# RHEUMATIC HEART DISEASE DETECTION USING DEEP LEARNING FROM SPECTRO-TEMPORAL REPRESENTATION OF UN-SEGMENTED HEART SOUNDS

**GUIDE: DR. M. Vimala Devi[1]**

**Agilan R[2], Elango M[3], Guna M[4], Sathishkumar Thangavel[5]**

*1. Associate Professor& Head, KSR Institute for Engineering and Technology, Tiruchengode.*
*[2,3,4,5] Department of Computer Science and Engineering,*
*KSR Institute for Engineering and Technology, Tiruchengode.*

**ABSTRACT**

*Rheumatic Heart Disease (RHD) is an autoimmune response to a bacterial attack which deteriorates the normal functioning of the heart valves. The damage on the valves affects the normal blood flow inside the heart chambers which can be recorded and listened to via a stethoscope as a phonocardiogram. However, the manual method of auscultation is difficult, time consuming and subjective. In this study, a convolutional neural network based deep learning algorithm is used to perform an automatic auscultation and it classifies the heart sound as normal and rheumatic. The classification is done on un-segmented data where the extraction of the first, the second and systolic and diastolic heart sounds are not required. The architecture of the CNN network is formed as an array of layers. Convolutional and batch normalization layers followed by a max pooling layer to down sample the feature maps are used. At the end there is a final max pooling layer which pools the input feature map globally over time and at the end a fully connected layer is included. The network has five convolutional layers. This current work illustrates the use of deep convolutional neural network using a Mel Spectro-temporal representation. For this current study, an RHD heart sound data set is recorded from one hundred seventy subjects from whom one hundred twenty four are confirmed RHD patients. The system has an overall accuracy of 96.1% with 94.0% sensitivity and 98.1% and specificity.*

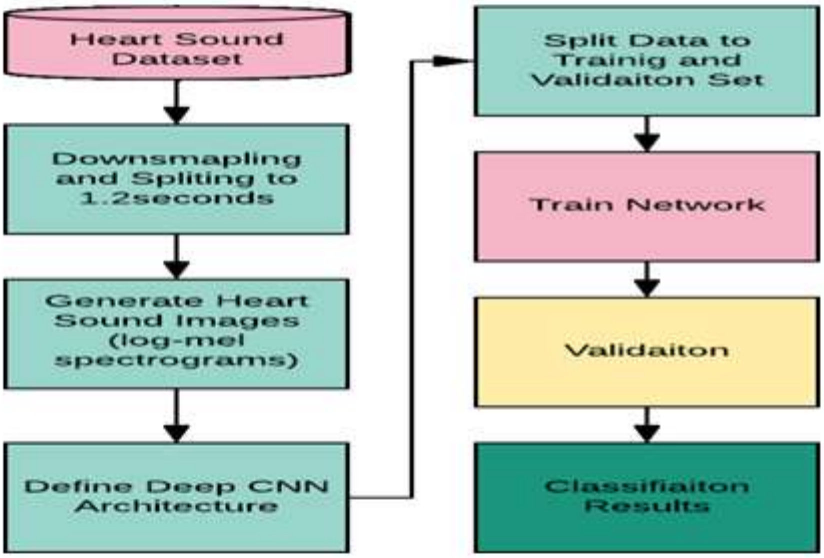
***Keywords*: Bacterial attack, CNN network, Deep learning algorithm, Heart sound data, pooling layer**

## I. INTRODUCTION

Cardiovascular diseases (CVDs) are world's leading cause of death. The majority of cardiovascular disease causes can be prevented and treated if they are detected early. In developing countries, the major cause of CVDs is RHD. According to the Global Burden of Disease report RHD is a significant and preventable cause of morbidity and mortality for people in developing nations that, even with incomplete data, affects at least 33 million people and causes more than 300 000 deaths annually, particularly among vulnerable and marginalized ups including children, adolescents and pregnant women. This chronic heart disease, which is caused by a repeated infection of group a streptococcal bacteria predominantly affects the heart valves. The damage on the valves affects the normal blood flow in the heart chambers. This causes a turbulent blood flow in the heart which is referred to as murmur. This murmur can be recorded and listened to using a stethoscope as a form of phonocardiogram (PCG). Yet the manual method of auscultation is complex and difficult to master. It is subjective and time-consuming too.

