International Journal of Informatics and Applied Mathematics
e-ISSN:2667-6990 Vol. 5, No. 1, 62-73

PCG Classification Using Scalogram And CNN With DAG Architecture

Mohammed Saddek Mekahlia, Mohammed Fezari and Ahcen Aliouat

Laboratory of Automatic and Signals of Annaba (LASA), Electronics department, Badji Mokhtar University, BP 12, Sidi Amar, Annaba, 23000, Algeria

Abstract. Cardiovascular diseases (CVDs) are the most leading causes of death every year in the world. The threat of CVDs can be decreased and controlled with early diagnoses. Therefore, interpreting heart sounds is considered as one of the common ways to diagnose the cardiovascular system. Heart sound signal as known as phonocardiogram (PCG) provides useful information about the heart condition, which can be used in the diagnostic, and helps the physicians in the detection of several cardiovascular abnormalities. The technology development helped in the appearance of new diagnosis techniques, which combines new advanced signal processing techniques and deep learning algorithms. Thus, the heart sound classification is becoming a crucial task in the modern healthcare field. In this work a deep learning-based classification method was proposed. Using PCG database which contains five different classes taken from different cases of heart valve defects. Scalogram of heart sound signals was used as time-frequency representation to create a scalogram image database extracted from the PCG database. A convolutional neural network with Direct Acyclic Graph structure (DAG CNN) was used in the classification of the scalogram image database. The evaluation of the classification performance indicated that the accuracy was about 99,6%. A comparative results manifest that the proposed method had a better performance compared to other previous works in which the same database was used.

Keywords: Scalogram · CNN · DAG · Feature Extraction · Heart Sound Classification · PCG · Heart Valve Diseases.

1 Introduction

Cardiovascular diseases take about 17,9 million lives every year which became the most leading cause of death globally [1]. Heart valve diseases are diseases of the cardiac valves, mainly of the mitral valve or the aortic valve. It may occur as a consequence of congenital deformities or even physiologic processes including pregnancy and rheumatic heart disease, but usually it results because of aging [2]. The structure of heart is responsible for delivering the circulation of blood through all parts of the body. The hearts blood-pumping cycle, called cardiac cycle, ensures that blood is distributed throughout the body. Thus, a regularly abnormal heartbeat or beats per minute are characteristic of a heart-related illness [3, 4]. This is because a heartbeat is made of two phases: systole and diastole. The systole-diastole relationship is the reference in measuring blood pressure. Other ways of physically determining the regular functioning of the heart are through examining the pulse rate (beats per minute) [5–7]. Many signal activities between the first and the second heart sound can be seen. These extraneous activities occurring between S1 and S2 are referred as two abnormal sound signals S3 and S4. Therefore, murmurs occurring in the heart sound are signs of a pathological case [8].

Heart sounds are low-frequency signals and it is hard for human ears to distinguish it, especially murmurs. Therefore, due to limited ears capacity, different techniques using different items were used in cardiology to hear different heart sounds. Heart auscultation has been a very important method for the early diagnosis of heart diseases [9]. Auscultation using the stethoscope as many other techniques, has seen many technologic developments. Nowadays electronic stethoscope proved its efficiency, besides its reliability compared to the traditional stethoscope [10]. Many reports show that medical students and inexperienced doctors have made about three times more errors in analyzing heart sounds from specialist doctors [11].

Electronic stethoscope allows to record and visualize heart sounds as Phonocardiogram (PCG). Figure 1 shows PCG signal of healthy subject. Because the heart sound carries different information about the heart stat, PCG can be processed and used in diagnostic. Machine learning-based normal and abnormal heart sound classification using audio processing may be a possible method for automatic diagnosis of cardiac diseases or abnormalities without the help of trained medical professionals. Also, it may be a great help in primary health centers for early diagnosis and screening of cardiac disorders [12].

In this work scalogram representation of PCG is used. Image database of scalogram was created and extracted from an audio PCG database which contains five classes of five different types of heart sounds signals, one class is normal and the four other ones are considered to be abnormal. In the classification process, Convolutional neural network with Direct Acyclic Graph architecture (DAG CNN) was used for multi-class classification. DAG CNN was trained and tested using the scalogram images database without segmenting the data. The performance of the model used in classification was calculated, and an accuracy of 99,6% was achieved. After that the results obtained were compared to works in which the same PCG database was used. This paper is organized as follow: it begins with related works in heart sound classification. Then, the proposed method is described with database used in this work, scalogram and convolutional neural network. After that results and discussion. Finally, this paper ends with a concise conclusion part that sums up the main idea of this work.

