Training Algorithms on the Classification of Medical Databases Using Artificial Neural Networks, European Conference on Science, Art & Culture ECSAC 2018, Antalya, Turkey.
- Deperlioğlu, Ö. (2018, May). Classification of heart sounds with re-sampled energy method. In 2018 26th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.
- Deperlioglu, O. (2019a). Intelligent Techniques Inspired by Nature and Used in Biomedical Engineering. In Biotechnology: Concepts, Methodologies, Tools, and Applications (pp. 666-692). IGI Global.
- Deperlioglu, O. (2019b). Classification of Segmented Phonocardiograms by Convolutional Neural Networks. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 10(2), 5-13.

Deperlioglu, O., (2019c). Classification of Segmented Heart Sounds with Autoencoder Neural **144** | P a g e

Networks" VIII. International Multidisciplinary Congress of Eurasia (IMCOFE'2019), ISBN: 978-605-68882-6-7, pp. 122-128, April 24 to 26, 2019, Antalya.

- El-Segaier, M., Lilja, O., Lukkarinen, S., Sörnmo, L., Sepponen, R., & Pesonen, E. (2005). Computer-based detection and analysis of heart sound and murmur. Annals of Biomedical Engineering, 33(7), 937-942.
- Eslamizadeh, G., Barati, R. (2017) Heart murmur detection based on wavelet transformation and asynergy between artificial neural network and modified neighborannealing methods. Artificial Intelligence in Medicine, Volume: 78, pp. 23–40.
- Gharehbaghia, A., Borgaa, M., Sjöbergc, B. J., Ask, P. (2017) A novel method for discrimination between innocent and pathological heart murmurs. Medical Engineering and Physics, Volume: 37, 674–682.
- Goldberger A. L., Amaral L. A. N., Glass L., Hausdorff J. M., Ivanov P. C., Mark R. G., Mietus J. E., Moody G. B., Peng C.-K., Stanley H. E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23): 215-220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 [Last access: 30 July 2019].
- Gurney, K. (2014). An introduction to neural networks. CRC press.
- Hinton, G. E. (2016) Neural Networks for Machine Learning, Lecture 15a. http://www.cs.toronto.edu/~bonner/ courses/2016s/csc321/lectures/lec15.pdf, Last Access: [Last access: 23 January 2019].
- Kumar, A. K., & Saha, G. (2018). Improved computerized cardiac auscultation by discarding artifact contaminated PCG signal sub-sequence. Biomedical Signal Processing and Control, 41, 48-62.
- Saini, M. (2016). Proposed Algorithm for Implementation of Shannon Energy Envelope for Heart Sound Analysis. International Journal of Electronics & Communication Technology IJECT, 7(1), 15-19.
- Shankar, N., & Sangeetha, M. S. (2013). Analysis of phonocardiogram for detection of cardiac murmurs using wavelet transform. International Journal of Advanced Scientific and Technical Research, 1(3), 350-357.
- Sharma, P. K., Saha, S., & Kumari, S. (2014). Study and design of a Shannon-energy-envelope based phonocardiogram peak spacing analysis for estimating arrhythmic heart-beat. International Journal of Scientific and Research Publications, 4(9), 1-5.
- Point, T. (2017). Artificial Neural Network, Tutorials Point (I) Pvt.
- Valentine, A. P., & Trampert, J. (2012). Data space reduction, quality assessment and searching of seismograms: autoencoder networks for waveform data. Geophysical Journal International, 189(2), 1183-1202.
- WHO, Cardiovascular Diseases, Fact Sheet, World Health Organization, 2017. http://www.who.int/mediacentre/factsheets/fs317/en/ [Last access: 24 June 2019]
- Zhang, W., Han, J., Deng, S. (2017) Heart sound classification based on scaled spectrogram and tensor decomposition. Expert Systems with Applications, Volume: 84, pp. 220–231.