### 1.1 Process & Solution:

PCG is an acoustic recording of the heart vibrations using a microphone transducer. Nowadays there exist several electronic stethoscopes which record digital heart sound data. A PCG-based automatic detection tool provides a low cost and non-invasive solution which avoids the subjectivity. Due to the availability and accessibility of advanced health care to the general public, this approach will be particularly interesting for developing countries.Automatic audio recognition is a growing area of research with various applications in the real world. While there is a wide range of studies in speech and music, work on classifying heart sounds is comparatively limited. However, automatic heart sound data classification has become a major topic of research especially after the PhysioNet/CinC Challenge2016. However the research works were hampered by the unavailability of large and reliable data. It should be noted that for noisy signals most current heart sound classification techniques don't work properly. Potes et al. have developed a decision rule for classifying normal / abnormal heart sounds based on a classifier set that combines AdaBoost outputs and CNN training using 124 time frequency features. The algorithm was trained on a training data set (normal=2575, abnormal=665) provided by PhysioNet's computing in the cardiology challenge. With 94.2% sensitivity and 77.8% specificity and overall accuracy of 86.02 percent, the classifier achieved the highest score of the competition.



## II. MATERIALS AND METHODS

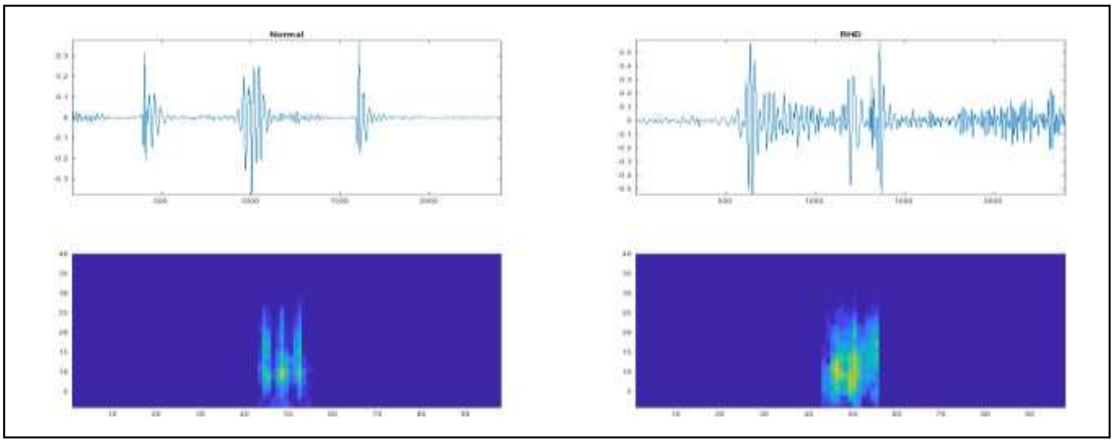
The boom in deep learning is mainly driven by its success in computer vision and speech recognition. However, when it comes to biosignal processing there is less research and there are no pre-trained networks. To leverage the techniques and the insights brought by the recent developments of computer vision and speech recognition, this paper presents a novel approach to encode a time series heart sound signal as images applying 2D CNNs. The flow diagram of the whole system.First the raw heart sound signal is downscaled from 44.1 kHz to 2 kHz and split to 1.2seoncd of data. Then it each 1.2second segment is converted into its log-Mel-spectral representation; if required the data augmentation step can be incorporated to increase the training data. Then the deep CNN network architecture is defined followed by the actual training phase. After the validation step, the classification results are displayed. The overall classification system flow diagram.

### 2.1 Data Collection:

The preliminary data is collected in Ethiopia. The country is one of the endemic countries for RHD where the disease is seriously affecting the youth community. The heart sound data was collected at Tikur Anbessa Referral Teaching Hospital, College of Health Sciences, Addis Ababa University, and Addis Ababa, Ethiopia from August 2018 to July 2019. The study protocol was approved by the Research Ethics Committee of the Department of Internal Medicine (Ethical Clearance No: 014/2018). The data were recorded from 170 subjects, 124 were confirmed RHD patients (74 females, 50 males) with ages from 9 to 47 with mean and std of 22.9±8.9 years. The time since the first diagnosis is from two months to 20 years with mean and std of 3.3±3.1 years. Each diagnosis is confirmed by echocardiographic imaging and a cardiologist analysis. There were 46 normal subjects (15 females, 31, males) with age 5 to 37 years with mean and STD of 14.4 ±10.5 years. The electronic heart sound was recorded by ThinklabsOne™ digital stethoscope with a sampling rate of 44.1 KHz. The first experiment is done on our collected heart sound data after down sampling the data to 2 kHz. A total of 33453 (22156 RHD and 11297 Normal) records with 1.2-second duration are used. The corresponding log-Mel heart sound spectrogram is computed to generate a pictorial representation of the heart sound clips. The data is recorded in an uncontrolled environment and many of the recordings are corrupted by various types of noises such as movement artifacts, talking, mobile phone interference, traffic sound, coughing, lung sounds, gastrointestinal sounds, pounding and clicks due to high volume recording. Finally, approximately 20% of the data is hold-out by making sure that a heart sound record from the same person is not used for validation and training.