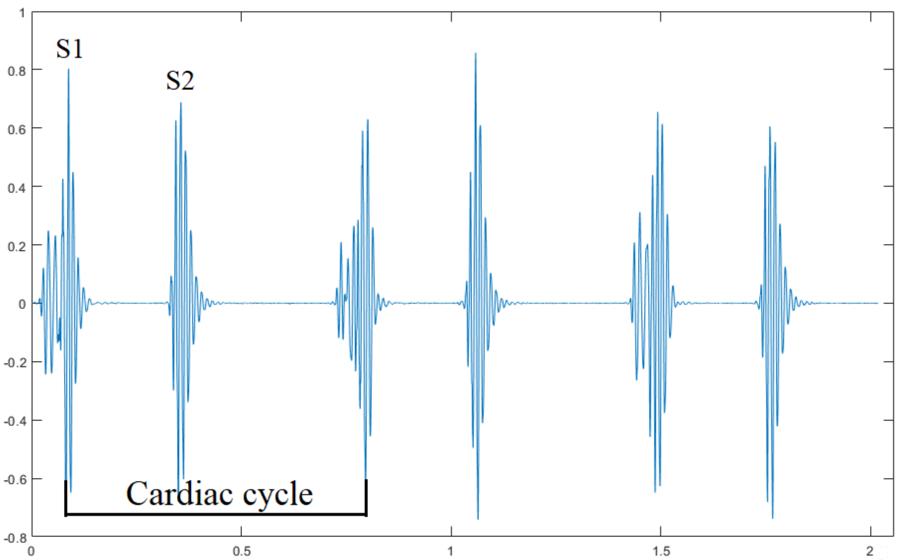


Fig. 1: PCG signal of healthy subject.

2 Related works

In the literature, several techniques have been used for classifying the normal and abnormal heart sounds. They are categorized based on different time-frequency that transforms disparate set of features and various classification methods. Some of these works are as follow: in [12] heart sounds were classified using a pretrained CNN model on 2 datasets of normal and abnormal heart sounds, with different sampling frequency 44,1 kHz for dataset A and 4 kHz for dataset 2, the spectrogram was used as time-frequency representation in order to get a precision of 80% and 79% respectively. In [13] 1D CNN model was used to classify 13015 PCGs of normal and abnormal cases from PhysioNet/ CinC challenge database using DAE network to extract features, to achieve an accuracy of 99,01%. In [14] 3240 PCGs from PhysioNet/ CinC challenge database were passed by ResHNet model to get 97,2% accuracy. Another method using CNN model was proposed in [15] with 327 heart sound from normal and abnormal subjects, FFT features were used for classification to achieve an accuracy of 98%. 1000 PCG samples with 5 different classes (Normal, Aortic Stenosis, Mitral Stenosis, Aortic Regurgitation, Mitral Regurgitation, Mitral Valve Prolapse) were used in [16], time frequency representation of Bispectrum images were used as inputs of the multiclassification with CNN model for an accuracy of 98,7%. In [17] Feedforward Neural Network was applied on PCG database of 1081 samples with 85,74% accuracy. In [18] SVM was used to classify MFCC+DWT features to classify 1000 PCGs of normal and abnormal subjects, and the accuracy was about 97,9%. In [19] a modified pre-trained Alexnet model was used to classify scalogram representation images, extracted from PhysioNet/ CinC challenge database containing 3240 PCG samples, to achieve an accuracy of 90%.

3 Proposed method

The method used in this work is based on the combination of CNN with DAG architecture and spectral images of scalogram representation. Figure 2 shows the block diagram of the proposed method. The database used is consisted of 1000 PCGs for five different valve heart cases that are divided into five classes. Each class contains 200 samples from each case. Thus, a multiclass classification with DAG CNN is used, evaluated, and compared to other works using the same database.

