Research & Reviews: Journal of
Embedded System & Applications
https://journals.stmjournals.com/rrjoesa

ISSN: 2395-6712 (Online)
ISSN: 2321-8533 (Print)
Volume 12, Issue 1, 2024
January—April
DOI (Journal): 10.37591/RRJoESA

Review RRJoESA

Classifying Abnormalities in Heartbeat Sound

Befina Y.[1,*], A. Chandrasekar[2]

Abstract

Heartbeat sounds play a major role in the detection of various diseases such as heart disease, hyperthyroidism, and high blood pressure in their early stages. In the proposed method, various abnormal and healthy heartbeat audio signals are given as input and the features are extracted using MFCC (mel-frequency cepstral coefficients). Then, a deep learning approach is applied in which the MFCC audio signals are sent to the CNN (convolutional neural network) + LSTM (long short-term memory) model to extract and classify the abnormal heartbeat sounds with its associated disease. Thus, a system for classifying abnormal heartbeat sounds is proposed by leading a way to identify and diagnose diseases that are identified using heartbeat sounds in an early stage.

Keywords: Heartbeat audio signals, deep learning, mel-frequency cepstral coefficients, convolutional neural network, long short-term memory

INTRODUCTION

During the cardiac cycle, the cardiac valves open and close producing heartbeat sounds. Heartbeat sounds play a major role in the detection of various diseases that are identified using heartbeat sounds in an early stage. Evaluation of the heart sounds produced by the beating heart and the blood flow that results from it is a useful diagnostic technique for various diseases such as heart disease, hyperthyroidism, and high blood pressure. There are a few types of heartbeat sounds such as normal, murmur, extrasystole, and extra heart sounds. Heart sounds that are normal and healthy fall into the normal group. An abnormal heart murmur indicates complications such as heart failure. An extra heart sound in some situations is an important sign of high blood pressure. Extrasystole can happen normally in an adult and can be very common in children. In some situations, extrasystoles can be dangerous as they can indicate hyperthyroidism. With reference to the papers for the classification of heartbeat sounds it has been found that deep learning techniques performed better than machine learning techniques. The existing papers performed the classification using single neural networks providing maximum and minimum accuracies based on the neural network. This motivated us to consider two neural networks combined to produce maximum accuracy for classifying the abnormalities in the heartbeat sound with its associated disease. The neural network used in the proposed method is convolutional neural network (CNN) + long short-term memory (LSTM). The pre-processing step in the proposed method is carried out with MFCC (mel-frequency cepstral coefficient) in which the power spectrum of the input audio signal is identified. Data augmentation of the extracted features is further done by adding noise, shifting, stretching time and pitch to it. The pre-processed data is then implemented in the neural network. CNN's primary benefit is its ability to learn from unprocessed data, eliminating the need for manual feature engineering or pre-processing. Additionally, lowering the dimensionality and complexity of the input data improves the speed and efficiency of inference and training. By finding patterns and retaining pertinent information, LSTMs effectively increase performance. As a result, the

***Author for Correspondence**
Befina Y.
E-mail: befina20@gmail.com

[1]Student, Computer Science and Engineering, St.Joseph's College of Engineering, Chennai, Tamil Nadu, India
[2]Professor, Computer Science and Engineering, St.Joseph's College of Engineering, Chennai, Tamil Nadu, India

Received Date: March 14, 2024
Accepted Date: March 30, 2024
Published Date: April 04, 2024

Citation: Befina Y., A. Chandrasekar. Classifying Abnormalities in Heartbeat Sound. Research & Reviews: Journal of Embedded System & Applications. 2024; 12(1): 24–31p.

CNN may concentrate on extracting spatial information from the LSTM network's output for use in classification and other applications. In section I, it contains the overview of the paper. The next section presents the existing methodology. The third section contains the methodology of the paper. The fourth section contains the results and discussion of the paper. The penultimate section presents the conclusion and the final section presents the future development of the paper.

