Classification of Segmented Phonocardiograms by Convolutional Neural Networks

Omer Deperlioglu
Afyon Kocatepe University, Vocational School of Afyon
Erenler Mahallesi, Gazlıgöl Yolu Rektörlük E Blok, 03200
Afyonkarahisar Merkez/Afyonkarahisar, Turkey
Phone: +90 272 228 13 50
deperlioglu@gmail.com

Abstract

One of the first causes of human deaths in recent years in our world is heart diseases or cardiovascular diseases. Phonocardiograms (PCG) and electrocardiograms (ECG) are usually used for the detection of heart diseases. Studies on cardiac signals focus especially on the classification of heart sounds. Naturally, researches generally try to increase accuracy of classification. For this purpose, many studies use for the segmentation of heart sounds into S1 and S2 segments by methods such as Shannon energy, discreet wavelet transform and Hilbert transform. In this study, two different heart sounds data in the PhysioNet Atraining data set such as normal, and abnormal are classified with convolutional neural networks. For this purpose, the S1 and S2 parts of the heart sounds were segmented by the resampled energy method. The images of Phonocardiograms which were obtained from S1 and S2 parts in the heart sounds were used for classification. The resized small images of phonocardiogram were classified by convolutional neural networks. The obtained results were compared with the results from previous studies. The classification with CNN has performance as classification accuracy of 97.21%, sensitivity of 94.78%, and specificity of 99.65%. According to this, CNN classification with segmented S1-S2 sounds showed better results than the results of previous studies. In studies carried out, it has been seen that segmentation and convolutional neural networks increases the accuracy of classification and contributes to the classification studies efficiently.

Keywords: Deep Learning; Convolutional Neural Networks; Heart Sounds Segmentation; Re-Sampled Signal Energy; Heart Sounds Classification.

1. Introduction

The early diagnosis of abnormal heart disease is vital to preventing sudden cardiac death by identifying heart conditions. An auscultation technology is a fast, widely used, non-invasive method for the assessment of cardiac functions by listening to heart sounds. Abnormal heart sounds indicate malfunctions in heart function, such as inadequate heart valves or murmur. A cardiac auscultation can give clues to many cardiac abnormalities, and at the same time is a traditional way of distinguishing heart murmurs from heart sounds. Heart sounds or phonocardiograms (PCG) may also provide additional diagnostic tests for advanced medical assessments. However, obtaining accurate diagnoses from a cardiac auscultation depends on the knowledge and experience of doctors. Doctors who do not have sufficient knowledge and experience may not be able to diagnose correctly. For this, there is a need for a clinic decision support system that will help doctors to assess the past cases and correct diagnoses, along with the sounds heard in cardiac auscultation. For this purpose, there are many studies including classification and segmentation (Deperlioglu, 2018; Zheng et al., 2015; Vafaie et al., 2014; Zabihi et al., 2016; Deperlioglu, 2018b).

Identification of precise locations of the S1 and S2 sounds in the PCG's segments is an important process in the analysis of heart sounds. Thus, the classification of pathological events can be made more productive. A correct detection of heart sounds, identification of systolic or diastolic regions of a PCG is very important for the diagnosis of heart diseases. Thus, the classification of pathological murmurs can be made by taking advantage of them (Springer et al., 2016). The approaches of signal analysis of PCGs are based on temporal segmentation. They define cardiac

cycles and point the location of S1, systolic start and S2, systole end primary heart sounds. Changes in S1 and S2 sounds durations and their intensity are explained as definitive indications of heart diseases (Zabihi et al., 2016).

