International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

Multi-Class Cardiac Diagnostic Decision Support System Based On Phonocardiogram Signal

Lijisha T[1], Ramitha MA[2]

[1]*PG Scholar, Dept. Computer Science & Engineering, M.Dasan Institute Technology Ulliyeri, Koyilandi, Kozhikoode ,673620*

[2]*Assistant Professor, Dept. Computer Science & Engineering, M.Dasan Institute Technology Ulliyeri, Koyilandi, Kozhikoode ,673620*

ABSTRACT-
A Phonocardiogram (PCG) signal is the representation of signal like murmurs and sounds which made by the vibrations caused for the period of cardiac cycle . Where the heart beat is recorded using a low-cost small handled digital device called as stethoscope . By using this device it provide information the heart rate , intensity , tone , frequency , quality and the location of the various components of the cardiac sound . Due to these characters , phonocardiogram signals ca detect the heart status at the early state and go for the treatment . Diagnoses at the early stage is the only way to decrease the death rate due to cardiac vascular diseases (CVD) . There are many invasive and non invasive method to diagnoses the cardiac vascular diseases . The facilities are no available in the low and middle income areas where the lack of availability of facilities or lack of money the facilities are no easy available and it case death . In previous studies , it uses convolutional neural network (convnet) , which is trained by hybrid constant – Q transform (HCQT) for heart beat sound classification and most studied architecture . Constant Q – transform (CQT) , variable – Q transform (VQT) and hybrid constant Q-transform which is extracted from phonocardiogram signals as the acoustic features , which includes the domains of Mel Frequency Cepstral Coefficients (MFCCs) where audio or speech signal processing . In the proposed system convolutional neural network & CQT, Variable-Q Transform (VQT), and HCQT are extracted from each phonocardiogram signal as the acoustic features, including the dominant MFCC features, feed into five-layer regularized convnet like convolution layer , pooling layer and dense layer. After analyzing the literature in the same domain, it can be stated that this is the first time HCQT is being utilized for PCG signals. HCQT is more effective than standard CQT and other variants. Also, the accuracies of the system proposed in this work on the validation datasets are 94% in multi-class classification, which outperforms the proposed work relative to other models significantly.

INDEX TERMS - *Cardiovascular disease, convolutional neural network, decision support system deep learning, multi-class classification, phonocardiogram signal*.

I. INTRODUCTION

As per the fact sheet available with the world health organization (WHO) , cardiac vascular diseases (CVD) cause the death of around 17.9 million people each year ,and it is around 31% of overall death in a year , which shows that the CVD disease is the major reason for the death causes . Mainly the CVD death occurs in two countries , in low & middle income countries where the availability of the facilities is less and also having the facility with high cost . The only way to reduce the death rate due to CVD is diagnose at early stage . There are two methods to diagnose the CVD diseases , where one is invasive and other one is non-invasive method . Where the invasive techniques are very costly , pain full and readily unavailable at every places , which is especially not available in remote areas . The non- invasive method is less expensive and painless to diagnoses the CVD diseases . Where ECG and PCG are the two such non-invasive ways to diagnose the CVD . But their analysis requires an expert doctor of this domain which is not readily available in remote areas . A Phonocardiogram (PCG) signal is the representation of signal like murmurs and sounds which made by the vibrations caused for the period of cardiac cycle .

DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

Where the heart beat is recorded using a low- cost small handled digital device called as stethoscope . By using this device it provide information the heart rate , intensity , tone , frequency , quality and the location of the various components of the cardiac sound . Due to these characters , phonocardiogram signals ca detect the heart status at the early state and go for the treatment . Diagnoses at the early stage is the only way to decrease the death rate due to cardiac vascular diseases (CVD) . Recent advances in computing have enabled researchers to design decision support systems that can be utilized to diagnose CVD at an early stage, even in the absence of an expert. Machine learning and deep learning algorithms have allowed us to create decision support systems that can help doctors and can also be used by laypeople in the absence of doctors .

The hybrid constant-Q transform based classification model to acquire more detailed information from PCG signals in this work. Acoustic features from the PCG signal are fetched to the ConvNet model for learning.

The following are the key contributions of the proposed work:

• Propose hybrid constant-Q transform based (HCQT) acoustic features for PCG signals.

• Compare the HCQT features to other acoustic features and recommend the best feature set for PCG signal classification.

The following is the paper's structure : Discussion of different models found in the literature for automatic diagnosis of CVD from PCG is given in Section II . Details of sound features used with the model for classification, classifier, an insight view of the proposed model, and features of the phonocardiogram signal dataset used for the training and testing of the designed model are given in Section III . Detail of the simulation environment and result generated through the proposed model are given in Section IV . Discussion and analysis of results are presented in Section V . It is ended with the conclusive remarks given in Section VI .

II. LITERATURE REVIEW

An overview of different types of automatic heart disease diagnostic models from PCG signal along with datasets used and accuracy level achieved by them is given below Table 1 . Though in the last five years, a lot of research has been carried out in designing of automatic heart disease diagnosis model from PCG signal, yet there are many more areas that are yet to be explored.

III. MATERIAL AND METHODS

This section describe in detail all the materials that have been used to conduct a study as well as the procedures that are undertaken. As research writing should be orderly and organized therefore the materials in each of its sub-section should be presented in a logical manner. And also detailed overview of sound feature extraction methods, classification model, the dataset used, and proposed model utilized in this work.

