International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Improved Deep Learning and Feature Fusion Techniques for Chronic Heart Failure

Namrata Gawande[1], Dinesh Goyal[2], Kriti Sankhla[3]

Submitted: 20/12/2023 **Revised**: 31/01/2024 **Accepted**: 10/02/2024

Abstract: Early detection of heart problems is of paramount importance, given that chronic heart failure remains a leading cause of global mortality. Accurate forecasting of cardiac conditions is crucial for timely intervention and improved patient outcomes. While various machine learning (ML) and deep learning (DL) models have emerged for cardiac disease diagnosis, most struggle to effectively handle high-dimensional healthcare datasets and often fail to significantly enhance chronic heart failure (CHF) diagnosis performance. In this study, we propose a smart healthcare framework that integrates deep learning and feature fusion techniques to predict CHF. Leveraging the PhysioNet datasets, our approach amalgamates features extracted from phonocardiogram (PCG) data. The study introduces novel algorithms, including lightweight CNN, hybrid CNN-autoencoder, and parallel hybrid CNN-autoencoder, offering promising avenues for enhancing CHF detection accuracy and efficiency. The performance of our proposed system is rigorously evaluated against alternative approaches, including feature extraction, machine learning, and traditional deep learning classifiers, using heart sound data. This research aims to advance CHF prediction capabilities, bridging the gap between cutting-edge technology and early cardiac healthcare intervention.

Keywords: *Deep Learning, Machine Learning, Feature Extraction, PCG, Heart sound Classification, Healthcare*

1 Introduction

The heart is an extremely important organ in the human body, and cardiovascular disease (also known as CVD) is one of the main causes of illness and death on a global scale. Some signs that pertain to the heart include computer signals of the heart, sometimes known as electrocardiograms (ECGs), heart sound signals (referred to as PCGs), and so on, indicates that something is amiss with the circulatory system. When compared to electrocardiogram (ECG) readings, Changes in the status of the heart is clearly seen on a phonocardiogram (PCG), which is a visual record of the sounds the heart makes. As a result, PCG may be used to determine if heart components are out of shape or whether heart valves are damaged [2]. When someone has a myocardial attack, a greater decrease in cardiac reserve (CR) makes the heart less tight. This is more important than a lower heart's performance at rest. The main goal is to find out more about CHF diagnosis signs based on Heart Sound features. The primary method of auscultation is cardiovascular relatively simple diagnostic technique, used to assess and identify the heart's functionality and quality of proper functioning. Auscultation is a form of physical therapy in which heartbeats heard on the lateral and anterior chest walls, typically using a stethoscope [25]. Auscultation does have a few limitations, though. For instance, it solely depends on hearing ability and experience of the doctor. These factors highlight the requirement for systems-systems that can accurately process and identify cardiac sounds.

Heart sounds is affected in a variety of ways by disorders that affect the cardiovascular system. Timing, amplitudes, spectrum contents, and other aspects of a signal pattern are all areas in which a healthy signal pattern and an unhealthy signal pattern can be distinguished from one another.

This research paper presents the following significant work:

- Use Publicly Available PhysioNet Dataset as input and apply Feature Extraction and Feature Selection Techniques.
- To work on different classification algorithms, including Various ML and DL Techniques.
- To analyze the current model for Heart Failure Prediction using Heart Sound data and improve it with advance deep learning techniques.

2 Review of Literature

The objective of the study that Li, Liu, Zhao, and the other authors [1] set out to achieve was to maximise the effectiveness of one-dimensional convolutional neural networks (CNNs) in the field of heart sound analysis. Their primary goal was to improve the classification of heart sound abnormalities by leveraging the inherent

[1]*Department of Computer Science & Engineering, Poornima University, Jaipur, gawande.namrata@gmail.com*
[2]*Department of Computer Science & Engineering, Poornima Institute of Engineering & Technology, Jaipur, dinesh8dg@gmail.com*
[3]*Department of Computer Science & Engineering, Poornima University, Jaipur, kriti.sankhla@poornima.edu.in*

ability of CNNs to automatically extract pertinent features from complex and non-stationary heart sound data. Similarly, Krishnan Palani et al. in [19] also delved into unsegmented PCG signal categorization, employing 1D-CNNs and feed-forward neural networks (F-NN). Their project aimed to streamline the feature engineering and feature selection processes integral to PCG signal analysis.