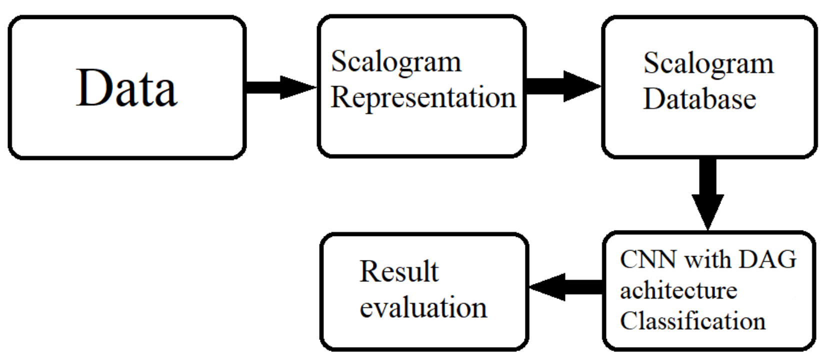


Fig. 2: Block diagram of the proposed method.

3.1 Database

PCG recordings used in this work were obtained from the database created by Yassen et al [18]. The data was collected from random sources (Auscultation skills CD, Heart sound made easy, 48 different websites provided the data including Washington, Texas, 3M, and Michigan and so on). After collection, files with extreme noise were excluded, sampled to 8 kHz and converted to mono channel. The database contains 1000 PCGs and it was divided into 5 classes (Normal, AS, MS, AR, MR and MVP) with 200 samples for each class. In this work a scalogram images database was created from the audio database, keeping the same number of samples (1000) without segmenting the data, and it is saved as JPEG. After that, data was used as input in the classification process.

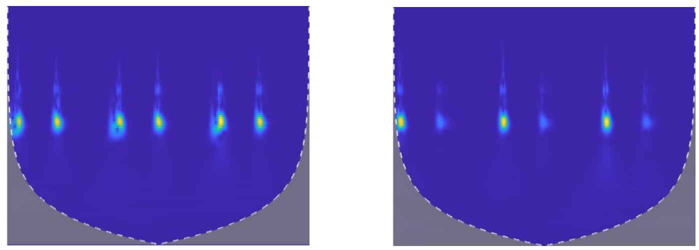
3.2 Scalogram

The scalogram is the time-frequency representation of the signal by Continuous Wavelet Transform (CWT). The coefficient values of time-frequency locations are indicated using the colour or brightness. Scalogram is compared to spectrograms, where only a fixed time and frequency resolution is given [20]. A scalogram is created using a short time Fourier transform (STFT). The CWT and by allowing different size analysis windows at different frequencies, it provides superior time and frequency resolution to the STFT. The frequencies present at different times in the signal are represented by the scalograms which gives a visual representation that can be used to distinguish between the heart sounds [21]. Figure 3 shows the scalogram representations extracted from each of the five classes in Yassen et al database [18], two scalogram images from each of the five different cases (Normal, AS, MS, MR and MVP). Viewed with human eyes, the difference between the images from different classes can be distinguished. Thus, the scalogram can be a good tool in heart diseases diagnosis.

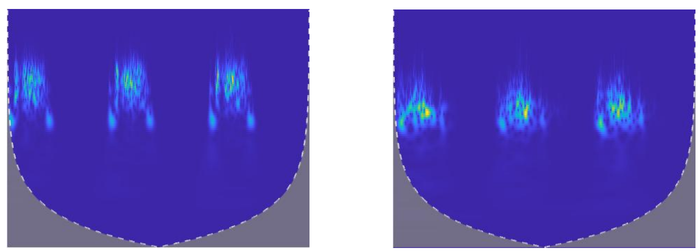
As CNN have proven its efficiency for visual recognition tasks, The scalogram images were proposed to be used in this work to improve the performance of the CNN, which can distinguish the different features corresponding to each class based to the clear presentation of coefficient values of the time-frequency resolution on scalogram. Thus, the classification process gives better results.

