PHONOCARDIOGRAM HEART SOUND SIGNAL CLASSIFICATION USING DEEP LEARNING TECHNIQUE

Nishant Sanjay Indalkar [1], Shreyas Shrikant Harnale[2], Prof. Namrata[3]

[1]Computer Engineering, Pimpri Chinchwad College of Engineering Pune, India
[2] Computer Engineering, Pimpri Chinchwad College of Engineering Pune, India
[3]Computer Engineering, Pimpri Chinchwad College of Engineering Pune, India

---***---

Abstract— *Most common reason for human mortality in today's world that causes almost one-third of deaths is especially due to heart disease. It has become the most common disease where in every 5 people 4 of them are dealing with this disease. The common symptoms of heart diseases are breath shortness, loss of appetite, irregular heartbeat, chest pain. Identifying the disease at early stage increases the chances of survival of the patient and there are numerous ways of detecting heart disease at an early stage. For the sake of helping medical practitioners, a range of machine learning & deep learning techniques were proposed to automatically examine phonocardiogram signals to aid in the preliminary detection of several kinds of heart diseases. The purpose of this paper is to provide an accurate cardiovascular prediction model based on supervised machine learning technique relayed on recurrent neural network (RNN) and convolutional neural network (CNN). The model is evaluated on heart sound signal dataset, which has been gathered from two sources:*

1. *From general public via I Stethoscope pro iPhone app.*

2. *From clinical trials in the hospitals. Experimental results have shown that number of epochs and batch size of the training data for validation metrices have direct impact on the training and validation accuracies. With the proposed model we have achieved 91% accuracy.*

Keywords— *CNN, RNN, Epochs, Deep Learning*

I. INTRODUCTION

Heart Disease is an illness that causes complications in human being such as heart failure, liver failure, stroke. Heart disease is mainly caused due to consumption of alcohol, depression, diabetes, hypertension [2]. Physical inactivity increase of cholesterol in body often causes heart to get weaken. There are several types of heart diseases such as Arrhythmia, congestive heart failure, stroke, coronary artery disease and many more. Identification of cardiovascular disease can be done by using the widely known auscultation techniques based on echocardiogram, phonocardiogram, or stethoscope. Machine learning and deep learning is a widely used method for processing huge data in the healthcare domain. Researchers apply several different deep learning and machine learning techniques to analyze huge complex medical data, to predict the abnormality in heart disease. This research paper proposes heart signal analysis technique based on TFD (Time Frequency Distribution) analysis and MFCC (Mel Frequency Cestrum Coefficient). Time Frequency Distribution represents the heart sound signals in form of time vs frequency simultaneously and the MFCC determines a sound signal in the form of frequency coefficient corresponding to the Mel filter scale [3]. A quite Helpful method was used to improve the accuracy of heart disease model which is able to predict the chances of heart attack in any individual. Here, we present a Deep Learning technique based on Convolutional Auto-Encoder (CAE), to compress and reconstruct the vital signs in general and phonocardiogram (PCG) signals specifically with minimum distortion [4]. The results portray that the highest accuracy was achieved with convolution neural network with accuracy 90.60% with minimum loss and accuracy achieved through recurrent network was about 67% with minimum loss percentage.

II. LITERATURE REVIEW

Ryu et al. [5] Studied about cardia diagnostic model using CNN. Phonocardiograms(PCG) were used in this model. It can predict whether a heart sound recording is normal or not. First CNN is trained to extract features and build a classification function. The CNN is trained by an algorithm called back propagation algorithm. The model then concludes between normal and abnormal labels.

Tang et al. [6] Combined two methods i.e. deep learning and feature engineering algorithms for classification of heart sound into normal and abnormal. Then features were extracted form 8 domains. Then, these features were fed into convolution neural network(CNN) in such a way that the fully connected layers of neural network replaces the global average pooling layer to avoid over fitting and to obtain global information. The accuracy, sensitivity and specificity observed on the PhysioNet data set were 86.8%, 87%, 86.6% and 72.1% respectively.