There are many studies to improve the classification accuracy of heart sounds using different classification algorithms and segmentation with different envelope extraction methods. Puri et al. used Hidden Markov Model based on Springer's improved version of Schmidt's method to segment S1 and S2 sounds. Then they made feature selection and classification with support vector machines (Puri et al., 2016). In another study, cardiac sounds were classified using convolutional neural networks with the same segmentation method (Rubin et al., 2016). Tschannen and his colleagues used a hidden semi-Markov model and Viterbi decoding to segment the heart sounds and used a wavelet-based deep convolutional neural network to classify the heart sounds (Tschannen et al., 2016). Ryu et al. classified segmented heart sounds with convolutional neural network (CNN). Yang et al. segmented in to the S1-S2 sounds with peak-finding method and then classified them support vector machines (Ryu et al., 2016). The provided state of the art PCG segmentation algorithm and the bagging Meta algorithm with decision trees also used to analyze heart sounds (Yazdani et al., 2016). In the Deperlioglu study, the classification accuracy in practice with S1 and S2 sounds segmented by the resampled energy method is increased (Deperlioglu, 2018c).

As shown in the examples above, segmentation of heart sounds and different classification algorithms have been used to improve the classification performance. The purpose of this study is to show that convolutional neural networks can be used to classify segmented S1 and S2 sounds, and that higher classification success can be achieved with deep learning methods to enhance classification success. The phases of the study are explained in detail in the following sections.

2. Material and Methods

In this study, a deep learning method is proposed to increase the classification success. First, the heart sound samples in the PhysioNet atraining data set were normalized and filtering. After preprocessing, S1 and S2 sounds of each signal are segmented by the resampled energy method. PCG images were generated for each obtained segment. To reduce the amount of computer memory required during learning, the sizes of these images were reduced from 520x460 pixels to 100x75 pixels. PCG images are classified by convolutional neural networks. The block diagram of the constructed processes is given in Figure 1. These processes are explained in detail in the following sections respectively.

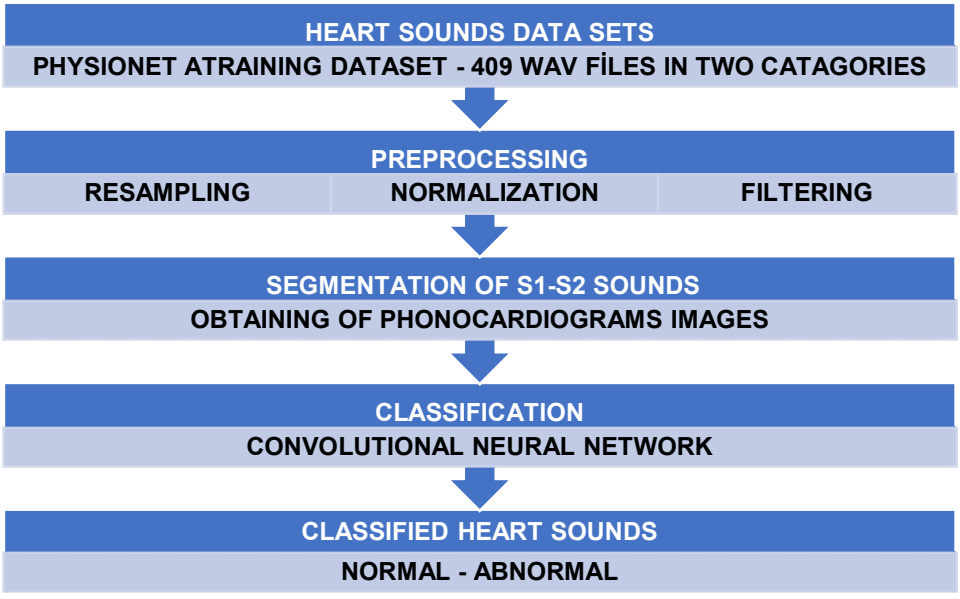


Figure 1. Block diagram of the classification of segmented heart sounds

2.1. PhysioNet Training Data Set

There are both healthy and pathological records in the data set. The persons whose heart sounds recording consist of children and adults. One to six heart sounds recorded from each patient. The duration of the recordings varies from a few seconds to a hundred seconds. All recordings were resampled to 2000 Hz and recorded in .wav format. (Chengyu et al., 2016; Goldberger et al. 2016). The dataset contains 409 files in 2 categories such as normal, and abnormal. Table 1 shows the general characteristics of the heart sounds files.