3.3 Convolutional Neural Network

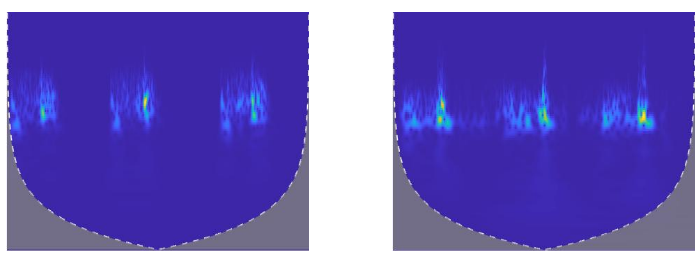
Deep learning algorithms have become commonly used in the medical field, that covers the disease recognition, diagnosis and classification. A part of this deep learning algorithms is convolutional neural network which is widely used. CNN consisting of several layers is an artificial neural network that handles 2D images as well as 1D signals. Layers that are usually used in CNN are represented by. input layer, convolution layer, RELU layer fully connected layer, classification layer, and an output layer [17]. ReLU layers are non-linear activation functions. Convolutional layers consist of spatial filters that are convolved across the width and height of the input, as well as pooling layers that downsample the input to reduce the number of parameters and also to reduce overfitting. Between convolutional and ReLU layers, the batch size normalization layers are positioned. These layers normalize the activation of each channel, reducing training time



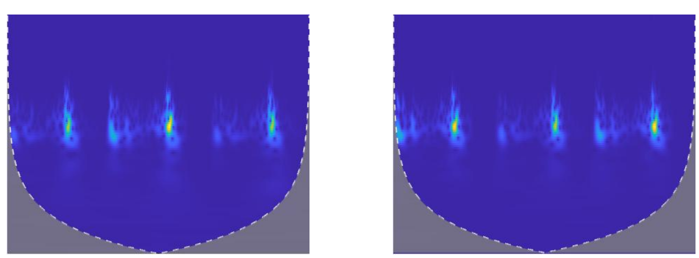
(a) Scalogram representation of normal heart sound



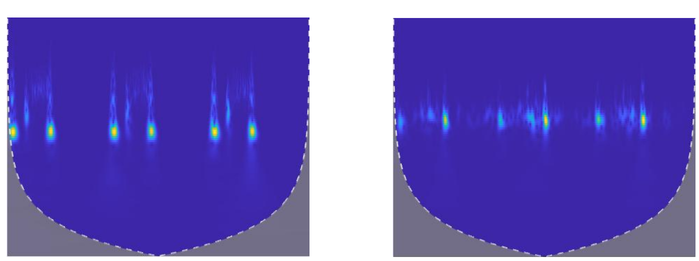
(b) Scalogram representation of Aortic Stenosis heart sound



(c) Scalogram representation of Mitral Regurgitation heart sound



(d) Scalogram representation of Mitral Stenosis heart sound



(e) Scalogram representation of Mitral Valve Prolapse heart sound

Fig. 3: (a) Scalogram representation of normal heart sound. (b) Scalogram representation of Aortic Stenosis heart sound. (c) Scalogram representation of Mitral Regurgitation heart sound. (d) Scalogram representation of Mitral Stenosis heart sound. (e) Scalogram representation of Mitral Valve Prolapse heart sound.

and sensitivity to network initialization. The final layer of a CNN used for classification purposes is usually a layer of fully connected neurons that computes the class scores [21]. CNN, unlike traditional feature extraction methods, does not use manual features selection which respond to the problem of the limited types of features. It also allows spectral and textural features to be explored and used. Using supervised learning and as the CNN doesnt require manual feature selection, the network must be trained to be allowed to extract the different types of features [22]. In the proposed method, CNN with DAG structure was used. DAG architecture helps to improve the processing time and allows the input image to maintain a larger size through convolution layers. Therefore, the number of features learned by the network will increase [23]. In consequence the performance of the CNN model will improve. The database was divided into 80% (800 samples, 160 samples from each class) for training process, and 20% for testing (200 samples, 40 samples from each class).

4 Results and discussion

The proposed method in this work based on combining CNN with DAG structure and scalogram representation images of heart sound signals were implemented using Matlab under Intel Core-i5 computer with 4Gb RAM. Using the 5 K-fold cross validation technique. The CNN model proposed was trained and tested using scalogram images database. The scalogram representation images was extracted from database containing 5 classes of PCG (Normal, AS, MS, MR, MVP), created by Yaseen el al [18]. 20% of samples from each class of the database was used for testing and the other 80% for training. The network was trained with 500 iterations, 10 epochs and 50 iterations per epoch where the minibatch size was set to 16.