Jia Xin et al. [7] Proposed a system in which heart sounds are segmented and converted using two classification method: simple softmax regression network (SMR) and CNN. Features were determined automatically through training of the neural network model instead of using supervised machine learning features. After working on both Softmax regression and Convolutional neural network(CNN) they found out CNN gave the highest accuracy. The accuracy achieved through CNN model is 93%.

Mehrez Boulares et al. [8] developed a model based for cardiovascular disease(CVD) recognition based on unsupervised ML and supervised ML methods based on CNN. The datasets are taken from PASCAL and PhysioNet. They have worked on PCG signals mainly and enhanced their focus on denoising and preprocessing. For feature selection they have used MFCC as it is the best out there now-a-days and similarly for classification they've gone with CNN. Classification results of defined models had overall accuracy of 0.87, and overall precision of 0.81, and overall sensitivity 0.83.

Suyi Li et al. [9] proposed a paper to provide an overview of computer aided sound detection techniques. They have worked on PCG signals and characteristics of heart sounds introduced first. They did a thorough review on preprocessing and analyzing techniques that have developed over the last five years. They've further done a deep research on denoising, feature extraction, segmentation, classification and most importantly computer aided heart detection techniques.

Raza et al. [10] Proposed a framework for denoising the heart sound by applying band filter, them the size of sample rate of each sound is fixed. Then the features were extracted using sampling techniques and reduce the dimension of frame rate. RNN method is used for classification. RNN using Long Short-Term Memory (LSTM), Dropout, Softmax and Dense layers used. Hence, the method is more accurate compared to other methods.

Perera et.al [11] developed a software tool to predict heart abnormalities which can be recognized using heart sounds.

The audio inputs are taken through e- stethoscope and then entered into a database with symptoms of each patient. Feature extraction is done using MATLAB "MIR toolbox" and prominent features and statistical parameters are extracted.

Segmentation is done by tall peak and short peak process followed by classification of S1 and S2 systole and diastole square of wavelet fourth detail coefficient method was used for further classification process.

Yadav et al. [12] They proposed a model to extract discriminatory features for machine learning which involves strategic framing and processing of heart sound. They trained a supervised classification model based on most prominent features for identification of cardiac diseases. The proposed method achieved the accuracy 97.78% with error rate of 2.22% for normal and abnormal heart disease classification.

Khan et al. [13] Proposed a model based on different classifiers such as (KNN, Bagged Tree, Subspace, subspace Discriminant, LDA, Fine Tree and Quadratic SVM) to obtain and accuracy and results. Kaggle dataset was used to extract features from the sets of different domains i.e. frequency domain, Time domain and statistical domain to classify the heart sounds in two different classes i.e. normal and abnormal. Out of 6 classifiers the highest accuracy of 80.5% was obtained using Bagged tree.

III. PROPOSED MODEL

The study aims in classifying the heart sounds of heart disease patients into normal and abnormal heart sound based on Phonocardiogram signals.

The dataset was obtained from Physionet website which contains physionet challenge heart dataset which was providd publicly. There were two challenges related with this competition. Dataset contains heart sounds of 3 to 30 sec in length [4]. The proposed model described ahead is divided into three main parts. These three parts are preprocessing, train-test and classification.