Various feature extraction techniques for different medical applications are describe by several researchers. Yang et al. [11] employ a novel fuzzy matching method utilizing Gaussian wavelet and Convolutional features in the time and frequency domains to diagnose heart problems using PCG signals, subsequently applying ML algorithms like SVM, KNN, MLP and RF to assess their feature extraction method's effectiveness. Li, Fan et al. [22] propose a deep learning-based data pre-processing technique for effectively diagnosing healthy and abnormal cardiac recordings, employing a global average pooling layer within a CNN to comprehensively analyze features from eight domains. Pedro Narváez et al. [24] recognise systolic and diastolic intervals in cardiac sound recordings using the modified empirical wavelet transform (EWT) and normalised Shannon average energy, extracting power parameters to optimize computational efficiency for potential real-time applications.

In the field of medical diagnostics, several approaches have been explored to enhance the accuracy and efficiency of heart sound analysis. Alkhodari et al. [12] concentrated on early VHD diagnosis using PCG heart sounds and novel approaches including the maximum overlap discrete wavelet transform (MODWT) and -score normalisation. Their NN architecture, a hybrid of CNN with Bi-directional LSTM, was designed to make complicated studies easier to understand. In a similar vein, Al-Issa, Y. et al. [14] introduced a system for heart problem diagnosis that effectively distinguished between various heart conditions using CNN and LSTM, showcasing excellent results. On the data side, C. Liu et al. [15] helped by creating an open-access database of heart sounds, which made it possible to test heart sound algorithms. Furthermore, Fahime Khozeimeh et al. [17] introduced the CNN-AE method for predicting the survival chances of COVID-19 patients, emphasizing the importance of data augmentation. Yadav Anjali et al. [20] strategically processed heart sounds to extract discriminative features for ML-based classification of heart disorders. Lastly, Humayun, Ahmed Imtaaz, and collaborators [21] proposed a novel CNN layer with moment components for identifying ECG signal abnormalities, surpassing previous systems in public cross-sample datasets. These studies collectively demonstrate the diverse approaches and data resources used in the advancement of cardiac diagnostics, offering both advantages and challenges for improving healthcare outcomes.

In recent research, several papers have focused on lightweight models for various medical diagnostic applications. CardioXNet, a new lightweight CRNN architecture for the automated diagnosis of cardiac auscultation classes utilising raw PCG data, was presented by Shuvo, Samiul Based, et al. [13]. Their method had two learning steps: representation learning and sequence residual learning. During these steps, coarse and fine-grained features from PCG data were taken from parallel CNN paths. Haval I. et al. [16] created two lightweight deep learning models for early COVID-19 detection using chest X-ray images, obtaining outstanding accuracy rates in binary and multiclass classification tests. In contrast, Md Nahiduzzaman et al. [18] presented a framework for their diagnostic application. This framework combines the powerful feature extraction techniques and lightweight parallel CNN model. This framework is referred to as CNN-ELM for short. These studies demonstrate the rising interest in lightweight models for efficient and precise medical diagnostics.

3 Proposed Work

In our proposed methodology heart failure detection using the Physionet dataset is performed. We have devised a comprehensive approach encompassing several crucial steps. First, we employ advanced feature extraction techniques to distill meaningful information from the physiological data, ensuring that essential patterns and attributes related to heart health are captured effectively. To address temporal aspects of the data, we implement padding sequences, allowing us to maintain the temporal context and integrity of the signals. Subsequently, we perform a rigorous train-test split to ensure model evaluation's reliability and prevent data leakage. For the heart of our methodology, we leverage a combination of traditional ML models and state-of-the-art deep learning techniques [23], aiming to capture both linear and complex nonlinear relationships within the data. We have also introduced novel improvements in the deep learning architecture, including lightweight models and hybrid model to enhance model performance and interpretability. This holistic methodology represents a robust framework for accurate and reliable heart failure detection, with the potential to significantly impact clinical decision support systems and patient care. Figure 1. illustrates the sequence of steps involved in our heart failure detection methodology using the Physionet dataset.