Table 1. The general characteristics of the heart sounds files in PhysioNet Dataset

Categories	**Sampling Frequency**	**Number of Files**
Normal Category	2000 Hz.	294
Abnormal Category	2000 Hz.	115
Total		**409**

2.2. Preprocessing

Because heart sounds are recorded with different devices in different environments, they can have different sampling frequencies. This requires resampling at the pre-processing stage. In addition, there are some noises during recording due to environmental sounds and body sounds. These noises should be removed from heart sounds. All heart sounds in the data set used in this study were re-sampled at 2000 Hz. For this reason, no resampling has been done. Only the normalization was done, then the noise was removed with the Elliptic filter.

2.3. Segmentation of S1-S2 Sounds and Obtaining of Phonocardiograms Images

The energies of heart sound signals were calculated using the resampled energy method. A short-time energy account is used to calculate the energy given by the sound signal at a given time. For this reason, square total energy function is used with resampling of the filtered signal. The signal energy can be calculated as Equation 1 (Deperlioglu, 2018d).

$$E = \sum_{i=1}^{N} x(i)^2 \qquad (1)$$

Figure 2 shows the energy image obtained by filtering with the original signal and elliptic filter. As seen in the energy graph, the S1 sounds in the heart sound correspond to the large triangles, and the S2 sounds correspond to the small triangles. In this context, segmentation of S1 and S2 sounds was performed by taking advantage of the base lengths of the triangles in the energy curves. The phonocardiogram images of the obtained segmented heart sounds were created with a size of 520x460 pixels. To reduce the amount of required computer memory during learning, the sizes of these images were reduced from 520x460 pixels to 100x75 pixels. Obtained PCG images were classified with convolutional neural networks.

2.4. Classification

Convolutional Neural Networks (CNN) is an alternative neural network type that can be used to reduce spectral variations and spectral correlation models in sound signals. CNNs are a more effective model for heart sound signals than other Deep Neural Networks (DNNs) because heart sound signals carry both features (Sainath et al., 2013).

2.4.1. Convolutional Neural Network

A CNN has a multi-layered structure, such as feed forward networks in general. Unlike feed forward networks, it can be composed of several convolutional layers with a sub-sampling section. In this multi-layer structure, the layers are completely interconnected. CNN is designed to take two-dimensional images and easily handle them. It provides local links and their varying weights for

easy processing of images. Thus, CNNs can have the ability to produce parameters with less training (ufldl.stanford, 2018).

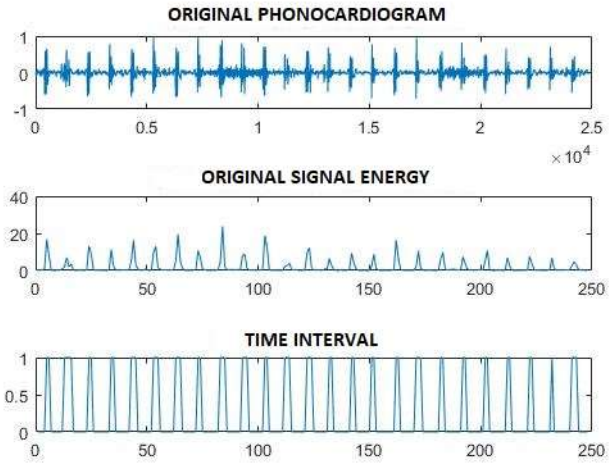


Figure 2. Example scheme of resampled energy

CNNs are structurally similar to multi-layered perceptron networks. CNNs use a local correlation with relative links by applying a local linking model between nodes in neighboring layers. That is, entries of hidden nodes in a layer are made up of a subset of the related nodes in the previous layer. Thus, bringing so many layers together results in non-linear spiral filtering. On a CNN, each filter scans the input images. These nodes form feature maps using their shared weights and bias. Gradient methods can be used to learn shared weights and bias.