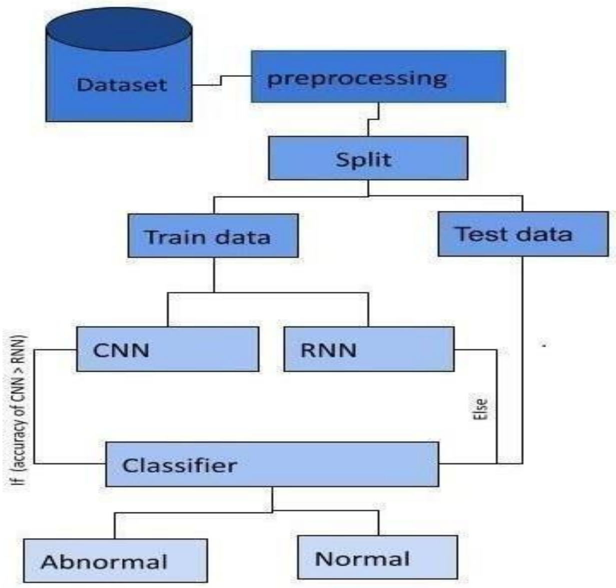


Fig 1: Proposed Architecture

The classification was performed using the most robust techniques which gave the highest accuracy among CNN and RNN. The main motive of this paper is to predict whether the heart sound is normal or abnormal with its confidence value. The first step is preprocessing which includes data compression, feature extraction where denoising was performed in order to remove the unwanted noise and were enhanced to remove the unwanted frequencies. The dataset used for training and testing in this model consist of phonocardiogram sound signal files which contains normal and abnormal heart sounds. These files are audio recordings of heart sounds at various different stages of heart beat.

The denoising and compression of heart sound in this model takes place while building the model by using convo 2D autoencoder[18]. Using the functional API convolution autoencoder was build and once the model was build autoencoder was trained using train_data as both our input data and target data [14]. The convolution encoder consists of MaxPooling 2D layers and 2D convo stack for max down sampling. Below shown are the graphs of loss and accuracy achived after using convo2D autoencoder with 100 epochs.

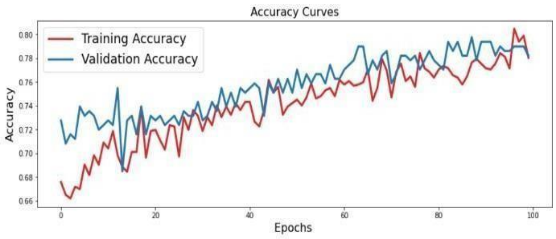


Fig 2 : Loss curve with 100 epochs

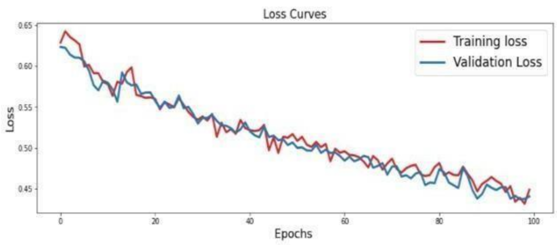


Fig 3: Accuracy curve with 100 epochs

In feature extraction the informative and relevant features were extracted from PCG signals using Mel-scaled power spectrogram and Mel-frequency cepstral coefficients (MFCC) which are then fed into a classification model to classify each PCG signal into an abnormal or normal heart sound [15]. The MFCC which was introduced by Mermelstein in the 1980s is widely used in automatic speech recognition. The mel-frequency analysis is based on human acoustic perception and experimental results have shown that human beings ear acts as a filter that focuses on certain level of frequency components. It transmits audio signal of certain frequency level and directly ignores the unwanted and undesired signals. In mfcc it converts the audio signal from analog to digital format with sampling frequency. It basically includes:

a. Filter-Bank: Filtering out high frequency sound signals to balance the sound wave.

b. Windowing the signal: The sound signals of time varying signal. For sound, signal needs to be examined over a short period of time. Therefore, speech analysis is to be carried out on short segments across which the speech signal is assumed to be stationary. Short-term spectral measurements are typically carried out over 20 ms windows, and advanced every 10 ms.

c. Applying DFT: DFT is applied on windowed frame to convert it into magnitude spectrum.

d. Mel-spectrum: Fourier transformed is applied to Mel spectrum signal through a set of filters known as Mel-filter bank and by applying inverse DCT frequencies are wrapped on a mel-scale.

e. DCT: As sound signals are smoothened; the energy levels are correlated. so a set of cepstral coefficients are produced by Mel frequency coefficients.

f. Dynamic features: As cepstral coefficients contain information from a given frame they are referred as static features. The extra information about the temporal dynamics of the signal is obtained by computing first and second derivatives of cepstral coefficients.