A. MEL FREQUENCY CEPSTRAL COEFFICIENTS (MFCCs)

In audio or speech signal processing, The short-term power spectrum of sound is represented by MFC. It is based on a non-linear Mel frequency scale and a linear cosine translation of the logarithmic power spectrum. Collectively MFCCs coefficients make up MFC. The feature extraction process of MFCC is composed of the following steps [23], [24]:

1. Pre-emphasis: It amplifies high frequencies by passing phonocardiogram signals from a high pass filter.

2. Framing: Phonocardiogram signals are separated into overlapping frames. It is implemented to fetch local spectral properties.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

IMPORTANT FEATURES	CLASSIFIERS	DATASET	ACCURACY
Classification Of Heart Sound Signal Using Multiple Features	SVM , DWT & Centroid Displacement Base Kknn	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	Centroid Displacement Based Knn The Highest Accuracy Achieved Is 97.4%
Heart Diseases Diagnosis Using Intelligent Algorithm Based On Pcg Signal Analysis	SVM DWT	Pascal Pcg Signal Database Was Used For Training And Testing The Proposed Algorithm	Accuracy Of 97% For Pascal Heart Sound Database.
Multi-Class Heart Sounds Classification Using 2d - Convolutional Neural Network	2D -CNN	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	Achieves An Accuracy Of 83%.
Early Detection Of Heart Valve Disease Employing Multiclass Classifier	SVM LOGFBANK MFCC	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	The Method Achieved An Accuracy Of 97.50 % During The Classification Process.
Phonocardiogram Signal Processing For Automatic Diagnosis Of Congenital Heart Disorders Through Fusion Of Temporal And Cepstral Features	1D-LTPS MFCC SVM	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	Achieves A Mean Accuracy Of 95.24% In Classifying Asd, Vsd, And Normal Subjects.
Accurate Classification Of Heart Sounds For Disease Diagnosis By A Single Time-Varying Spectral Feature: Preliminary Results	SVM KNN	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	The Highest Accuracy For Both Type Of Classification Was Obtained With KNN Classifier.It Yielded An Accuracy Value Of 99.60% For two Class Classification And 96.50 % For Multiclass Classification.
Systolic Murmurs Diagnosis Improvement By Feature Fusion And Decision Fusion	SVM MLP	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	A Total Accuracy Of 92.5% And A Total Validity Of 92.4% Are Achieved.
Diagnosis Of Heart Diseases By A Secure Internet Of Health Things System Based On	SVM ANN	Pascal Dataset, Aen, Pascal B-Training And Physiobank –Physionet A-Training Heart Sound Datasets Were Used Accordingly.	Achieving The 100% And 99.8% Overall Accuracy Rates For The Two Most Commonly Used The Data Sets Of Heart Sounds Shows

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

Autoencoder Deep Neural Network			That The Obtained Results Are Not Random.
Use Of Machine Learning Techniques In Healthcare: A Brief Review Of Cardiovascular Disease Classification.	SVM CNN	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	Highest Accuracy Of Over 95% Were Attained By Ensemble Techniques
Deep Learning Based Cardiovascular Disease Diagnosis System From Heartbeat Sound	SVM DWT	Database Containing 5 Categories Of Heart Sound Signal (Pcg Signals)	It Has Found That Proposed Model Has Shown The Average Accuracy Of 94% While Doing The Classification Of Pcg Sound In Five Classes.

Table 1 : An overview of PCG signal based heart disease diagnosis models.

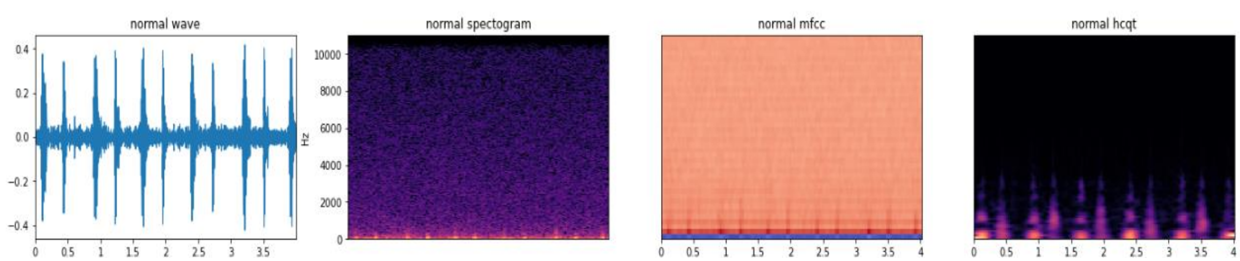
3. Windowing: It is implemented on frames for the minimization of discontinuities around edges. An example of a widely used technique is Hamming windowing.
4. Discrete Fourier Transformation: DFT is applied to the sound signal after the third step to obtain the frequency domain signal from the time domain.
5. Mel-Frequency Warping: It's used to calculate the quantity of energy that occurs in various locations of a frequency domain. Mel in this case is a pitch unit. A pitch of 1000 Mels is a pure tone at 1000Hz with a 40 dB strength over the listener's threshold. Mel-scale is used to determine this non-linear frequency result, as presented in (1).

$$M(f) = 1125\log(1 + f/700) \quad (1)$$

Here, the frequency term is denoted by f, while the Mel-scale frequency is denoted by M(f).
6. Discrete Cosine Transform and Log Compression: In this step, the logarithmic function IFFT is applied on filtered bank energies received in step 5. The DCT follows it. Finally, MFCC(n) is computed as shown in (2).

$$MFCC(n) = 1/T \sum_{r=1}^{R} \log [MF(t)] \cos [2\pi/T (r + 1/2) n] \quad (2)$$

where MFCC(n) is the nth MFCC coefficient derived from specific audio sections using T triangular filters, and MF(t) is the t-th filter's Mel-spectrum. The heartbeat spectrogram obtained by MFCC is shown in Fig. 1.



(a) (b) (c) (d)

FIGURE 1 : (a): A sample waveform for normal phonocardiogram signal, (b): heat map visualization for spectogram of a PCG signal segment (c): heat map visualization for MFCC of a

DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

PCG signal segment , (d): heat map visualization for HCQT of a PCG signal segment Sliding windows, x, and filter-bank frequencies, y, are represented on the horizontal and vertical axes. MFCC energy information, Ex,y, is represented by pixel color in the heat map. The MFCC is generated with the number of frequency bins = 84 and hop length = 512.

B. CONSTANT-Q TRANSFORM (CQT) , VARIABLE - Q TRANSFORM , AND HYBRID CONSTANT-Q TRANSFORM (HCQT)

J.C. Brown, in 1988 has introduced CQT. It refers to a technique that transforms a signal from time to frequency domain. However, it is different from Fourier transformation as central frequencies are geometrically spaced, and corresponding Q-factors are equal. CQT is defined as a 1/24 octave filter bank, but it is not restricted to 24 only; it can be varied to 12, 36, or 48 bins per octave also. Unlike DFT, central frequencies of analysis are not uniformly distributed but aligned with equally tempered scale notes; this makes CQT suitable for the processing of sound [25], [26]. Furthermore, the frequency resolution of CQT has a constant Q-factor, which effectively improves resolution accuracy in low-frequency regions. Under the N-th frame of CQT, the frequency component of the K-th semitone can be stated in (3).

$$X_n^{cqt}(k) = \frac{1}{N}\sum_{m=0}^{N_k-1} x(m)w_{N_k}(m)e^{-j2\pi mQ/N_k} \quad (3)$$

where Q is a constant whose value depends on the number of spectral lines of a single octave (β).