Figure 3 shows an example of a convolutional layer. In the hidden layer *m*, there are two feature maps, h^0 and h^1. The outputs of the nodes in h^0 and h^1 are calculated from the outputs of the nodes entering the 2x2 receive area on the layer beneath the *m-1* layer. Thus, the weights are W^0, and W^1 *and* h^0 and h^1 are 3D weight tensors. The mean dimension summarizes the map entries summaries, pointing to the other two neuron outputs. To combine all of them, W_{ij}^{kl}, with each pixel of the layer's minus *k* map, it refers to the weight of the link between neuron outputs (*i, j*) in the coordinates of the feature map *m-1* (deeplearning.net, 2018).

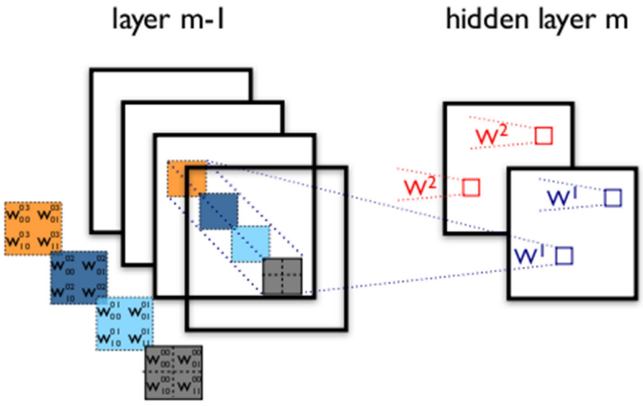


Figure 3. An example of a convolutional layer (deeplearning.net, 2018)

2.5. Performance Evaluation

An accuracy, a sensitivity and a specificity are commonly used performance measures in medical classification studies. These measures were used to assess the precision of the proposed

method (Gharehbaghia et al., 2017; Zhang et al., 2017). The PhysioNet/Computing in Cardiology (CinC) Challenge 2016 suggests public heart sound database (Clifford et al. 2016). The scoring algorithm was that originally proposed by the Challenge organizers defined as the average of specificity (Sp) plus sensitivity (Se) as follow:

$$score = \frac{S_e + S_p}{2} \tag{2}$$

where

$$S_e = \frac{A_a}{A_a + A_n} \tag{3}$$

$$S_p = \frac{N_n}{N_a + N_n} \tag{4}$$

In the equations, A_a, N_n represent correctly classified abnormal and normal recordings respectively and A_n and N_a represent incorrectly classified abnormal and normal recordings respectively (Langley and Murray, 2016).

3. Classification of Heart Sounds

The segmented heart sound images were classified by convolutional neural networks (CNN). CNN is a type of deep learning that is particularly suitable for image recognition and classification. CNNs receive and process image data as objects with the label. In this application, CNN has 8 layers. These layers are composed of an image input layer, a convolutional layer, a ReLU layer, a cross channel normalization layer, a max pooling layer, a fully connected layer, a softmax layer, and a classification layer. The details of CNN structure were given in Figure 4. In all classification study, MATLAB r2017a software was used.

In order to evaluate the results, non-segmented phonocardiograms and S1-S2 segmented phonocardiograms were classified by a CNN. Both classification studies were conducted for the PhysioNet atraining data set. From a total of samples files, 80% samples were used for training data, 20% samples were used for testing data. The CNN neural network was retrained and run 20 times with the different training and test data sets. The average accuracy of classification of non-segmented phonocardiogram images with the CNN was 96.28%. The sample confusion matrix of classification is given in Figure 5. The overall accuracy of classification of S1-S2 segmented phonocardiogram images with CNN was 97.21%. The sample confusion matrix of classification of S1-S2 segmented is given in Figure 6.