EXISTING METHODOLOGY

The existing methods of the classification of abnormal heartbeat sounds paved a way for finding a solution for improving the model performance in the proposed methodology. Ul Haq et al. [1] aimed to classify various heart rate signals by optimizing the deep learning model with EfficientNet architecture. The mel spectrogram of the heart rate signals were given as input to the model. This resulted in minimum accuracy and the paper suggested to use different neural networks for better accuracy. The system was lacking in robustness as it failed working in noisy conditions. Jadhav et al. [2] described a novel approach towards classifying heartbeat sounds. An algorithm using Shannon energy is used for segmentation and neural networks were used for classification of heartbeat sounds. First, an algorithm based on Shannon energy envelope calculation and feature construction for systoles and diastoles segments of a heart sound is implemented. The proposed method however had a drawback of segmenting between diastoles and systoles and produced a lesser accuracy for classifying heartbeat sounds. Ryu et al. [3] worked with a cardiac diagnostic model using CNN, which is useful in predicting whether a heartbeat sound is healthy or abnormal. The heartbeat recordings are filtered using Windowed-sinc Hamming filter and classified using CNN. The method, however, presented some disadvantages such as classification produced a minimum accuracy. It lacks the usage of other feature extraction techniques other than CNN helpful for classification. Malik et al. [4] researched based on the multi-classification of the heartbeat audio signals. The proposed method has used recurrent neural network (RNN) model using LSTM and different filtering techniques such as bandpass filter has been used to remove noises. The method led to a disadvantage of applying many filtering techniques which eventually result in degrading the performance of the system while predicting the heartbeat sounds under noisy conditions. Gomes et al. [5] have used a method for heartbeat sound classification. The result includes two steps such as segmentation and classification of heartbeat sounds. Shannon algorithm is used for segmentation and classification is done using multilayer perceptron (MLP) algorithm. The method provided better precision but was comparatively less than the University College London team mentioned in the paper.

METHODOLOGY

System Overview

The system architecture of the proposed method as shown in Figure 1 explains that the input audio signal data is taken as wav format from Kaggle sponsored by Pascal, the dataset contains both normal and abnormal heartbeat sounds. Features of the input audio signals are extracted using MFCC [6]. The extracted MFCCs are pre-processed by adding noise, shifting, applying time stretch and pitch shifting and then they are passed to two deep learning neural network for extraction and classifying the heartbeat sound with its associated diseases. The deep learning neural networks used here are CNN + LSTM where CNN plays a role of extraction and LSTM plays the role of classification. The model is then evaluated and passed as the final output. The final output from the model is the classification system of abnormal heartbeat sounds [7].

Data Understanding

The data is collected from Kaggle in which a challenge was sponsored by Pascal, where the heartbeat sounds are collected from general public and hospitals using iStethoscope and digital stethoscope, respectively. The aim is to classify the heartbeat sound abnormalities [8]. The audio files are represented in wav format (waveform format). As shown in Figure 2, the dataset consists of 40 artifacts (sounds other than heartbeat), 351 normal, 129 murmur, 46 extrasystole and 19 extra heartbeat sounds. The heartbeat sounds in the first heart sound S1 and second heart sound S2 pattern are monitored to check the heartbeat signal. If the S1 and S2 are in regular pattern, then they are called normal heartbeat sounds. If S1 and S2 are together repeated, then there is an abnormality in the heartbeat sound resulting in murmur heartbeat sound. If S1 and S2 occur in a separated manner and if they are repeated, then they are extrasystole and extra heartbeat sounds [9–12].

Feature Extraction – MFCC

Feature extraction from audio data is done by computing MFCCs to create 2D image-like patches as shown in Figure 3.

The audio signal given to the MFCC will first break the audio signal waves into an overlapping frame. The audio signals are then passed to fast Fourier transform where the window length is given for segmentation of the frames to extract the features. MFCC captures the shape of the power spectrum of sound signal [13–15]. They are obtained by first utilizing a method such as the discrete Fourier transform (DFT) to convert the raw audio signal into a frequency domain, and then using the mel-scale a filter bank to simulate how the human ear perceives sound frequency. Therefore 52 MFCC features from each audio data were created using 512 hop length and segmentation of frames using windows and passed to DFT thus filtering of the audio data takes place.