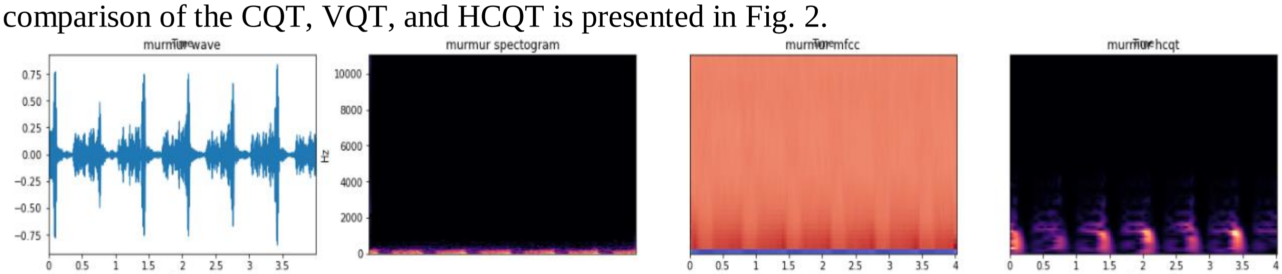
$$Q = \frac{1}{2^{\frac{1}{\beta}}}-1$$

The ability of the constant-Q transform to provide equal frequency support to all semitones and a variable number of bins among them is its main advantage. However, it has drawbacks, one of which being the absence of consistent temporal resolution at lower frequencies. This trade-off can be alleviated by introducing variants of CQT i.e., VQT and HCQT. When compared to the CQT transformation, the VQT transformation provides better temporal resolution at lower frequencies. A new parameter is introduced to allow for an equitable drop of the bins' Q-factors as it approach low frequencies [27], [28].

$$Bk = \alpha fk + \gamma$$

When γ = 0, the Q-factor in the constant-Q situation is a constant. The additional parameter γ might be understood as a Hertz offset, and it is normally set to be as low as possible, e.g., around 30 Hz. Instinctively, γ has a stronger relative influence at lower frequencies where the bandwidth is insufficient, but fades at higher frequencies. Hybrid CQT, on the other hand, is made up of two CQT varieties. In the temporal domain, the frameshift is thought to include L samples. Then, select the kc-th filter that fulfills the condition N [kc] = 2L [29], [30].

High frequencies are those that exceed f_kc, whereas low frequencies are those that are less than f_kc. The high frequency section of hybrid CQT uses the filter bank of the high-frequency part of CQT to filter the short-term Fourier transform-based spectrogram. The regular CQT is used directly for the low-frequency section of HCQT. In compared to CQT, HCQT is more computationally capable. A visualized comparison of the CQT, VQT, and HCQT is presented in Fig. 2.



(a) (b) (c) (d)

FIGURE 2 : (a): A sample waveform for murmur phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT base spectrograms, respectively. Sliding

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

windows, x, and filter-bank frequencies, y, are represented on the horizontal and vertical axes. MFCC energy information, Ex,y is represented by pixel color in the heat map. The MFCC is generated with the number of frequency bins = 84 and hop length = 512 .

C. CONVOLUTIONAL NEURAL NETWORK (ConvNet)

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of artificial neural network (ANN), most commonly applied to analyze visual imagery.[21] CNNs are also known as Shift Invariant or Space Invariant Artificial Neural Networks (SIANN), based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation-equivariant responses known as feature maps. Counter-intuitively, most convolutional neural networks are not invariant to translation, due to the down sampling operation that apply to the input. It have applications in image and video recognition, recommender systems , image classification , image segmentation , medical image analysis , natural language processing , brain–computer interfaces , and financial time series.

CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks make them prone to overfitting data. Typical ways of regularization, or preventing overfitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNNs take a different approach towards regularization: It take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore , on a scale of connectivity and complexity , CNNs are on the lower extreme . Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that cover the entire visual field. CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns to optimize the filters (or kernels) through automated learning, whereas in traditional algorithms these filters are hand-engineered . This independence from prior knowledge and human intervention in feature extraction is a major advantage.

D. PROPOSED PCG SIGNAL CLASSIFICATION MODEL USING ACOUSTIC FEATURES

The offered method for phonocardiogram signal classification using ConvNet is depicted in Fig. 4. The raw data provided is in Waveform Audio File Format (WAV) format, encoding phonocardiogram signals. To pass these sound waves to ConvNet model, these phonocardiogram signals are converted into an image, i.e. 2-D spectrogram. Spectrograms are convenient for representing these heartbeat recordings because it capture the intensity of the frequencies throughout a given sound. Thus, these spectrograms are effective representations of an audio recording. In this work, the proposed the use MFCC, CQT, VQT, and HCQT based spectrograms for phonocardiogram signal classification.

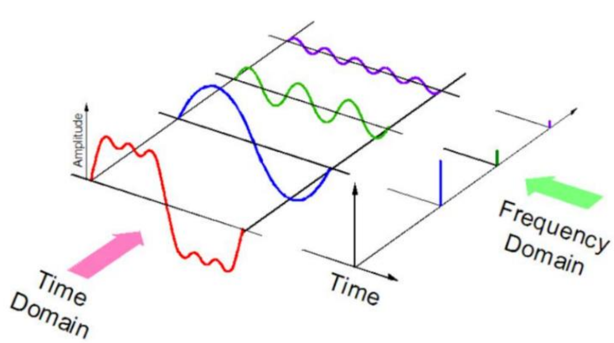


FIGURE 3 : Waveform Audio File Format (Time Domain & Frequency Doamian)

DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

The proposed method for phonocardiogram signal classification using ConvNet uses publicly available data. The raw data provided is in Waveform Audio File Format (WAV) for- mat, encoding phonocardiogram signals as shown in fig 3 .To pass these sound waves to ConvNet model, these phonocardiogram signals are converted into an image, i.e. 2-D spectrogram. Spectrograms are convenient for representing these heartbeat recordings because it capture the intensity of the frequencies through -out a given sound . Thus , these spectrograms are effective rep- representations of an audio recording. In this work , I use MFCC and HCQT based spectrograms for phonocardiogram signal classification . In this work , i take stethoscope sounds and even waveforms recorded using the microphone of a mobile phone as input and apply deep learning to the task of automated cardiac auscultation, i.e. recognizing abnormalities in heart sounds. It describe an automated heart sound classification algorithm that combines the use of time-frequency heat map representations with a deep convolutional neural network (CNN).