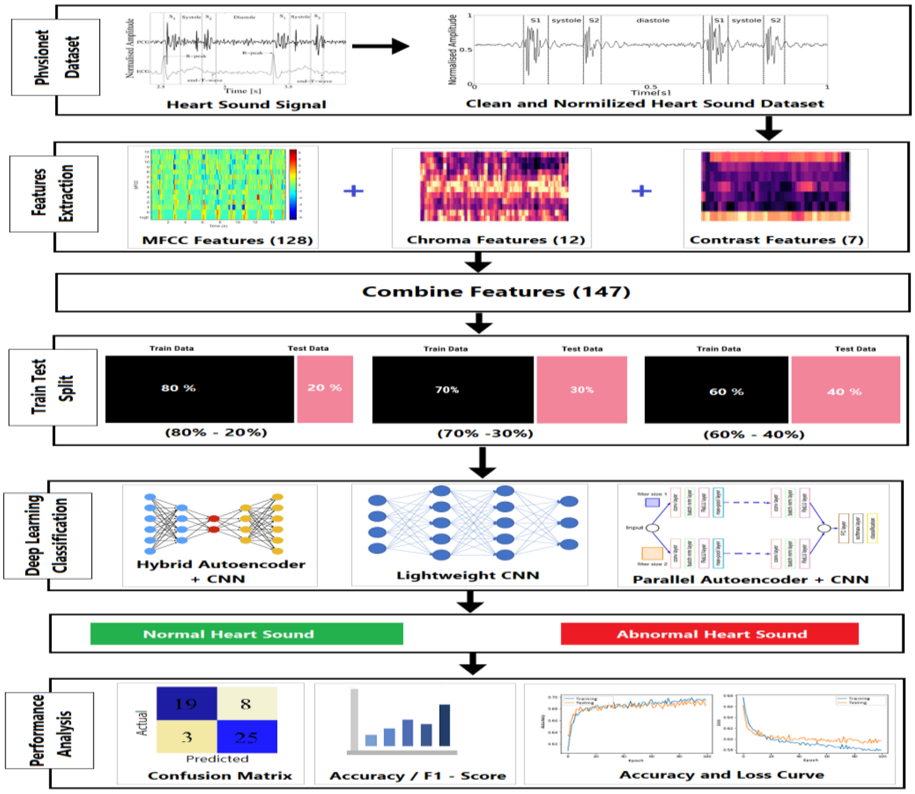

Fig. 1. System Architecture

3.1 Input Dataset: There are 3,240 heart sound recordings in the 2016 PhysioNet Challenge data spanning from 5 seconds to over 2 minutes in length. These recordings were collected from a wide variety of people, both well and ill, in a variety of situations outside of clinical settings. Although the dataset has many cardiac sound localizations, few record specifics are provided. Sampled at 2000 Hz, each WAV file has a single PCG lead. However, some recordings include noise interference from the recording settings, making it hard to tell the normal or abnormal.

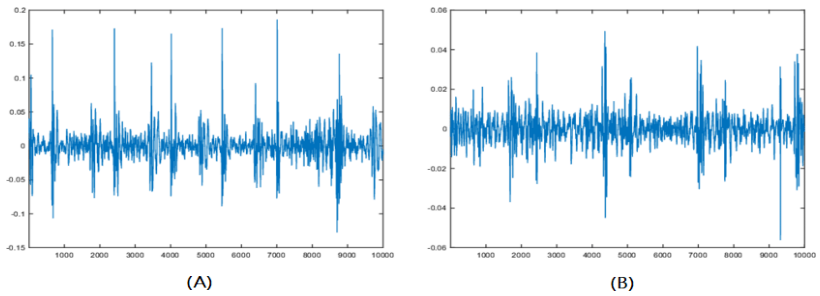

Fig. 2. (A) Normal (B) Abnormal Heart Sound Signal

3.2 Feature Extraction: Input dataset contain audio signal from which relevant features are extracted in time, frequency and time-frequency domain. Feature extraction techniques such as MFCCs, spectral contrast on the signals and chroma are used. By extracting these features from the audio signals, the important characteristics related to the audio content, such as spectral shape, pitch content, and timbral properties are extracted.