Librosa.display method was used to display the audio signals in different formats such as wave plot, spectrogram. The waveplot represents the graph of heart beat signals as shown below:

X- axis: Time in (Sec)

Y-axis: Frequency in (Hz)

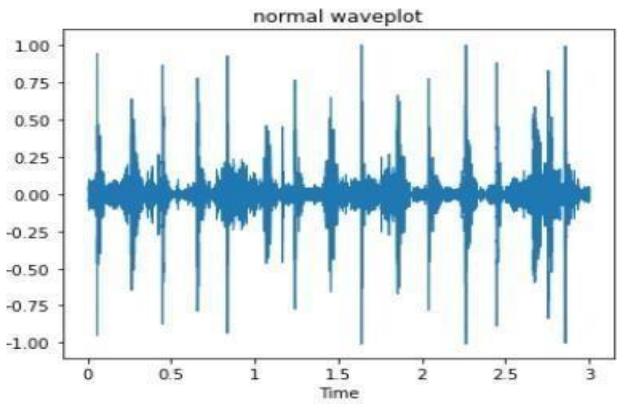


Fig 4: Normal heart sound waveplot

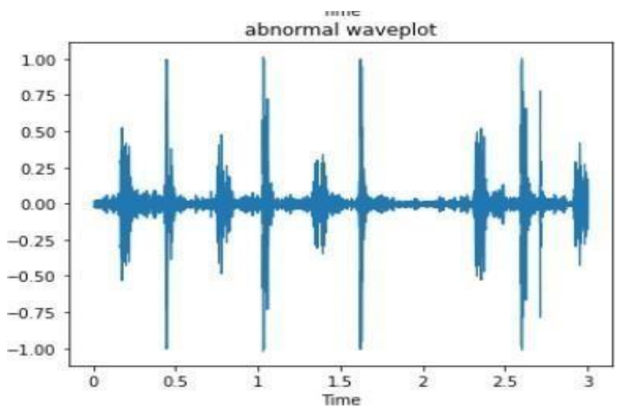


Fig 5: Abnormal heart sound waveplot

a) Mel-Scaled Power Spectrogram

Time period vs. Frequency representation of a sound signal is said to be spectrogram of signal. It graphically represents the change in frequency of a sound signal w.r.t time, which helps the building model to understand the sound accurately. The Mel-scaled filters present in Mel-scales are placed non- uniformly to mimic human ear properties in frequency axis [15].

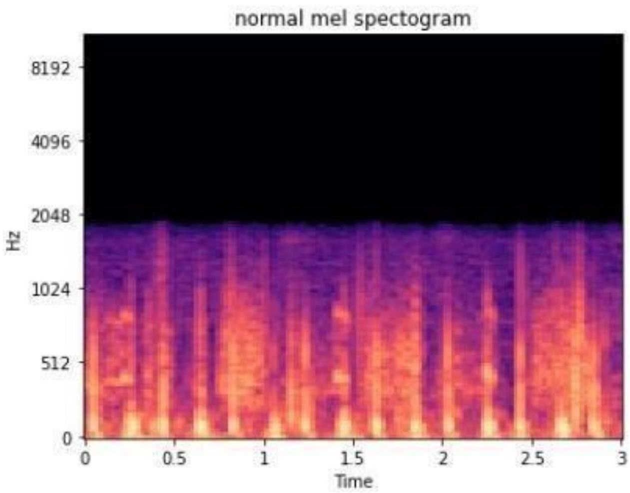
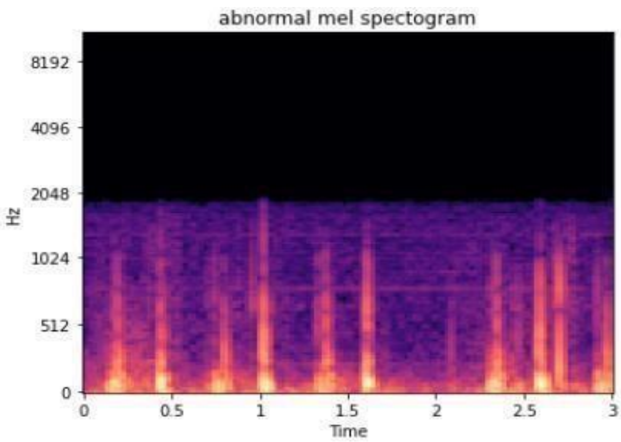