The original one-dimensional time series data is transformed into a two -dimensional time-frequency representation (i.e. spectrogram) , which allows each heart sound segment to be processed as an image. The Convolutional Neural Network (CNN) is one of the neural network architecture specifically used for image classification. Just like other neural network methods, CNN is also inspired by human brain tissue. Convolution neural network is mainly composed of two parts, feature extraction, and classification. The network architecture of a convolutional neural network that accepts as input a single channel 40x130 MFCC heat map and outputs a binary classification, predicting whether the input segment represents a normal or abnormal heart sound.

Convolutional Neural Network in this study uses 5 convolution layers, 4 max-pooling layers, 4 dropout layers, 1 global average pooling layer and finally a dense layer. The activation function in convolution layers uses Rectifier Linear Unit (ReLU) algorithm. The ReLU algorithm has advantages in time efficiency for training and testing.

The dropout layer : The term "dropout" refers to dropping out units (both hidden and visible) in a neural network. It is a very efficient way of performing model averaging with neural networks. Model averaging is a natural response to model uncertainty. The dropout layer allows for regularization by randomly setting some neurons in previous layers to zero during training.

Max Pooling : The objective of Max pooling is to down-sample an input representation. It helps in reducing the dimensionality and alleviate feature extraction. It reduces the computational cost-reducing the number of parameters to learned.

Dense layer : Here every input is connected to every output by weight. And using softmax as the non-linear activation function after this layer.

Adam method is used for the optimization process to update the weight on the Convolutional Neural Network. This method has efficient computation (memory and time), invariant to gradient scaling and suitable when applied to large data or parameters.

E. PHONOCARDIOGRAM SIGNAL DATABASE

Here used freely available open access dataset on Kaggle [33], originating through the PASCAL heart sounds classification challenge. Two datasets named A & B were generated through the PASCAL heart sound classification challenge [16]. Dataset A contains the variable-length (varying from 1 to 30 seconds) sounds recorded through a digital stethoscope in a real-time situation having background noise. Dataset A was partitioned into four classes named normal, extra heart sound, murmur, and artifact, while dataset B was partitioned into three classes: normal, extra-systole, and murmur. Here it have merged both datasets into a single dataset consisting of all five classes in this work.

This dataset was originally used in a Machine Learning challenge for the classification of heartbeat sounds by Mr Peter Bentley [2]. The dataset is divided into 2 sets depending on the sources from where

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

it was collected. Set A (set_a.csv) data was collected from the general public via the iStethoscope Pro & iPhone app and Set B (set_b.csv) from a clinical trial in hospitals using the digital stethoscope DigiScope. In this dataset there are 5 classes of heartbeat sounds:

1. Normal: healthy heart sounds
2. Murmur: extra sounds that occur when there is turbulence in blood flow hat causes the extra vibrations that can be heard
3. Extrahls: heartbeats with an additional sound
4. Extrasystoles: are additional heartbeats that occur outside the physiological heart rhythm and can cause unpleasant symptoms .

Artifact: disturbances in rhythm monitoring caused by movement of the electrodes.

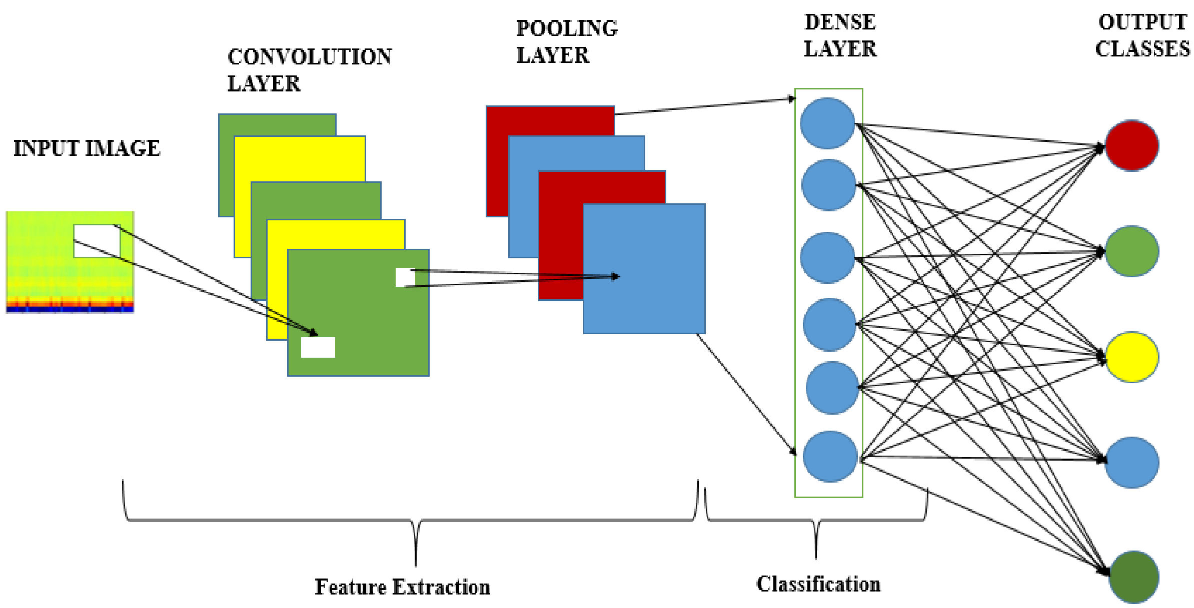
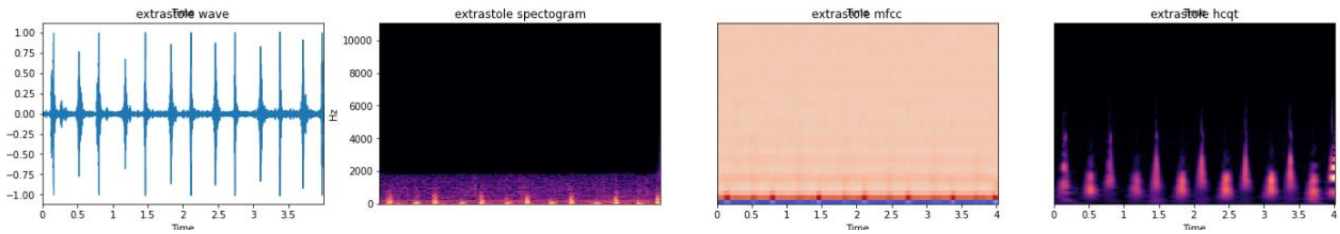