Image Input Layer	*409 .jpg images* *The image sizes of 75x100x3*
Convolutional Layer	*(4,16)* *The filter size is 4* *The number of filters 16*
ReLU Layer	*A nonlinear activation function* *The rectified linear unit function*
Cross Channel Normalization Layer	*The size of the channel window 2*
Max Pooling Layer	*The size of the rectangle is [4,3]* *The step size function Stride*
Fully Connected Layer	*The number of classes in the target table (2)* *Weight Learn Rate Factor 20* *Bias Learn Rate Factor 20*
Softmax Layer	*The softmax activation function for classification*
Classification Layer	*The Training Options* *The stochastic gradient descent with momentum (sgdm)* *The maximum number of epochs at 100* *The Initial learning rate of 0.00001*

Figure 4. The details of CNN structure

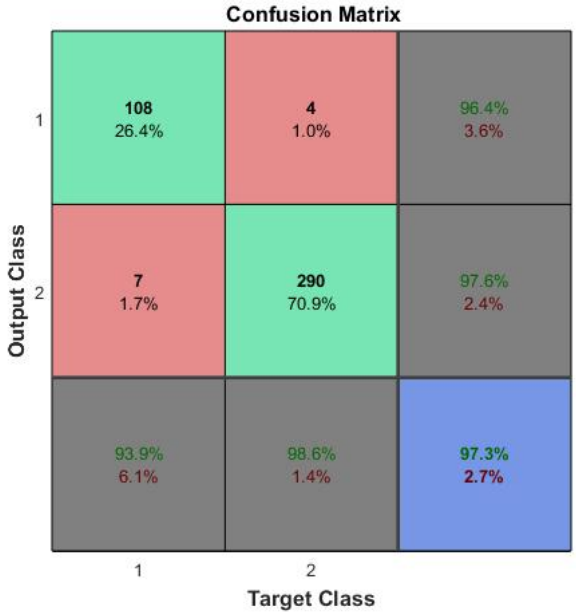


Figure 5. The sample confusion matrix of classification of non-segmented phonocardiograms

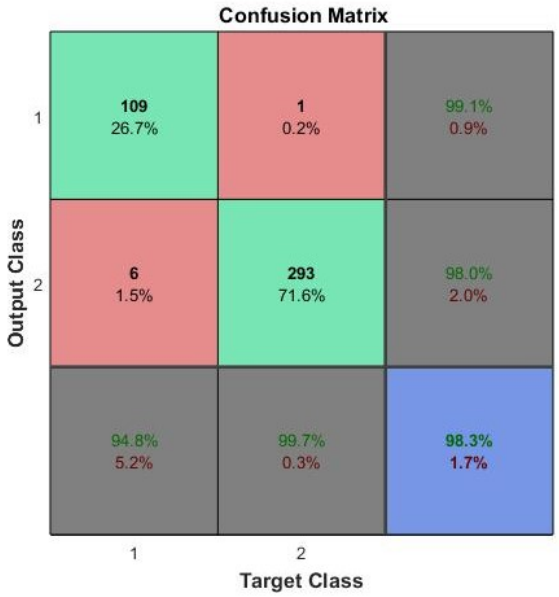


Figure 6. The Sample confusion matrix of classification of S1-S2 segmented phonocardiograms

4. Results and Discussion

When the performance evaluation is done; the average score for non-segmented PCGs was 96.28%, sensitivity was 93.91%, and specificity was 98.65%. The average score for segmented PCGs was 97.21%, sensitivity was 94.78%, and specificity was 99.65%. To compare with previous studies, PhysioNet / Computing in Cardiology (CinC) Challenge 2016 also benefited from the results of different methods of study with the same data set. The methods used in these studies and the results obtained are given in table 2. The best results in this table are shown as bold.

Table 2. The comparison data of the previous heart sound classification works

Classification method	Overall score (%)	Sensitivity (%)	Specificity (%)
Convolutional neural networks (non-segmented-this study)	96.28	93.91	98.65
Convolutional neural networks (segmented -this study)	**97.21**	**94.78**	**99.65**
Feedforward neural networks (Zabihi, 2016)	91.50	94.23	88.76
Deep Gated RNN (Thomae, 2016)	55.0	99.0	11.0
CNN (Nilanon 2016)	81.3	73.5	89.2
Wavelet Entropy (Langley and Murray, 2016)	77	98	56
Wavelet-based deep CNN (Tschannen, 2016)	82.8	85.5	85.9
Support Vector Machines (Yang et al., 2106)	83	70	96
Convolutional neural networks (Ryu , 2016)	79.5	70.8	88.2
Tape-long features (Yazdani et al., 2016)	70	90	80
ANN with Markov Feats (Vernekar, 2016)	81.75	79.2	84.3
Support Vector Machines (Puri, 2016)	78.20	77.49	78.91
CNN (Puri, 2016)	75	100	88
AdaBoost and CNN (Potes, 2016)	94.24	77.81	86.02