Mel-frequency cepstral coefficients (MFCCs): MFCC models the human auditory system's perception of sound to capture the spectrum features of an audio source. The MFCCs are computed by performing a number of mathematical operations on the audio stream. In these procedures, the signal is first segmented into a series of brief frames that overlap one another, and then the Fourier transform is performed on each frame individually in order to acquire the power spectrum, applying a filter to the power spectrum to extract relevant frequency bands, and finally taking the logarithm of the filterbank energies [4]. The resulting MFCCs provide a compact representation of the audio signal, capturing information about its spectral content and timbral characteristics. Mel scale [4] calculated as:

$$mel(f) = 2595\text{Log}_{10}(1 + \frac{f}{700}) \quad (1)$$

In heart sound classification, MFCCs were used to describe what each type of heart sound sounds like. Heart sounds with different rhythms in the MFCC traits could be caused by different problems with the heart. By training a model to recognise these trends, it is possible to use the heart sound to figure out what's wrong with the heart.

Chroma features: The pitch classes or musical notes in an audio source are shown by the chroma features. They tell us about the signal's musical content and how the tones fit together. The chroma features are found by matching the power spectrum of the audio input to the 12 pitch classes of the equal-tempered scale. For this mapping, you have to add up the energy in each frequency band that corresponds to a pitch class.

- librosa.feature.chroma_stft(S=stft, sr=sample_rate) uses a short-time Fourier transform (STFT) to figure out the chroma feature of an audio signal. Most of the time, chroma features are used to find information about music. They are basically a picture of the harmonic structure of music. Each frame of the audio information is used to figure out the chroma feature.
- `T` transposes the resulting matrix, so that the chroma features become the columns of the matrix.
- `np.mean (..., axis=0)` then computes the mean of these features over time. The `axis=0` argument specifies that the mean should be computed over columns (which, after the transpose operation, correspond to different frames in time). This results in a single chroma feature vector that describes the entire audio signal.

Spectral contrast: To compute spectral contrast, the power spectrum of the audio signal is divided into multiple frequency bands, and the difference between the maximum magnitude within each band and the minimum magnitude of neighbouring frequency bands is calculated. This difference is then averaged over the frequency bands to obtain a spectral contrast feature vector. Spectral contrast provides information about the spectral shape and emphasis in different frequency regions of the signal.

The 'spectral contrast' function computes the spectral contrast of an audio stream. The amplitude difference between peaks and troughs in a sound spectrum is measured as spectral contrast. It is often utilised as a characteristic in music and voice analysis. Let's break down this function call:

- `S=stft`: This is the input to the function. The `stft` here refers to the short-time Fourier transform of the audio signal. If you don't provide this, `librosa` will compute it for you from the raw audio.
- `sr=sample_rate`: This is the sample rate of the audio signal. It's needed to correctly compute the frequencies for the spectral contrast.
- The `. T` following the function call is a transpose operation. This is because `librosa` returns features where the columns correspond to frames, but we're going to average over frames so we want them to correspond to rows.
- Finally, `np.mean(..., axis=0)` computes the mean spectral contrast over all frames (because `axis=0` corresponds to columns in the 2D array). So `contrast` will be a 1D array where each element is the average spectral contrast for a different frequency band.

By extracting these features (MFCCs, chroma, and spectral contrast) from the audio signals, the important characteristics related to the audio content, such as spectral shape, pitch content, and timbral properties are extracted. These features serve as input to the ML and DL classifiers, enabling them to learn patterns and make predictions based on the extracted information from the audio signals.

3.3 Splitting Train and Test Data: The data may be used for either training or testing purposes. 80/20 split, 70/30 split, 60/40 split, respectively. The ML models are learned using the training set, and then their performance is evaluated using the testing set.

3.4 Machine Learning Models: ML algorithms play a pivotal role in various applications, including heart sound classification, by leveraging their unique approaches to data analysis and pattern recognition. SVM are known for their ability to find optimal hyperplanes that maximize the margin between different classes, making them suitable for distinguishing heart sound categories based on acoustic features. In contrast, Random Forest [27] is an ensemble learning approach that uses the combined wisdom of many decision trees to improve forecast accuracy while decreasing the likelihood of overfitting. Here we have used nine ML algorithms [26] including SVM, Random Forest, Adaboost, Extra Tree Classifier, XGB, MLP, KNN, SGD, GBM, etc. for classification.