### 2.2 Heart Sound Spectrograms:

The CNN model uses two dimensional images as an input hence the heart sound signal must be converted into a two dimensional representation. By transforming the onedimensional heart sound signal into two-dimensional representation using spectrograms, the need of feature extraction is not needed. The heart sound signal is first down sampled to 2 kHz and is converted into log-Mel spectrograms to prepare the data for a convolutional neural network's effective training. The log scale is preferred over the linear scale as it roughly resembles the resolution of the human auditory system. The duration of each heart sound clip is 1.2 seconds which is the maximum duration for a single heartbeat. Forty log-Mel filters are used in computing the spectrograms. The following Figure shows the plot of a few training examples with the waveforms and corresponding spectrograms.



### 2.3 Data Augmentation:

Data augmentation is one of the most effective strategies for improving the performance of classification in image classifiers. Data augmentation enlarges the training set which will increase the classification accuracy. It will also avoid over fitting. This step also creates different viewpoints of a single image which increases the robustness of the system. The regularly used techniques of augmentation such as rotation flipping and cropping do not work in spectrograms. Hence the augmentation is done by randomly translating the spectrogram up to 10 frames forward or backward in time and scaling up or down the spectrograms along the time axis by 20%.

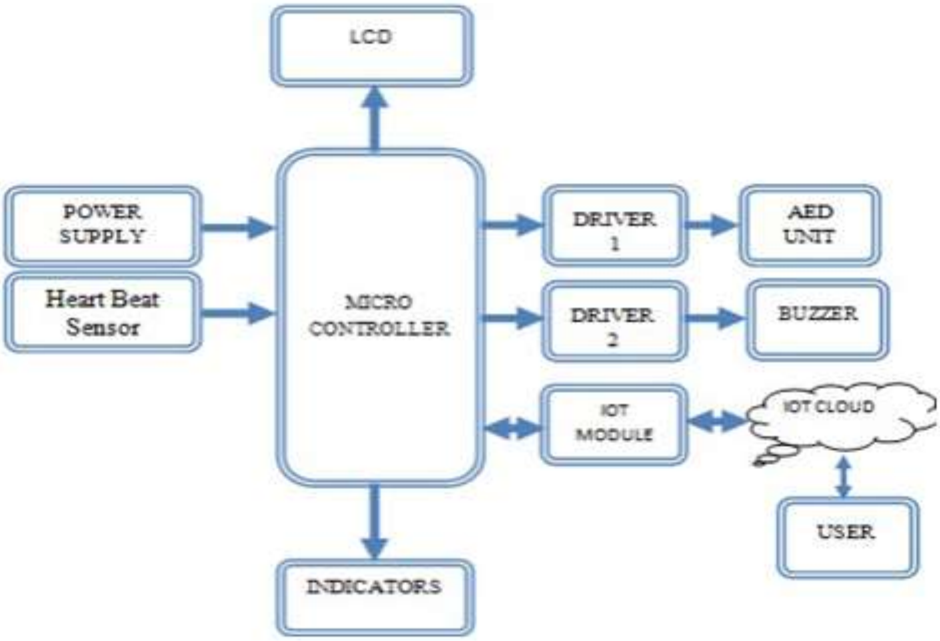
### 2.4 Neural Network Architecture:

The architecture of the CNN network is formed as an array of layers. The layers are convolutional, rectified linear unit (ReLU), batch normalization and max pooling layers. The directed acyclic graph (DAG) of the network architecture used in this work. The input image has a size of 40x98 pixels. The network has five convolutional layers with 12 filters at the start of the convolution layer with 3X3 kernel. The second convolution layer has 24 filters and the remaining three have 48 filters each. After convolutional and batch normalization layers, we down sample the feature maps using max pooling layers. A final max pooling layer that pools the input feature map globally over time. Global pooling is used to reduce the number of parameters in the final fully connected layer. To reduce the possibility of the network memorizing specific features of the training data, a 20% dropout to the input to the last fully connected layer is added.

### 2.5 Evaluation:

The heart sound data is split as training set and validation set. A heart sound record from the same person is not used for validation and training. To achieve this records from one subject are stored in one folder. Then 20% of the data is set aside for validation. This 20% validation data is predefined hence none of these validation records are used in the training.