Fig 6: Normal sound mel-spectogram (Time vs Frequency)



b) Mel-Frequency Cepstral Coefficients

Mel-frequency cepstrum is found by taking Discrete Cosine Transform of a log power spectrum on a nonlinear Mel- scale of frequency. It is the representation of the Mel-scaled power spectrogram [15] [16]. Most of the extracted features for PCG heart signal are computed mainly using time, frequency.

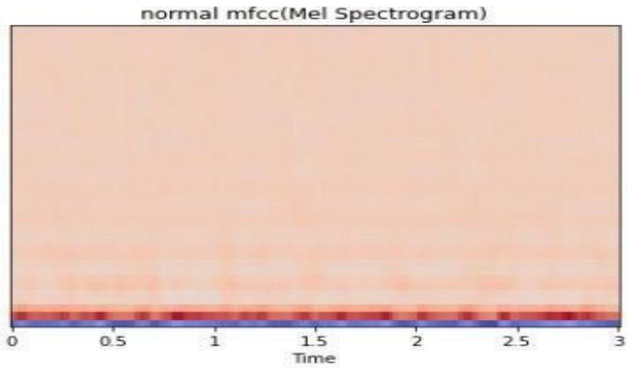


Fig 8: Normal sound mfcc (Time)

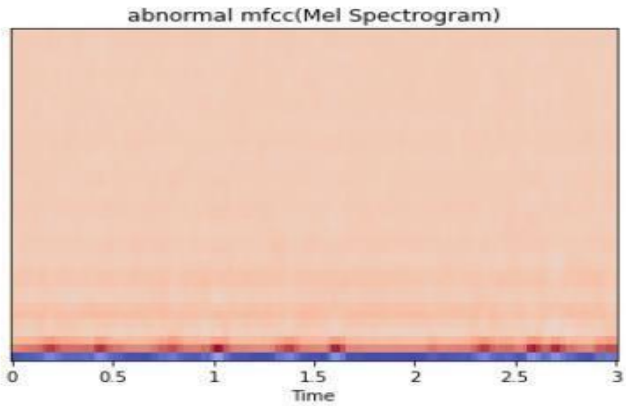


Fig 9: Normal sound mfcc (Time)

In classification stage the preprocessed data is fed to CNN [19] [20] [25] [26] and RNN [21] [29] [30], forTraining and testing. The model was trained using CNN and RNN and was build based on accuracy comparison of both the techniques. By comparing the accuracy and loss percentage of RNN and CNN as CNN has greater accuracy and less loss percentage than RNN thus using CNN for the prediction model was preferred. The model has an accuracy of 90.6% and test loss of 0.29 with 350 epochs and 128 batch size. The less test loss indicates that the model performs better after each iteration. The final prediction of model is categorized into normal and abnormal heart sound. Previous referred models from different researchers have had the same output which predictes if heart sound is normal or abnormal [20], what makes this model different from other researchers is the confidence value of normal or abnormal heart sound shown on the UI of classification model. Refer below figure

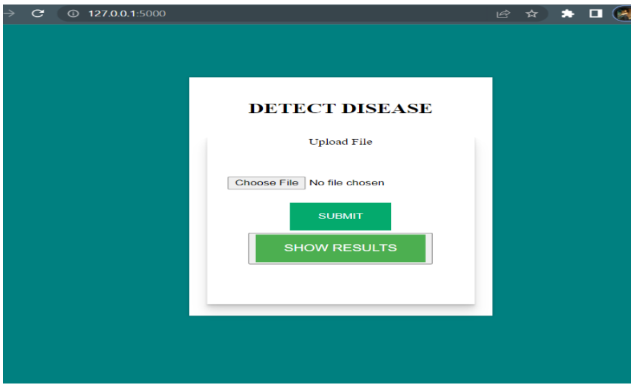


Fig 10: Proposed model (GUI)