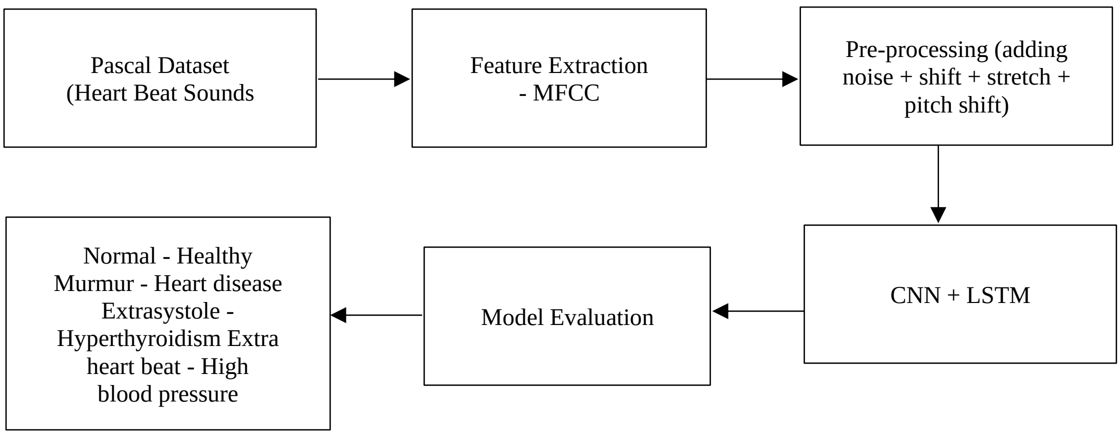

Figure 1. System architecture.

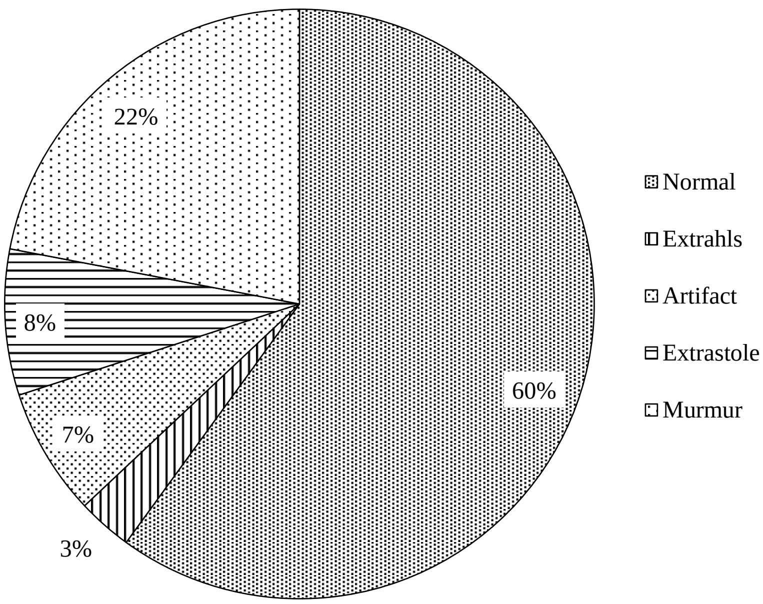

Figure 2. Exploration of the dataset.