3.4 Deep Learning Models:

a. Proposed Lightweight CNN Model

A lightweight CNN model has been designed to have fewer parameters and lower computational complexity than more complicated CNN models. Here are some benefits of using a lightweight CNN model: Memory and storage efficiency, Faster inference, training efficiency. With the use of lightweight CNN models provides a trade-off between model complexity and performance, offering a balance between accuracy and computational resources. Figure 3 shows Lightweight CNN Model architecture.

b. Hybrid CNN – Autoencoder Model Evaluation

A Hybrid model that combines an Autoencoder and a CNN for a binary classification task. The Autoencoder is used as a pre-training step to learn a compressed representation of the input data. The encoded representation learned by the Autoencoder is then reshaped and passed as input to the CNN. The CNN [28] utilizes the learned features to perform classification. During training, the model contains a Dropout layer that randomly sets a percentage of the input units to 0. This reduces the model's dependence on particular input characteristics, which helps to avoid overfitting. A dense layer equipped with a softmax activation function makes up the Output Layer. This layer is responsible for generating predicted probabilities for the binary classification task. Figure 4 shows Hybrid CNN – Autoencoder Model architecture.

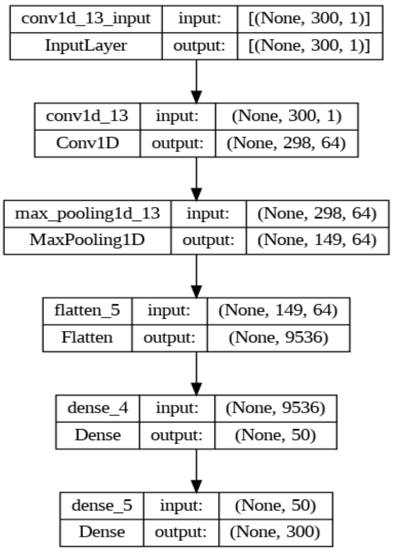

Fig. 3. Lightweight CNN model

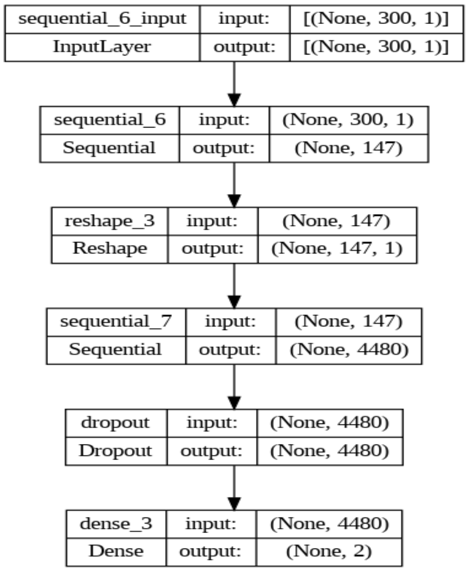

Fig. 4. Hybrid CNN – Auto encoder model

c. Parallel Hybrid CNN – Autoencoder Model

A parallel hybrid CNN-Autoencoder model combines the strengths of Convolutional Neural Networks (CNNs) and Autoencoders. By using parallel processing, the training process can be distributed across multiple processing units. The autoencoder and CNN branches can be trained simultaneously, reducing the overall training time. The autoencoder and CNN branches operate on different input streams, with each branch extracting distinct features. With parallel processing, the feature extraction process can be performed concurrently, allowing for efficient utilization of available resources.