4.1 Results

The training progress of accuracy and loss is shown in Figure 4 After the training process the network was tested with the rest 20% of the scalogram images database.

The performance of the classifier is measured with different parameters like accuracy, sensitivity specificity, precision and F1 Score, which can be calculated as follow [24]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Sensiticity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Precision = \frac{TP}{FP + TP} \tag{4}$$

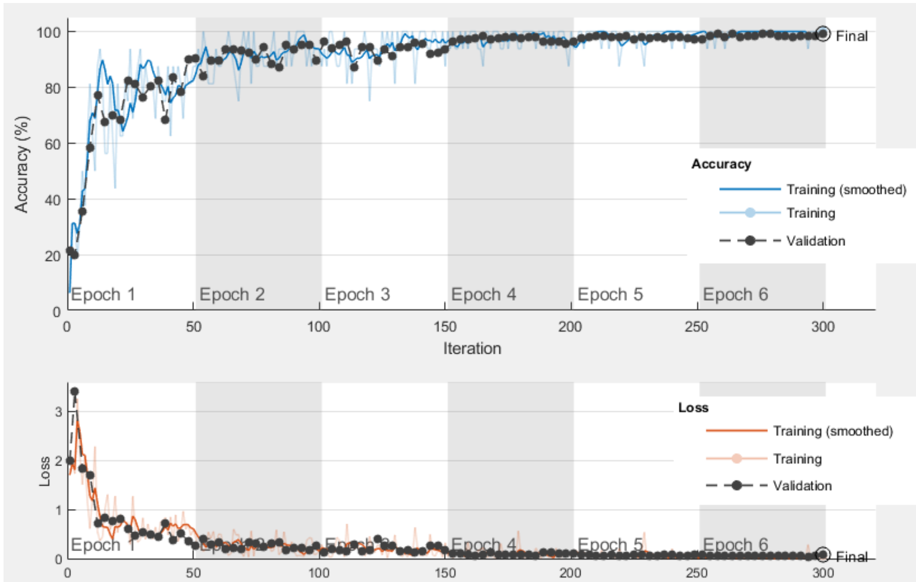


Fig. 4: the training progress of accuracy and loss.

$$F1Score = 2 * \frac{Precision * Sensitivity}{Precision + Sensiticity} \tag{5}$$

the results obtained were about 99,6% of accuracy. Table 1 shows the results of the proposed method. Also, a confusion matrix was generated. Figure 5 presents the confusion matrix containing the classification results of the proposed method.

Table 1: Results of the proposed method.

Accuracy	Sensitivity	Specificity	Precision	F1score
99,6%	99,6%	99,9%	99,59%	99,59%

4.2 Discussion

The proposed method was compared to other previous works in which the same database was used. In the first work MFCC+DWT were used for feature extraction and SVM as classifier [18]. Also, another method compared to this work was based on bispectrum representation and AOCNet CNN for classification [16]. the problem posed by yassen et al [18] is that data (1000 samples) wasnt enough for DNN to classify MFCC+DWT extracted features with an accuracy of 92,1%,