FIGURE 4 : The Archietutre Of Cnn & Spectogram Based Phonocardiogram Signal Classification Model . Input Are The Spectogram Generated Through MFCC , CQT , VQT , HCQT And Output Is One Of The Five Classes : Artificat , Murmur ,Extra Systole ,Extrahls , Normal .

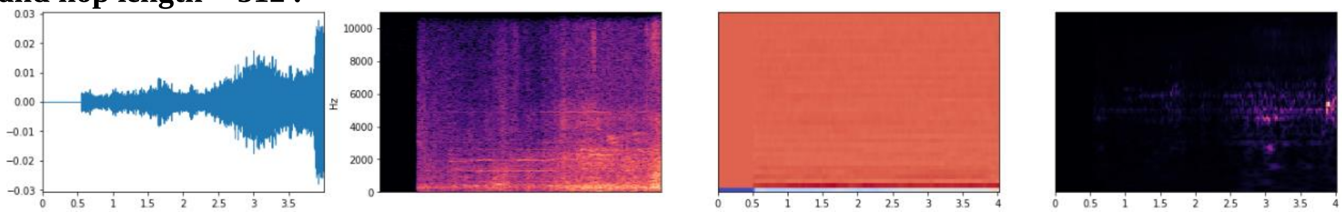
The number of phonocardiogram signals in normal, murmur, artifact, extra-systole, and extrahls classes are 255, 114, 40, 37, and 16. Since the number of heartbeat signals in each class is very low, audio augmentation is performed over raw audio signals. It have applied noise injection, shifting time, varying pitch, and speed to generate augmented data for phonocardiogram signals. After audio augmentation, the number of phonocardiogram signals in normal, murmur, artifact, extra-systole, and extrahls classes are 2555, 1146, 400, 378, and 158, respectively. The augmented dataset is partitioned into training and testing datasets with an 80:20 ratio. A spectrogram represents the PCG signal waves, as shown in Fig. (5-9), that presents five types of HCQT spectrograms for the artifact, extrahls, extra-systole, murmur, and normal in that order. Red shades described the amplitude of a PCG signal in a spectrogram. The spectrogram of a normal PCG signal is a strong sequence of amplitude, i.e., lub dub. It displays a noise sequence of amplitude in the murmur PCG signal greater than normal and extra-systole PCG signals. The amplitude of a PCG signal is greater than the normal PCG signal but lesser than the murmur PCG signal in the extrasystole PCG signal.

Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621



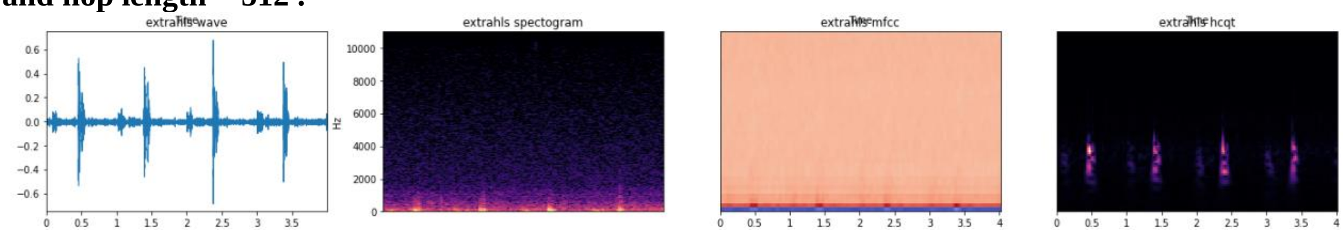
(a) (b) (c) (d)

FIGURE 5 : (a): A sample waveform for extrahls phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT power spectrogram for extrahls phonocardiogram signal . the spectrogram is generated with the number of frequency bins = 84 and hop length = 512 .



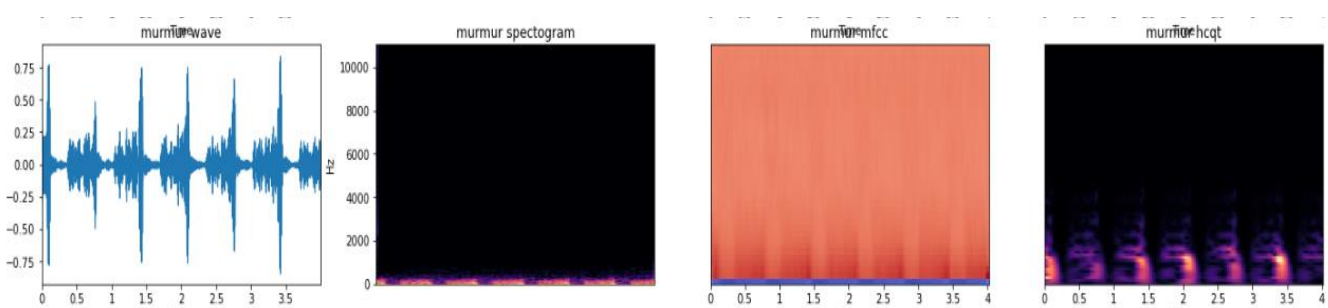
(a) (b) (c) (d)

FIGURE 6 : (a): A sample waveform for Artifact phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT power spectrogram for artificat phonocardiogram signal . the spectrogram is generated with the number of frequency bins = 84 and hop length = 512 .