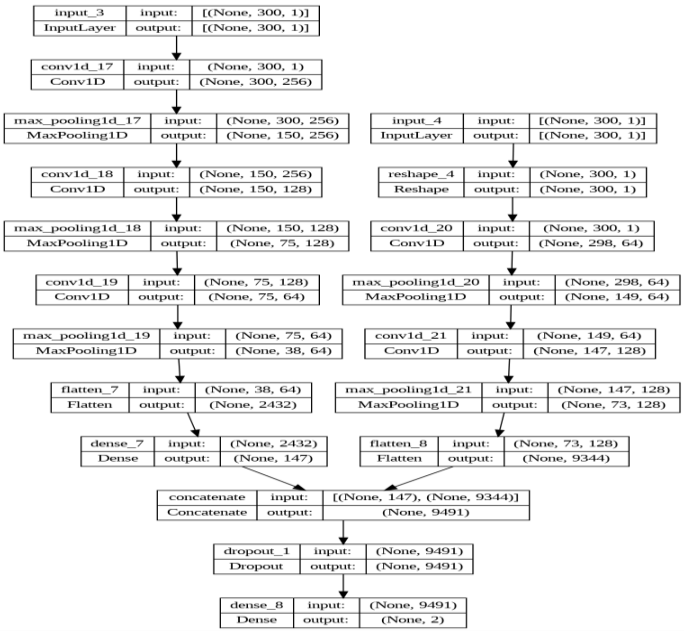

Fig. 5. Parallel Hybrid CNN – Auto encoder model

4 Result & Analysis

4.1 Dataset Description

Six files, ranging in duration from 5 seconds to over 2 minutes, make up the 2016 PhysioNet Cardiology Challenge [15] archive. There are six research organisations that send out 3,153 heart-sound videos. In both clinical and non-clinical situations, recordings were made from various locations on the bodies of both healthy and diseased individuals. It might be any of nine spots, but it's most likely the aorta, lungs, tricuspid, and mitral valve regions. There are both children and adults in the steady subject and the patient pathology samples. Each person may have made between one and six records. But the file doesn't give much information about who the record is for. Both samples were found to have a frequency of 2000 Hz and saved in the "Wav" format.

4.2 Performance Parameters

Accuracy, Specificity, Sensitivity, Precision, Recall, and F1-score are some of the metrics used to evaluate a classifier's efficacy [28].

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Recall \,/\, Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

$$F1 - Score = 2 * \frac{(Precision + Recall)}{(Precision * Recall)} \quad (6)$$

Receiver Operating Characteristic (ROC) Curve: As the discriminating threshold of a binary classification model changes, the ROC curve shows how the True Positive Rate (Recall) and the False Positive Rate change in conjunction.

Logarithmic Loss (Log Loss): Log Loss is a commonly used loss function to measure the performance of a classification model. It quantifies the difference between the true class probabilities and the predicted class probabilities

For a binary classification problem with two classes (0 and 1), the formula is:

$$\text{Log loss} = -\frac{1}{N}\sum_{k=0}^{n} yi \log\big(p(yi)\big) + (1 - yi).\log\big(1 - p(yi)\big) \quad (7)$$

Where N is the number of samples in the set. Yi is the real name for the i-th sample (0 or 1). Pi is the expected chance that the i-th sample is from class 1.

Matthews Correlation Coefficient (MCC)

The Matthews Correlation Coefficient is a metric for assessing the accuracy of binary classification predictions. TP, TN, FP, FN are all considered.

$$\text{MCC} = \frac{TP.TN - FN.FP}{\sqrt{(TP+FN)(TN+FP)(TP+FP)(TN+FN)}} \quad (8)$$

The MCC ranges from -1 (perfect inverse prediction) to +1 (perfect prediction) with 0 representing random or chance prediction. A higher MCC indicates a better classification performance.

The classifier's performance is tested on a number of different factors, and the best classification performance is used to sort heart beats.

4.3 Results

There are three possible training-test splits for the heart-sound dataset: 80%-20% and 70%-30%, 60%-40%. All databases were either used for training or testing purposes. Both the training and testing sets are independent populations (no data from the same subject/patient appears in both sets).