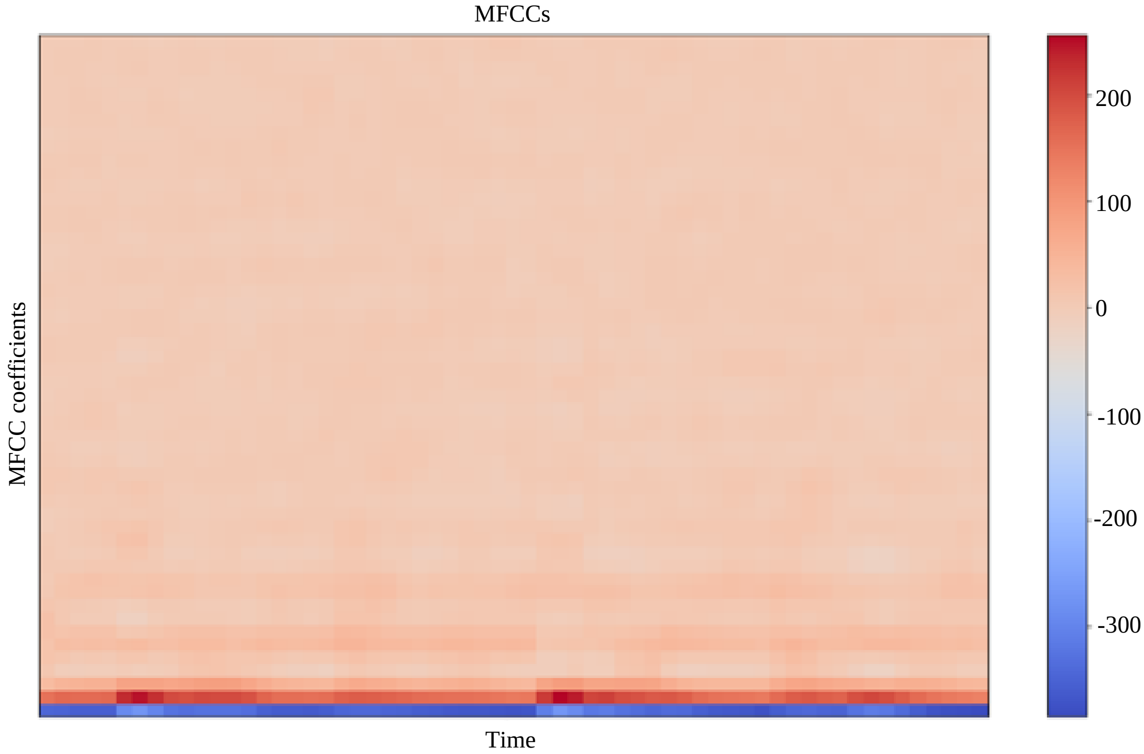

Figure 3. Mel frequency cepstral coefficient (MFCC) representation.

Pre-Processing

Pre-processing or data augmentation step involves a set of techniques that are applied to raw data that are coherent to be given as input to deep learning neural networks. To make the system robust by exposing the model with different range of inputs, we can inject noise to the audio data, shift the time of the audio, change the pitch and speed of the audio. NumPy is helpful in injection of the noise and shifting time whereas librosa is useful in processing audio data. The first step for pre-processing used here is to add noise to the 52 features extracted audio data from MFCC and see how the models perform under these noisy conditions. The next step is to apply shift to both the extracted audio data and added noise that is applied, time stretch is applied to the audio data and finally, pitch shift is applied to it.

CNN + LSTM

A CNN+LSTM network is used for extraction and classification of audio data. CNN works with sequential data by extracting features of the input. It can learn features from both spatial and time dimensions. An LSTM network works with sequential data by looping leading to coherent classification of the audio signals.

CNN + LSTM network as shown in Figure 4 uses three convolutional and three LSTM layers to perform extraction and classification. To train a CNN + LSTM network with audio data, we can extract audio-based spectrograms such as MFCC to be given as input to the neural network. Sequence of input layer and input size are given to the neural network. Activation functions such as ReLU and SoftMax layer are included. CNN includes convolutional 1D layer with input shape (52,1), Max pooling layer is given with pool size 2 and a batch normalization is added. Totally, three CNN layers were added. Three layers of LSTM which includes two ReLU activation functions and one SoftMax activation function. All the LSTM layers included dense and dropout layers.

RESULTS AND DISCUSSION

In this study, as shown in Table 1, the classification of abnormal heartbeat sound is examined using CNN + LSTM. It was also compared with CNN + gated recurrent unit (GRU) in order to identify the best suitable deep learning model for the classification. After examining the accuracies, it has been found that CNN + LSTM is the most suitable deep learning neural network.

The accuracy graph of the CNN + LSTM model as shown in Figure 5 consists of both training and validation accuracies that have been performed with 80 epochs and the loss graph of the model as shown in Figure 6 has been obtained while performing with 80 epochs. The model evaluation is implemented by the confusion matrix as shown in Figure 7, which is helpful in model evaluation. Confusion matrices are generated by a binocular network, where the vertical coordinate corresponds to ground truth labels and the horizontal coordinate corresponds to prediction results labels. The model is finally integrated to a web application as given in Figure 8 in which classification of the abnormal heartbeat sound takes place.

CNN (CONV1D, Max Pooling, Batch Normalization)
CNN (CONV1D, Max Pooling, Batch Normalization)
CNN (CONV1D, Max Pooling, Batch Normalization)
LSTM (ReLU Activation function)
LSTM (ReLU Activation function)
LSTM (SOFTMAX Activation function

Figure 4. Convolutional neural network (CNN) + long short-term memory (LSTM) architecture.