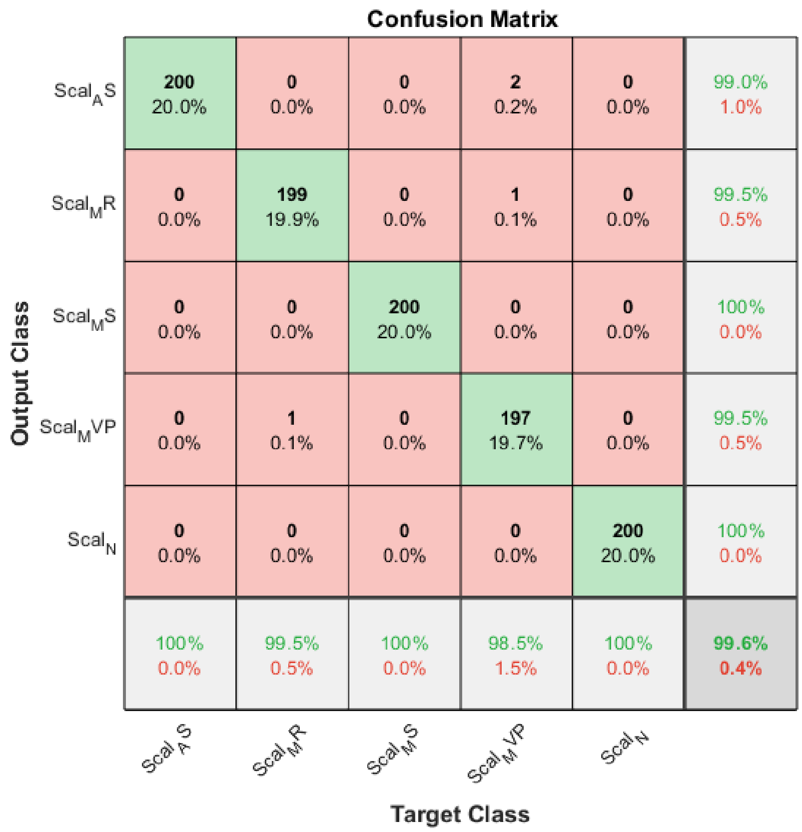


Fig. 5: Confusion matrix for testing results.

but while using SVM an accuracy of 97,9% was achieved. Thus, a larger dataset was required for deep learning. Alqudah et al [16] achieved 98,7% accuracy using bispectrum representation images extracted from the same database used in this work, and AOCT CNN. In this work, scalogram representation images were extracted from the audio database and were classified using a CNN with DAG architecture. With measuring the performance of the network on the database, an accuracy with 99,6% was obtained. Based on the previous results it can be observed that the scalogram representation can be a better tool for PCG interpretation, providing a lot of information about the heart sound signals. The multiple features carried by the heart sound signal can be more visible in time-frequency domain with the CWT. Thus, scalogram representation helps more the classifier in the features extraction process. In consequence the classification results are higher compared with other works. Table 2 presents a performance comparison between the proposed method in this work and previous works. CNN with DAG architecture gave better results than the two-precedent works without segmenting data.

Table 2: The performance comparison of the proposed method with other works.

Author	classifier	Accuracy	Sensitivity	Specificity	F1score
Proposed method	DAG CNN	99,6%	99,6%	99,6%	99,59%
Yaseen et al [18]	AOCT CNN	97,9%	98,2%	99,4%	99,7%
Alqudah et al [16]	SVM	98,7%	98,7%	99,67%	98,71%

5 Conclusion

Heart sound signals are important in the diagnosis of heart diseases. PCG contains several information about the heart stat. Thus, extracting this information using signal processing is crucial to better understand the nature of the sound. In this work a new method was proposed in heart sound classification, based on scalogram representation images and CNN with DAG structure. The scalogram images were extracted from an open source PCG database containing five classes. Each class contains PCG samples concerning one of the five heart diseases cases (Normal, AS, MS, MR, MVP). After that, the scalogram images were used to train as well as to test the CNN model. The performance of the DAG CNN gave a better result compared with other works using the same database reaching an accuracy of 99,6%. Form the results illustrated in this paper the scalogram can be a good tool in heart sound classification beside the CNN with DAG architecture which give a good performance with high accuracy.

References

1. Cardiovascular diseases (CVDs), https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Last accessed 22 jan 2022
2. R.O. Bonow et al.: ACC/AHA 2006 guidelines for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the Society of Cardiovascular Anesthesiologists: endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. J Am Coll Cardiol. **48**(3), pp. 1148, (2006)
3. How the Heart Works, https://www.nhlbi.nih.gov/health-topics/how-heart-works. Last accessed 22 jan 2022
4. Heart Beat, https://my.clevelandclinic.org/health/articles/17064-heart-beat. Last accessed 22 jan 2022
5. E. Delgado-Trejos, A.F. Quinceno-Manrique, J.I. Godino-Llorente, M. Blanco-Velasco, G. Castellanos-Dominguez.: Digital Auscultation Analysis for Heart Murmur Detection. Annals of Biomedical Engineering, **37**(2), pp. 337353, (February 2009)