As seen in the table, the highest overall score, sensitivity, and specificity rates were obtained in the proposed study that is classification of heart sounds which was segmented by the resampled energy method with CNN. Also, considering the other studies made with the same data set, it seems that classification of segmented S1-S2 sounds PCG images is more successful. From these results, it can be said that classification of segmented heart sounds with CNN has a much more efficient classification performance.

5. Conclusion

The results of classification studies of heart sounds wich were segmented into S1-S2 parts by resampling energy method are examined. The segmented phonocardiogram images of heart sound data in two different groups such as normal and abnormal in the PhysioNet atraining data set were obtained. Later, to speed up the learning process, the sizes of these images are reduced. These images are classified by a CNN. It was compared with the different studies conducted with the results of the other studies using the same data set to evaluate the obtained results. The proposed classification method has the highest accuracy, overall score, the sensitivity, and the specificity ratios according to other classification methods for the PhysioNet heart sounds data sets. From these results, it can be concluded that phonocardiogram images of segmented S1-S2 by resampling energy method can be classified effectively and efficiently with CNN.

6. Acknowledgment

This paper has been supported by Afyon Kocatepe University Scientific Research and Projects Unit with the Project number 18.KARİYER.44.

References

Clifford, G. D., Liu, C., Moody, B., Springer, D., Silva, I., Li, Q., & Mark, R. G. (2016, September). Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016. In *2016 Computing in Cardiology Conference (CinC)* (pp. 609-612). IEEE.

Convolutional Neural Networks, Last access date: 10.01.2018, http://ufldl.stanford.edu/tutorial/supervised/ ConvolutionalNeuralNetwork/.

Convolutional Neural Networks (LeNet), Last access date: 10.01.2018 http://deeplearning.net/tutorial/lenet.html.

Deperlioglu, O. (2018), *Intelligent Techniques Inspired by Nature and Used in Biomedical Engineering* ", Chapter 3 in *Nature-Inspired Intelligent Techniques for Solving Biomedical Engineering Problems*, ISBN13: 9781522547693.

Deperlioglu, O. (2018b). Classification of phonocardiograms with convolutional neural networks. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *9*(2), 22-33.

Deperlioglu, O. *Classification of heart sounds with segmented S1 and S2 sounds*, in proceedings of "*7th International Conference on Advanced Technologies (ICAT'18)*", Antalya, Turkey on 28 April - 1 May, 2018c.

Deperlioglu, O. (2018d). Segmentation of heart sounds by re-sampled signal energy method. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *9*(1), 17-28.

Gharehbaghi, A., Borga, M., Sjöberg, B. J., & Ask, P. (2015). A novel method for discrimination between innocent and pathological heart murmurs. *Medical engineering & physics*, *37*(7), 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 [Accessed 20 February 2017].

Langley, P., & Murray, A. (2016, September). Abnormal heart sounds detected from short duration unsegmented phonocardiograms by wavelet entropy. In *2016 Computing in Cardiology Conference (CinC)* (pp. 545-548). IEEE.

Liu, C., Springer, D., Li, Q., Moody, B., Juan, R. A., Chorro, F. J., ... & Syed, Z. (2016). An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, *37*(12), 2181.

Nilanon, T., Yao, J., Hao, J., Purushotham, S., & Liu, Y. (2016, September). Normal/abnormal heart sound recordings classification using convolutional neural network. In *2016 Computing in Cardiology Conference (CinC)* (pp. 585-588). IEEE.