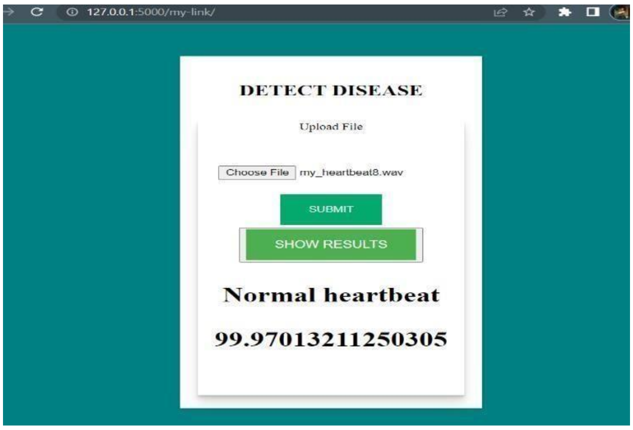


Fig 11: Normal heart sound with 99.97% confidence value

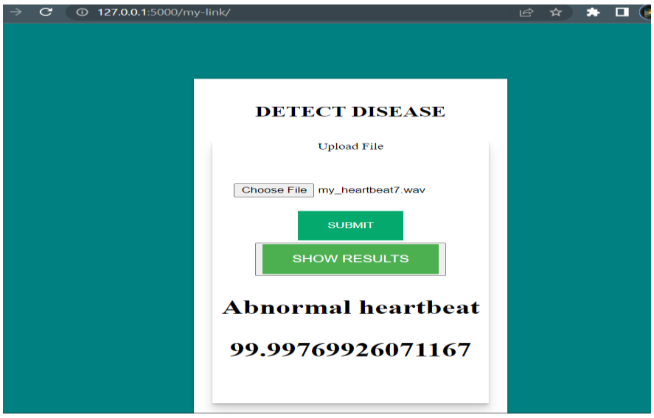


Fig 12: Abnormal heart sound with 99.99% confidence value

The confidence value of heart sound PCG signal will help the medical practitioners to identify the patients with greater risk of heart disease. The actual confidence value of heart sound makes a huge difference on the results because the higher or lower percent could help the physician to make better decisions.

IV. RESULT

Performance of the proposed Heart Sound Classification Techniques is shown in table 1. We applied Convolution neural network (CNN) and Recurrent neural network(RNN) in order to classify heart sound dataset described in above section. The dataset used for this project consists of Phonocardiogram signals of heart sounds containing heart sounds from 3 to 30 seconds in length. Preprocessing was performed in order to filter out the noisy data using auto-encoder and relevant features were extracted using mfcc. In order to test and train the proposed model the dataset was split into training and testing in the proportion of 80% - 20%. The primary objective of this paper is to examine the effects of the hidden layers of a CNN and RNN to check the overall performance of the neural network. To demonstrate this, we have applied CNN and RNN with different number of epochs on the given dataset and also to observed the variations in accuracy of both the techniques based on different number of epochs and batch size.

	Accuracy % of CNN	Loss % Of CNN	Accuracy % of RNN	Loss % Of RNN
300 epochs	90.82	0.32	73	0.57
350 epochs	90.60%	0.29	67	0.22

Table 1: Accuracy and loss percentage Convolution neural network outperformed recurrent

neural network having same number of epochs and batch size with accuracy 90.82% and 90.60% with 300 and 350 epochs. The precision, recall and F1 score was calculated for normal and abnormal heart sounds using CNN. Precision, recall, F1 of normal heart sounds were 0.83, 0.96, 0.89 and those for abnormal heart sounds were precision 0.97, recall 0.89 and f1 0.93

V. CONCLUSION

In this paper, we have compared the accuracies of two different neural networks i.e. Convolution Neural Network and Recurrent Neural Network based on the implemented model and proposed a model based on the technique which has performed well i.e. CNN with 90.60%. The model was trained using different number of epochs to ensure the model was not overfitted or underfitted and a constant number of epochs were chosen to train both the algorithms where they have performed well with having highest accuracy and less loss percentage with highest precision value.