(a) (b) (c) (d)

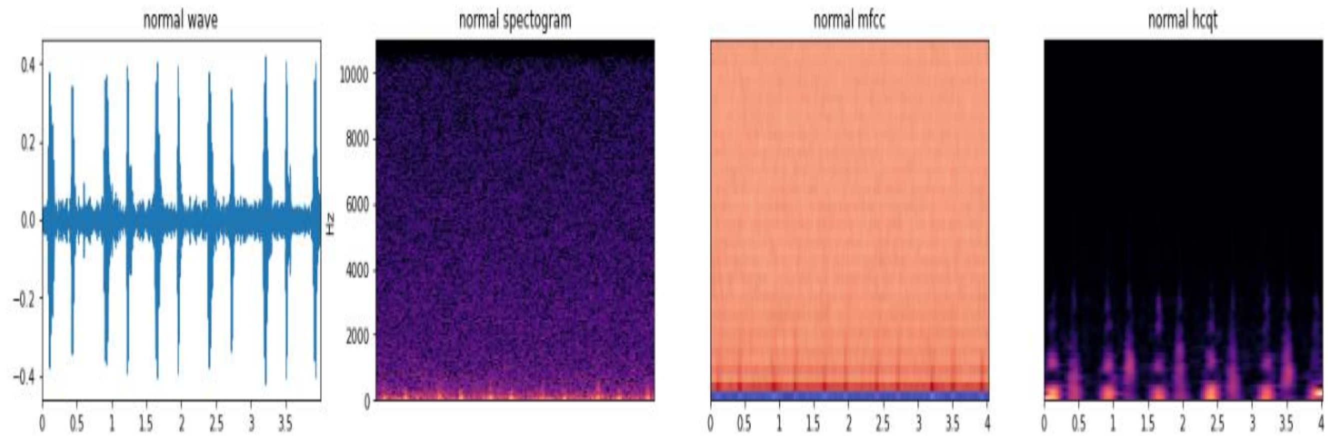
FIGURE 7 : (a): A sample waveform for Extra systole phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT power spectrogram for for Extra systole phonocardiogram signal . the spectrogram is generated with the number of frequency bins = 84 and hop length = 512 .



(a) (b) (c) (d)

FIGURE 8 : (a): A sample waveform for Murmur phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT power spectrogram for for Murmur phonocardiogram signal . the spectrogram is generated with the number of frequency bins = 84 and hop length = 512 .

Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621



(a) (b) (c) (d)

FIGURE 9 : (a): A sample waveform for Normal phonocardiogram signal, (b-d): heat map visualization for SPECTOGRAM , MFCC, and HCQT power spectrogram for for Normal phonocardiogram signal . the spectrogram is generated with the number of frequency bins = 84 and hop length = 512 .

IV. EXPERIMENT & RESULTS

Four separate ConvNet models termed ConvNet-MFCC, ConvNet-CQT, ConvNet-VQT, and ConvNet-HCQT are designed with MFCC, CQT, VQT, and HCQT spectrograms, respectively. To build the proposed ConvNet models, Keras, an open-source Python library, has been used that can run on top of different machine learning libraries like TensorFlow. In addition, the Librosa library in Python is used for generating MFCC, CQT, VQT, and HCQT spectrograms.

ConvNet models used in this phonocardiogram signal classification model using these spectrograms have four convolutional layers. The first convolution layer has a size of 32-5×5, the second convolution layer has a size of 64-5 × 5, the third convolution layer has a size of 64-5 × 5, and the last layer has a size of 32-5 × 5. A subsampling layer using max-pooling follows the first two convolution layers. The size of these max-pooling layers is 2 × 2 with a stride of size 2 × 2. The final layer of the ConvNet model is a fully connected layer with a softmax non-linear activation function with five units. These five units in the last layer are essential for this five-class phonocardiogram signal classification problem. Additionally, two dropout layers are also used to avoid overfitting with a 0.4 drop rate. The size of the MFCC spectrogram images is 128 × 130. The model is compiled after design. The optimizer is the gradient descent algorithm based on 'Adam' optimizer and cross-entropy loss to calculate the prediction error rate. The values 0.0001 are used as the learning rate. This optimizer uses backpropagation to update the weights of the neurons. It computes the derivative of the loss function regarding each weight and deducts it from the weight. A categorical cross-entropy loss function is utilized due to the multi-class nature of the problem, which has the form given by (7):

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N} log \frac{e^{W_{y_i}^T x_{i}+b_{y_i}}}{\sum_{j=1}^{n} e^{W_j^T x_i + b_j}} \quad (7)$$

W = weight matrix, xi = i th training sample, yi = class label for the ith training sample, b = bias term, N = sample count, Wj, and Wyi are the jth and yth i column of W. 300 epochs with batch size 128 are used for training. Fig. 10 , shows the accuracy and loss curves for the train and test set during the training of ConvNet models. The shape and dynamics of these learning curves are studied to diagnose the behavior of a ConvNet model. Three common dynamics observed in these learning curves are under-fitting, overfitting, and optimal fitting. From these plots, it can be verified that the ConvNet-HCQT model has offered optimal fit in comparison to other models.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