Potes, C., Parvaneh, S., Rahman, A., & Conroy, B. (2016, September). Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. In *2016 Computing in Cardiology Conference (CinC)* (pp. 621-624). IEEE.

Puri, C., Ukil, A., Bandyopadhyay, S., Singh, R., Pal, A., Mukherjee, A., & Mukherjee, D. (2016, September). Classification of normal and abnormal heart sound recordings through robust feature selection. In *2016 Computing in Cardiology Conference (CinC)* (pp. 1125-1128). IEEE.

Rubin, J., Abreu, R., Ganguli, A., Nelaturi, S., Matei, I., & Sricharan, K. (2016, September). Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients. In *2016 Computing in Cardiology Conference (CinC)* (pp. 813-816). IEEE.

Ryu, H., Park, J., & Shin, H. (2016, September). Classification of heart sound recordings using convolution neural network. In *2016 Computing in Cardiology Conference (CinC)* (pp. 1153-1156). IEEE.

Sainath, T. N., Mohamed, A. R., Kingsbury, B., & Ramabhadran, B. (2013, May). Deep convolutional neural networks for LVCSR. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 8614-8618). IEEE.

Springer, D. B., Tarassenko, L., & Clifford, G. D. (2016). Logistic regression-HSMM-based heart sound segmentation. *IEEE Transactions on Biomedical Engineering*, *63*(4), 822-832.

Thomae, C., & Dominik, A. (2016, September). Using deep gated RNN with a convolutional front end for end-to-end classification of heart sound. In *2016 Computing in Cardiology Conference (CinC)* (pp. 625-628). IEEE.

Tschannen, M., Kramer, T., Marti, G., Heinzmann, M., & Wiatowski, T. (2016, September). Heart sound classification using deep structured features. In *2016 Computing in Cardiology Conference (CinC)* (pp. 565-568). IEEE.

Vafaie, M. H., Ataei, M., & Koofigar, H. R. (2014). Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. *Biomedical Signal Processing and Control*, *14*, 291-296.

Vernekar, S., Nair, S., Vijaysenan, D., & Ranjan, R. (2016, September). A novel approach for classification of normal/abnormal phonocardiogram recordings using temporal signal analysis and machine learning. In *2016 Computing in Cardiology Conference (CinC)* (pp. 1141-1144). IEEE.

Yang, X., Yang, F., Gobeawan, L., Yeo, S. Y., Leng, S., Zhong, L., & Su, Y. (2016, September). A multi-modal classifier for heart sound recordings. In *2016 Computing in Cardiology Conference (CinC)* (pp. 1165-1168). IEEE.

Yazdani, S., Schlatter, S., Atyabi, S. A., Vesin, J.-M. (2016), *Identification of Abnormal Heart Sounds,* In *2016 Computing in Cardiology Conference (CinC)* (pp. 1165-1168). IEEE.

Zabihi, M., Rad, A. B., Kiranyaz, S., Gabbouj, M., & Katsaggelos, A. K. (2016, September). Heart sound anomaly and quality detection using ensemble of neural networks without segmentation. In *2016 Computing in Cardiology Conference (CinC)* (pp. 613-616). IEEE.

Zhang, W., Han, J., & Deng, S. (2017). Heart sound classification based on scaled spectrogram and tensor decomposition. *Expert Systems with Applications*, *84*, 220-231.

Zheng, Y., Guo, X., & Ding, X. (2015). A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification. *Expert Systems with Applications*, 42(5), 2710-2721.



Omer Deperlioglu received his BSc in Electric and Electronic (1988) Gazi University, MSc in Computer Science (1996) from Afyon Kocatepe University, PhD in Computer Science (2001) from Gazi University in Turkey. Now he is associate professor of Computer Programming in Department of Science, Vocational School of Afyon, Afyon Kocatepe University of Afyonkarahisar, Turkey. His current research interests include different aspects of Artificial Intelligence applied in Power Electronics, Biomedical, and Signal Processing. He has edited 1 book and (co-) authored 3 books and more than 20 papers, more than 30 conferences participation, member in International Technical Committee of 4 conferences and workshops.