V. REFERENCES

[1] M. Tschannen, T. Kramer, G. Marti, M. Heinzmann and T. Wiatowski, "Heart sound classificationusing deep structured features," 2016 Computing in Cardiology Conference (CinC), 2016, pp.565-568.

[2] Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande and Parneet Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning",Volume 2021 Article ID 8387680.

[3] I. Kamarulafizam, Shussain Salleh, Mohd Najeb Jamaludin, " Heart Sound Analysis Using MFCC and Time Frequency Distribution" doi: 10.1007/978-3-540-68017-8_102.

[4] W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, C. Xu, "Deep Learning Methods for Heart Sounds Classification: A Systematic Review" Entropy (Basel). 2021;23(6):667 Published 2021 May 26 doi:10.3390/e23060667.

[5] Heechang Ryu, Jinkyoo Park, and H. Shin, "Classification of heart sound recordings using convolution neural network," 2016 Computing in Cardiology Conference (CinC), 2016, pp. 1153- 1156.

[6] Hong Tang, Ziyin Dai, Yuanlin Jiang, Ting Li and Chengyu liu, "PCG Classification Using Multidomain Features and SVM Classifier" Volume 2018 |Article 4205027 | doi: 10.1155/2018/4205027.

[7] Jia Xin L and Keng Waah Choo, "Classification of Heart Sounds Using Softmax Regression and Convolutional Neural Network" ICCET '18: Proceedings of the 2018 International Conference on Communication Engineering and Technology February 2018 Pages 18–21.

[8] Mehrez Boulares, Reem Al-Otaibi, Amal Almansour and Ahmed Barnavi, "Cardiovascular Disease Recognition Based on Heartbeat Segmentation and Selection Process" October 2021 International Journal of Environmental Research and Public Health 18(20):10952. Suyi Li, Feng Li , Shijie Tang and Wenji xiong, "A Review of Computer-Aided Heart Sound Detection Techniques" Volume 2020 Article ID 5846191.

[9] Suyi Li, Feng Li, Shijie Tang and Wenji Xiong, "A Review of Computer-Aided Heart Sound Detection Techniques" JF - BioMed Research International PB – Hindawi.

[10] A. Raza, A. Mehmood, S. Ullah, M. Ahmad, G. S. Choi, and B.-W. On, "Heartbeat Sound Signal Classification Using Deep Learning," Sensors, vol. 19, no. 21, p. 4819, Nov. 2019, doi: 10.3390/s19214819.

[11] I. S. Perera, F. A. Muthalif, M. Selvarathnam, M. R. Liyanaarachchi and N. D. Nanayakkara, "Automated diagnosis of cardiac abnormusing heart sounds," 2013 IEEE Point-of-Care Healthcare Technologies (PHT), 2013, pp. 252-255, doi: 10.1109/PHT.2013.6461332.

[12] A. Yadav, A. Singh and M. K. Dutta,"Machine learning-based classification of cardiac diseases from PCG recorded heart sounds" Neural Comput & Applic 32, 17843–17856 (2020) doi: 10.1007/s00521-019-04547-5.

[13] Younas Khan, Usman Qamar, Nazish Yousaf, Aimal Khan, "Machine Learning Techniques for Heart Disease Datasets: A Survey", ICMLC '19: Proceedings of the 2019 11th International Conference on Machine Learning and Computing February 2019 Pages 27–35.

[14] Ying-Ren Chien, Kai-Chieh Hsu and Hen Wai Tsao, "Phonocardiography Signals Compression with Deep Convolutional Autoencoder for Telecare Applications" Appl. Sci. 2020, 10, 5842; doi:10.3390/app10175842.

[15] T. H. Chowdhury, K. N. Poudel and Y. Hu, "Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals," in IEEE Access, vol. 8, pp. 160882-160890, 2020, doi: 10.1109/ACCESS.2020.3020806.