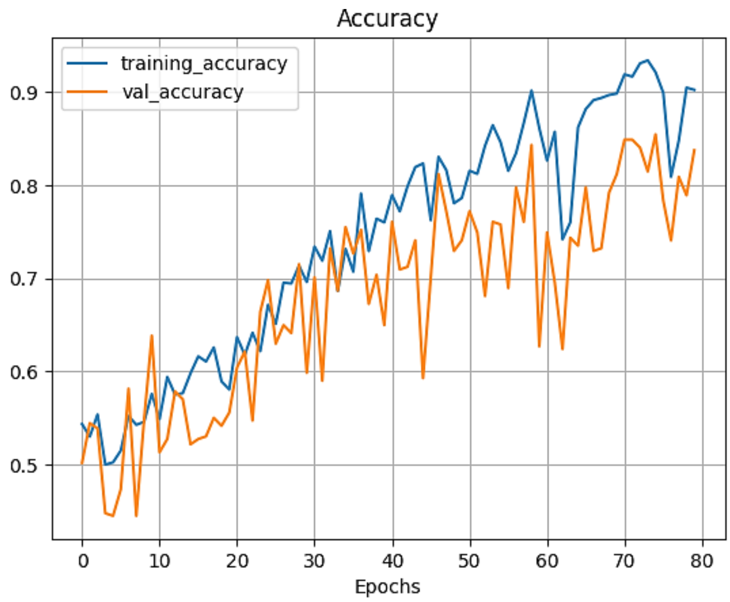

Figure 5. Accuracy graph.

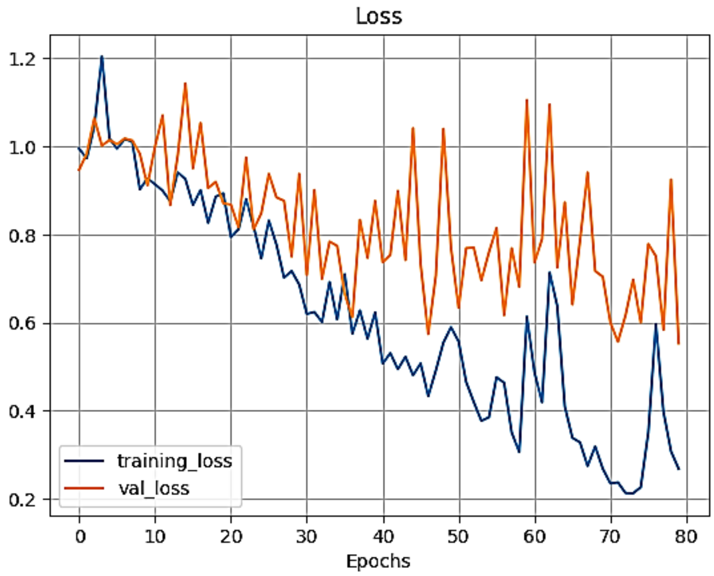

Figure 6. Loss graph.

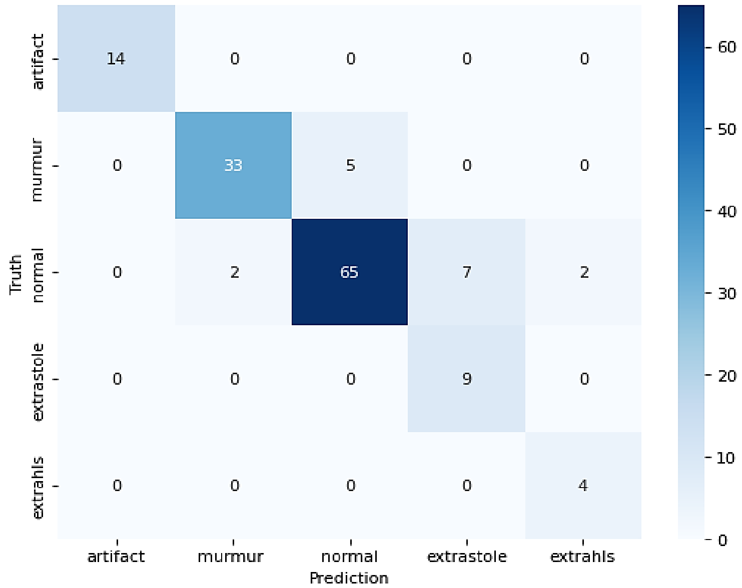

Figure 7. Confusion matrix.

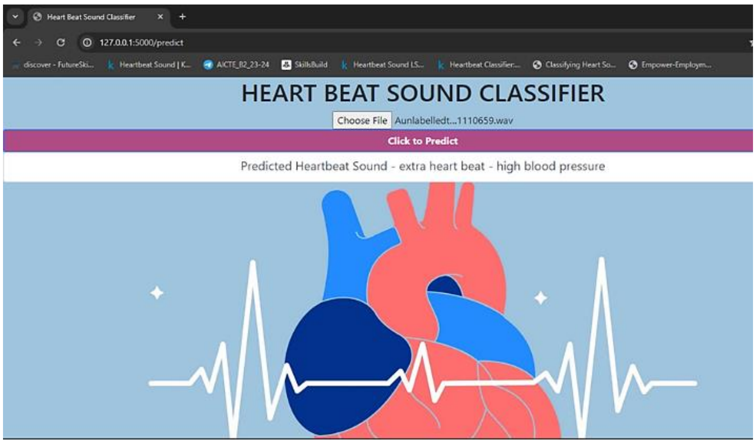

Figure 8. Classification web application.