## III. RESULTS:




### 3.1 Experiment I:

In this first experiment, a data augmentation step as discussed in the methodology is incorporated. The system has generated an overall accuracy of 96.1% with 94.0% sensitivity and 98.1% specificity as shown in Table 1.

### 3.2 Experiment II:

The second experiment is done by removing the augmentation trainings data. For this case, the system has overall accuracy of 96.7% with 95.2% sensitivity and 98.2% specificity as shown in Table 1.

| **Experiments** | **Sensitivity** | **Specificity** | **Accuracy** |
|---|---|---|---|
| I | 94.0% | 98.1% | 96.1.0%. |
| II | 95.2% | 98.2% | 96.7% |

**TABLE I. EXPERIMENTAL RESULTS**

## IV. CONCLUSION:

It has been demonstrated that deep convolutional networks which are designed for image recognition can be successfully trained to classify heart sound spectral images. Our trained model obtained an overall accuracy of 96.7% with 95.2% sensitivity and 98.2% specificity in detecting RHD. This can help to develop timely, affordable and reliable access to cost-effective technologies for the detection and prevention of rheumatic heart disease.

## V. ACKNOWLEDGEMENTS:

I would like to thank my major professor DR. M. VIMALA DEVI for her constant guidance and help throughout the project. I would also like to thank Mrs. Hamsareka (Assistant Professor) for graciously accepting to be on my committee.

## VI. REFERENCE:

[1] Emelia J. Benjamin, et al., Heart Disease and Stroke Statistics—2019 Update: A Report from the American Heart Association, Circulation, Volume 139, No. 10, 2019.

[2] The Global Burden of Disease Study 2010 (GBD 2010)

[3] Watkins DA, Johnson CO, Colquhoun SM, Karthikeyan G, Beaton A, Bukhman G, et al. Global, regional, and national burden of rheumatic heart disease, 1990-2015. N Engl J Med. 2017; 377(8):713-22.

[4] Cristhian Potes , Saman Parvaneh , Asif Rahman, Bryan Conroy, Ensemble of Feature-based and Deep learning-based Classifiers for Detection of Abnormal Heart Sounds, PhysioNet/CinC Challenge, 2016.

[5] Bernhard Suhm (2020). Heart Sound Classifier (https://www.mathworks.com/matlabcentral/fileexchange/65286- heart-sound-classifier), MATLAB Central File Exchange. Retrieved January 12, 2020.

[6] Yu Su, Ke Zhang, Jingyu Wang and Kurosh Madani, ―Environment Sound Classification Using a Two-Stream CNN Based on DecisionLevel Fusion‖ Sensors 2019

[7] Venkatesh Boddapati, Andrej Petef Jim Rasmusson, Lars Lundberg, ―Classifying environmental sounds using image recognition networks‖ , Procedia Computer Science, 2048–2056. 2017.

[8] Chen Hao, Sandi Wibowo, Maulik Majmudar and Kuldeep Singh Rajput, ―Spectro-Temporal Feature Based Multi-Channel CNN for ECG Beat Classification‖ IEEE, 2019

[9] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, ―Cardiologist-level arrhythmia detection with convolutional neural networks, ‖ arXiv preprint arXiv: 1707.01836v1, 2017.

[10] H. Gao, L. Zhuang, L. V. D. Maaten, and K. Q. Weinberger, ―Densely connected convolutional networks, ‖ 2017.

[11] Yadeta D, et al. Spectrum of cardiovascular diseases in six main referral hospitals of Ethiopia, Heart Asia 2017; 9:1–5. Doi: 10.1136/heartasia-2016-010829. 2017.

[12] M. Brookes, VOICEBOX: A speech processing toolbox for MATLAB," Imperial College, Software Library, 2011. [Online]. Available: http://www.ee.imperial.ac.uk/hp/sta_/dmb/voicebox/voicebox.html

[13] Agnieszka Mikołajczyk, Michał Grochowski, ―Data augmentation for improving deep learning in image classification problem‖ International Interdisciplinary PhD Workshop (IIPhDW), 2018.

[14] https://archive.physionet.org/pn3/challenge/2016/, last accessed on January 5, 2020.

[15] Thiyagaraja, S.R.; Dantu, R.; Shrestha, P.L.; Chitnis, A.; Thompson, M.A.; Anumandla, P.T.; Sarma, T.;Dantu, S. A novel heart-mobile interface for detection and classification of heart sounds. Biomed. Signal Process. Control 2018, 45, 313–324.

[16] Mostafa Abdollahpur, Ali Ghaffari, Shadi Ghiasi, M J M, "Detection of pathological heart sounds", Physiol. Meas, vol. 38, 2017.