6. Shindler, Daniel M. MD, FACC.: Practical Cardiac Auscultation. Critical Care Nursing Quarterly. **30**(2), pp. 166-180 (April 2007)
7. Pediatric Cardiology A Chapter in Core Concepts of Pediatrics, 2nd Edition, https://www.utmb.edu/pedi_ed/CoreV2/CardiologyPart1/CardiologyPart12.html. Last accessed 22 jan 2022
8. L. Bahekar, A. Misal, G.R. Sinha.: Heart Sound Segmentation Techniques: A Survey. IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE). **2**(11), pp 46-49, (2014)
9. S. Leng , R.S. Tan, K.T Chuan Chai, C. Wang, D. Ghista, L. Zhongl.: The electronic stethoscope. BioMed Eng OnLine. **14**(66), (2015)
10. J.P. Tourtier, N. Libert, P. Clapson, K. Tazarourte, M. Borne, L. Grasser, B. Debien, Y. Auroy.: Auscultation in flight: comparison of conventional and electronic stethoscopes. Air Med J. **3**(2), pp. 158-160 (2011)
11. A. Yadav, A. Singh, M.K Dutta, C.M. Travieso.: Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. Neural Computing and Applications 32. (2020)
12. F. Demir, A. engr, V. Bajaj, K. Polat.: Towards the classification of heart sounds based on convolutional deep neural network. Health Information Science and Systems **7**. (2019)
13. F. Li, M. Liu, Y. Zhao, L. Kong, L. Dong, X. Liu, M. Hui.: Feature extraction and classification of heart sound using 1D convolutional neural networks. EURASIP Journal on Advances in Signal Processing. **2019**(59), (2019)
14. A. Balamurugan, S.G. Teo, J. Yang, Z. Peng, Y. Xulei, Z. Zeng.: ResHNet: Spectrograms Based Efficient Heart Sounds Classification Using Stacked Residual Networks. In: 2019 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). pp. 1-4. IEEE, Chicago, IL, USA (2019). https://doi.org/10.1109/BHI.2019.8834578
15. T. Li, C. Qing, X. Tian.: Classification of Heart Sounds Based on Convolutional Neural Network. In: International Conference on Internet Multimedia Computing and Service. **819**, pp 252-259. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-8530-7_24
16. A.M. Alqudah, H. Alquran, I. Abu Qasmieh.: Classification of heart sound short records using bispectrum analysis approach images and deep learning. Network Modeling Analysis in Health Informatics and Bioinformatics **9**. (2020)
17. P.T. Krishnan, P. Balasubramanian, S. Umapathy.: Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network. Physical and Engineering Sciences in Medicine **43**. pp. 505515 (2020)
18. Yaseen, G.Y Son, S. Kwon.: Classification of Heart Sound Signal Using Multiple Features. Appl. Sci. **8**(12), (2018)
19. S.A. Singh, S. Majumder. M. Mishra.: Classification of short unsegmented heart sound based on deep learning. In: 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). pp. 1-6. IEEE, Auckland, New Zealand (2019). https://doi.org/10.1109/I2MTC.2019.8826991
20. Z. Ren, K. Qian, Z. Zhang, V. Pandit, A. Baird and B. Schuller.: Deep Scalogram Representations for Acoustic Scene Classification. IEEE/CAA Journal of Automatica Sinica, **5**(3), pp. 662-669 (May 2018)
21. A. Meintjes, A. Lowe and M. Legget.: Fundamental Heart Sound Classification using the Continuous Wavelet Transform and Convolutional Neural Networks. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 409-412. IEEE, Honolulu, HI, USA (2018). https://doi.org/10.1109/EMBC.2018.8512284

22. Z.C. Horn, L. Auret, J.T. McCoy, C. Aldrich, B.M. Herbst.: Performance of Convolutional Neural Networks for Feature Extraction in Froth Flotation Sensing. IFAC-PapersOnLine. **50**(2), Pages 13-18 (December 2017)
23. M.G.F. Costa, J.P.M. Campos, G.A. Aquino, W.C.A. Pereira, C.F.F.C. Filho.: Evaluating the performance of convolutional neural networks with direct acyclic graph architectures in automatic segmentation of breast lesion in US images. BMC Medical Imaging **19**(85), (2019)
24. Confusion Matrix for Your Multi-Class Machine Learning Model. https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826. Last accessed 22 jan 2022