Table 1. Accuracy results.

Neural Network	Accuracy
CNN + LSTM	83%
CNN + GRU	71%

CONCLUSION

In this study, coherent feature extraction from the given dataset has been done using MFCC to extract discriminant feature and pre-processing of augmenting the data to work even in noisy conditions has been done. Integrating of two neural networks such as CNN and LSTM for extraction and classification of the heartbeat sound, respectively, has been done. The extracted feature of the audio by MFCC is taken as the input to the model which is further trained, validated, and tested. The model is finally deployed in web for making an end-to-end deep learning model where the web application will be helpful in predicting and classifying the abnormalities in heartbeat sound. The study can be further improved by training the model using different heartbeat sound datasets that contains heartbeat sounds in different locations.

Future Work

Further, the study can be carried out by using different combinations of neural networks to improve the accuracy of the model. The study can be enhanced by using many new datasets and pre-processing techniques.

REFERENCES

1. Ul Haq HFD, Ismail R, Ismail S, Purnama SR, Warsito B, Setiawan JD, Wibowo A. EfficientNet optimization on heartbeats sound classification. In: 2021 5th International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, November 24–25, 2021. pp. 216–221.
2. Jadhav AR, Ghontale AG, Ganesh A. Heart sounds segmentation and classification using adaptive learning neural networks. 2017 International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, July 28–29, 2017. pp. 33–38. doi: 10.1109/CSPC.2017.8305881.
3. Ryu H, Park J, Shin H. Classification of heart sound recordings using convolution neural network. In: 2016 Computing in Cardiology Conference (CinC), Vancouver, British Columbia, Canada, September 11–14, 2016. pp. 1153–1156.
4. Malik H, Bashir U, Ahmad A. Multi-classification neural network model for detection of abnormal heartbeat audio signals. Biomed Eng Adv. 2022; 4: 100048. doi: 10.1016/j.bea.2022.100048.
5. Gomes EF, Bentley PJ, Pereira E, Coimbra MT, Deng Y. Classifying heart sounds – approaches to the PASCAL challenge. In: Proceedings of the International Conference on Health Informatics (HEALTHINF), Barcelona, Spain, February 11–14, 2013. pp. 337–340.
6. Raza A, Mehmood A, Ullah S, Ahmad M, Choi GS, On BW. Heartbeat sound signal classification using deep learning. Sensors. 2019; 19 (21): 4819. doi: 10.3390/s19214819.
7. Mukherjee U, Pancholi S. A visual domain transfer learning approach for heartbeat sound classification. Arxiv 2021. arxiv: abs/2107.13237v1. Available at https://arxiv.org/pdf/2107.13237
8. Kumar D, Carvalho P, Antunes M, Gil P, Henriques J, Eugenio L. A new algorithm for detection of S1 and S2 heart sounds. In: 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, Toulouse, France, May 14–19, 2006. pp. II–II. doi: 10.1109/ICASSP. 2006.1660559.
9. Deng SW, Han JQ. Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. Future Generation Computer Syst. 2016; 60: 13–21. doi: 10.1016/j. future.2016.01.010.
10. Zhang W, Han J, Deng S. Heart sound classification based on scaled spectrogram and tensor decomposition. Expert Syst Appl. 2017; 84: 220–231. doi: 10.1016/j.eswa.2017.05.014.
11. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Neural Networks. 2005; 18 (5–6): 602–610. doi: 10.1016/j.neunet. 2005.06.042.

12. Groch MW, Domnanovich JR, Erwin WD. A new heart-sounds gating device for medical imaging. IEEE Trans Biomed Eng. 1992; 39 (3): 307–310. doi: 10.1109/10.125016.
13. Babaei S, Geranmayeh A. Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals. Computers Biol Med. 2009; 39 (1): 8–15. doi: 10.1016/j.compbiomed.2008.10.004.
14. Debbal SM, Bereksi-Reguig F. Automatic measure of the split in the second cardiac sound by using the wavelet transform technique. Computers Biol Med. 2007; 37 (3): 269–276. doi: 10.1016/j.compbiomed.2006.01.005.
15. Randhawa SK, Singh M. Classification of heart sound signals using multi-modal features. Procedia Computer Sci. 2015; 58: 165–171. doi: 10.1016/j.procs.2015.08.045.