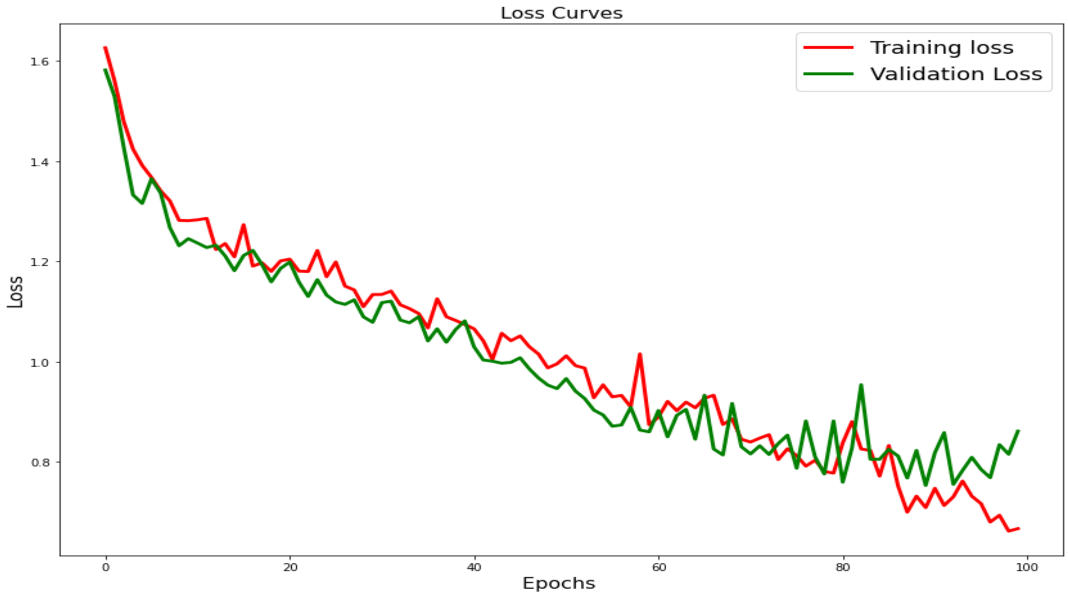


FIGURE 10 : Evolution of classification loss with training and validation image datasets throughout the training of ConvNet-HCQT model. Loss decreases abruptly for the first 200 repetitions and becomes stable after 250 repetitions.

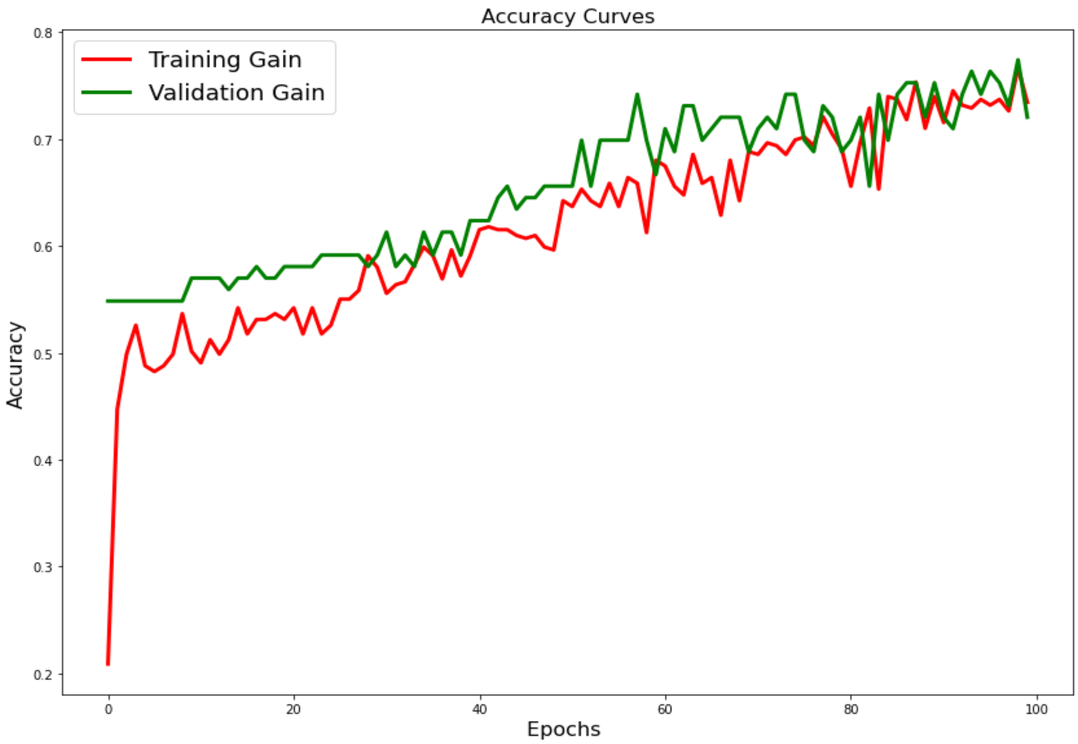


FIGURE 11 : Evolution of classification gain with training and validation image datasets throughout the training of ConvNet-HCQT model. Loss decreases abruptly for the first 200 repetitions and becomes stable after 250 repetitions.

V. RESULT ANALYSIS AND DISCUSSION

Commonly used time-frequency transformations and features such as DFT, DWT, and MFCC have extensively supported various acoustic recognition systems. Though the appreciated for most acoustic analyses, it is still not customized to any particular problem. So, it may be valuable to investigate features from other time-frequency transformations such as CQT, VQT, and HCQT. CQT is a dominant feature in acoustic signal processing analysis. CQT transforms a series of time-domain signals to the frequency domain signal. It is similar to the Short Term Fourier Transform (STFT) and almost identical to the complex Morlet wavelet transform. Hybrid CQT is a more computationally efficient version of CQT. It utilizes the pseudo-CQT for higher-order frequencies where the hop length is larger than half the filter size and full CQT for the lower frequencies. The findings of the experiments show that HCQT is more effective than traditional CQT and variable CQT. In this study, an effort is made to suggest the best acoustic features for phonocardiogram signal classification.

Django is a Python-based web framework, free and open-source, that follows the model–template–views architectural pattern. It is maintained by the Django Software Foundation, an independent organization established in the US as a 501 non-profit. Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing app without needing to reinvent the wheel.

PyCharm is an integrated development environment used in computer programming, specifically for the Python programming language. It is developed by the Czech company JetBrains. PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.

VI. CONCLUSION

While using the web application is can be used from anywhere at any time. A Web application (Web app) is an application program that is stored on a remote server and delivered over the Internet through a browser interface. Web applications include online forms, shopping carts, word processors, spreadsheets, video and photo editing, file conversion, file scanning, and email programs such as Gmail, Yahoo and AOL.