[16] M. Rahmandani, H. A. Nugroho and N. A. Setiawan, "Cardiac Sound Classification Using Mel-Frequency Cepstral Coefficients (MFCC) and Artificial Neural Network (ANN)," 2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE), 2018, pp. 22-26, doi: 10.1109/ICITISEE.2018.8721007.

[17] Pronab Ghosh, Sami Azam, Asif Karim, Mirjam Jonkman, MD. Zahid Hasan, "Use of Efficient Machine Learning Techniques in the Identification of Patients with Heart Diseases", ICISDM 2021: 2021 the 5th International Conference on Information System and Data MiningMay 2021 Pages 14–20.

[18] Abeer Z. Al-Marridi, Amr Mohamed and Aiman Erbad, "Convolutional Autoencoder Approach for EEG Compression and Reconstruction in m-Health Systems," 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC), 2018, pp. 370-375, doi: 10.1109/IWCMC.2018.8450511.

[19] Ximing Huai, Siriaraya Panote, Dongeun Choi,and Noriaki Kuwahara,"Heart Sound Recognition Technology Based on Deep Learning" Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Posture, Motion and Health: 11th International Conference, DHM 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020.

[20] Fan Li, Hong Tang, Shang and Klaus Mathaik," Classification of Heart Sounds Using Convolution Neural Neural Network" June 2020 Applied Sciences 10(11) 3956.

[21] M. F. Khan, M. Atteeq and A. N. Quereshi, "Computer Aided Detection of Normal and Abnormal Heart Sound using PCG" ICBBT'19: Proceedings of the 2019 11th International Conference on Bioinformatics and Biomedical Technology May 2019 Pages 94–99 doi: 10.1145/3340074.3340086.

[22] Suyi Li, Feng Li, Shijie Tang and Wenji Xiong, "A Review of Computer-Aided Heart Sound Detection Techniques" JF - BioMed Research International PB – Hindawi.

[23] I. Grzegorczyk, M. Solinski, M. Lepek, A. Perka, J. Rosinski, J. Rymko, K. Stepein and J. Gieraltowski "PCG classification using a neural network approach," 2016 Computing in Cardiology Conference (CinC), 2016, pp. 1129-1132.

[24] K. K.Tseng, C. Wang, Y.F. Huang, G. R. Chen, K. L. Yung and W. H. Ip, "Cross-Domain Transfer Learning for PCG Diagnosis Algorithm" Biosensors 2021, 11, 127. doi.org/10.3390/bios11040127.

[25] Noman Fuad, Ting Chee-Ming, S. Salleh and H. Ombao, "Short-segment Heart Sound Classification Using an Ensemble of Deep Convolutional Neural Networks," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 1318-1322, doi: 10.1109/ICASSP.2019.8682668.

[26] Sunjing, L. Kang, W. Wang and Songshaoshaui, "Heart Sound Signals Based on CNN Classification Research" ICBBS '17 Proceedings of the 6th International Conference on Bioinformatics and Biomedical Science June 2017 Pages 44–48 doi: 10.1145/3121138.3121173.

[27] Low, Jia Xin and Keng Wah Choo. "Automatic Classification of Periodic Heart Sounds Using Convolutional Neural Network." World Academyof Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering 12 (2018): 96-101.

[28] D. R. Megalmani, S. B. G, A. Rao M V, S. S. Jeevannavar and P. K. Ghosh, "Unsegmented Heart Sound Classification Using Hybrid CNN-LSTM Neural Networks," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 713-717, doi: 10.1109/EMBC46164.2021.9629596.

[29] Y. Chen, Y. Sun, and J. Lv, "End-to-end heart sound segmentation using deep convolutional recurrent network" Complex Intell. Syst. 7, 2103–2117 (2021). doi: 10.1007/s40747-021-00325-w.

[30] C. Thomae and A. Dominik, "Using deep gated RNN with a convolutional front end for end-to-end classification of heart sound," 2016 Computing in Cardiology Conference (CinC), 2016, pp. 625-62