This study was designed to classify the heartbeat sound into 5 different classes. The raw data was collected using Stethoscopes and heartbeats recorded through the microphone of a mobile phone. Classification of heartbeat sounds was conducted using a Convolutional Neural Network. It did not use any other time sequence based Neural Networks such as RNNs since the temporal behaviour of the heartbeat was repeated within the window of observation and different sequential patterns were not needed to be learnt. Diagnose at an early stage is the only way to decrease the mortality rate occurring due to CVD. However, due to a lack of awareness for routine health checkups and unavailability of all resources at low cost, there are major hurdles in the early diagnosis of CVD. The situation worsens in developing countries where population density is high, and a doctor is not available in remote locations. To target these issues, it have offered a design of a decision support system that utilizes the PCG signals for the early diagnosis of CVD. PCG signals can be captured by a small, low-cost handheld device called a stethoscope .

The work presented here is one of the first to apply deep convolutional neural networks to the task of automated heart sound classification of heartbeat sound recorded through a stethoscope. The proposed system developed a novel algorithm first transforms the one-dimensional time-series input into a two-dimensional time-frequency Mel spectrogram. It then trains a 5-layer CNN architecture on the MFCC obtained from the Mel spectrogram and Hybrid_CQT. The trained network automatically classify the heart beat sound into 5 classes. The epoch values used were 100,200,300. The best results were obtained with 300 epoch at 0.001 learning rate applied on batch size of 128. The training accuracy is 78.73, while the testing accuracy rate is 75%.

DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

REFERENCES

[1] (Jun. 2021). CVD Data as Cited on 27th. [Online] Available : https://www.who.int/news-room/fact-sheets/detail/cardiovasculardiseases-(cvds).

[2] A. Jain, S. Tiwari, and V. Sapra, ‘‘Two-phase heart disease diagnosis system using deep learning,’’ Int. J. Control Autom., vol. 12, no. 5, pp. 558–573, 2019.

[3] A. K. Dwivedi, S. A. Imtiaz, and E. Rodriguez-Villegas, ‘‘Algorithms for automatic analysis and classification of heart sounds—A systematic review,’’ IEEE Access, vol. 7, pp. 8316–8345, 2019.

[4] P. Rani, R. Kumar, N. M. S. Ahmed, and A. Jain, ‘‘A decision support system for heart disease prediction based upon machine learning,’’ J. Reliable Intell. Environ., pp. 1–13, Jan. 2021, doi: 10.1007/s40860-021- 00133-6.

[5] A. E. F. Malik, S. Barin, and M. E. Yuksel, ‘‘Accurate classification of heart sound signals for cardiovascular disease diagnosis by wavelet analysis and convolutional neural network: Preliminary results,’’ in Proc. 28th Signal Process. Commun. Appl. Conf. (SIU), Oct. 2020, pp. 1–4.

[6] G. Y. Son and S. Kwon, ‘‘Classification of heart sound signal using multiple features,’’ Appl. Sci., vol. 8, no. 12, pp. 2344–2358, 2018.

[7] P. Upretee and M. E. Yuksel, ‘‘Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results,’’ in Proc. Sci. Meeting Elect.-Electron. Biomed. Eng. Comput. Sci. (EBBT), Apr. 2019, pp. 1–4.

[8] S. Patidar, R. B. Pachori, and N. Garg, ‘‘Automatic diagnosis of septal defects based on tunable-Q wavelet transform of cardiac sound signals,’’ Expert Syst. Appl., vol. 42, no. 7, pp. 3315–3326, 2015.

[9] C. Cable. (1997). The Aucultation Assistant. Accessed: Apr. 3, 2014. [Online]. Available: www.med.ucla.edu/wilkes/intro.html

[10] Stillman. (2007). ALDMD Clinical Cardiology Tools from Hennepin County Medical Center. Accessed: Apr. 8, 2014. [Online]. Available: http://www.aldmd.com

[11] R. Lichtenberg. (May 1, 2014). Heart Sounds. [Online]. Available: http://www.loyolauniversity.adam.com

[12] (Mar. 16, 2014). University of Michigan Heart Sound and Murmur Library. [Online]. Available: http://www.med.umich.edu

[13] J. M. Wilson. (2009). Heart Sound Podcast Series. Accessed: Apr. 6, 2014. [Online]. Available: www.texasheartinstitute.org

[14] A. Gharehbaghi, A. A. Sepehri, A. Kocharian, and M. Linden, ‘‘An intelligent method for discrimination between aortic and pulmonary stenosis using phonocardiogram,’’ in Proc. World Congr. Med. Phys. Biomed. Eng., Toronto, ON, Canada, 2015, pp. 1010–1013.

[15] O. Deperlioglu, U. Kose, D. Gupta, A. Khanna, and A. K. Sangaiah, ‘‘Diagnosis of heart diseases by a secure Internet of Health Things system based on autoencoder deep neural network,’’ Comput. Commun., vol. 162, pp. 31–50, Oct. 2020.

[16] P. Bentley, G. Nordehn, M. Coimbra, S. Mannor, and R. Getz. Classifying Heart Sounds Challenge. Accessed: May 3, 2018. [Online]. Available: http://www.peterjbentley.com/heartchallenge/#downloads

[17] PhysioNet/Computing in Cardiology Challenge. Classification of Normal/Abnormal Heart Sound Recordings. Accessed: May 31, 2018[. Online]. Available: https://www.physionet.org/challenge/2016/

[18] M. Banerjee and S. Majhi, ‘‘Multi-class heart sounds classification using 2D-convolutional neural network,’’ in Proc. 5th Int. Conf. Comput., Commun. Secur. (ICCCS), Oct. 2020, pp. 1–6.

[19] S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, and A. Gumaei, ‘‘CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings,’’ 2020, arXiv:2010.01392. [Online]. Available: http://arxiv.org/abs/2010.01392

[20] G. Redlarski, D. Gradolewski, and A. Palkowski, ‘‘A system for heart sounds classification,’’ PLoS ONE, vol. 9, no. 11, pp. 1–12, 2014.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.124 ISSN: 2581-4621

[21] *Valueva, M.V.; Nagornov, N.N.; Lyakhov, P.A.; Valuev, G.V.; Chervyakov, N.I. (2020). "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation". Mathematics and Computers in Simulation. Elsevier BV.* ***177****:232243. doi:10.1016/j.matcom.2020.04.031. ISSN 03784754. S2CID 218955622. Convolutional neural networks are a promising tool for solving the problem of pattern recognition.*