Table 1. Performance Analysis of Machine Learning models for Various Train Test Split Ratio

Model	Accuracy	Precision	Sensitivity	Specificity	F1 Score	ROC	Log Loss	Mathew corrcoef
Train Test Split (80% – 20%)								
*Performance Analysis of Machine Learning models w*ithout Feature Extraction								
RF	0.731707	0.864198	0.521739	0.043478	0.463095	0.521739	9.266716	0.177959
MLP	0.487805	0.487805	0.484893	0.478261	0.461875	0.484893	17.690710	-0.027146
KNN	0.719512	0.359756	0.500000	0.000000	0.418440	0.500000	9.687930	0.000000
ETC	0.719512	0.359756	0.500000	0.000000	0.418440	0.500000	9.687930	0.000000
XGB	0.682927	0.242424	0.316384	0.000000	0.274510	0.518055	10.109115	0.033896

SVM	0.682927	0.457143	0.487841	0.043478	0.440126	0.487841	10.951534	-0.045656
SGD	0.500000	0.516148	0.519897	0.565217	0.482690	0.519897	17.269486	0.035850
ADB	0.707317	0.625731	0.610906	0.391304	0.615925	0.610906	10.109047	0.236173
GBM	0.768293	0.760135	0.613486	0.260870	0.622120	0.613486	8.003053	0.343637
Performance Analysis of Machine Learning models with Feature Extraction								
RF	0.804878	0.790412	0.691968	0.434783	0.715278	0.691968	6.739400	0.472228
MLP	0.817073	0.775132	0.806559	0.782609	0.786569	0.806559	6.318118	0.580841
KNN	0.768293	0.711290	0.693073	0.521739	0.700557	0.693073	8.002994	0.403953
ETC	0.780488	0.736318	0.675018	0.434783	0.691729	0.675018	7.581809	0.406744
XGB	0.634146	0.266667	0.293785	0.000000	0.279570	0.658069	8.424219	0.175169
SVM	0.804878	0.758290	0.758290	0.652174	0.758290	0.758290	6.739351	0.516581
SGD	0.719512	0.359756	0.500000	0.000000	0.418440	0.500000	9.687930	0.000000
AdB	0.731707	0.672281	0.680914	0.565217	0.676006	0.680914	9.266598	0.353089
GBM	0.743902	0.676215	0.649595	0.434783	0.658537	0.649595	8.845423	0.324721
Train Test Split (70% – 30%)								
Performance Analysis of Machine Learning models without Feature Extraction								
RF	0.780488	0.766892	0.619697	0.272727	0.632836	0.619697	7.581839	0.357470
MLP	0.455285	0.487062	0.483838	0.545455	0.440492	0.483838	18.813903	0.028921
KNN	0.731707	0.365854	0.500000	0.000000	0.422535	0.500000	9.266716	0.000000
ETC	0.731707	0.365854	0.500000	0.000000	0.422535	0.500000	9.266716	0.000000
XGB	0.666667	0.255452	0.303704	0.000000	0.277496	0.576768	9.266664	0.101112
SVC	0.715447	0.534188	0.508081	0.060606	0.466741	0.508081	9.828309	0.033243
SGD	0.471545	0.503639	0.504545	0.575758	0.457193	0.504545	18.252290	0.008134
ADB	0.699187	0.603759	0.592929	0.363636	0.596721	0.592929	10.389850	0.196390
GBM	0.756098	0.777119	0.555051	0.121212	0.533148	0.555051	8.424280	0.247026
Performance Analysis of Machine Learning models with Feature Extraction								
RF	0.780488	0.720063	0.686869	0.484848	0.698994	0.686869	7.581793	0.405576
MLP	0.788618	0.740698	0.778788	0.757576	0.752477	0.778788	7.300932	0.518087
KNN	0.804878	0.751928	0.741919	0.606061	0.746566	0.741919	6.739358	0.493746
ETC	0.796748	0.747475	0.697980	0.484848	0.714564	0.697980	7.020187	0.442696
XGB	0.577236	0.278431	0.262963	0.000000	0.270476	0.682323	9.266592	0.174831
SVC	0.796748	0.742337	0.755556	0.666667	0.748219	0.755556	7.020148	0.497717
SGD	0.747967	0.690211	0.712626	0.636364	0.698076	0.712626	8.704973	0.402213
ADB	0.715447	0.635474	0.632828	0.454545	0.634084	0.632828	9.828224	0.268289

GBM	0.772358	0.708801	0.700505	0.545455	0.704327	0.700505	7.862583	0.409222

In the first part of experimentation, we implemented the different ML algorithms, RF, MLP, KNN, ETC, XGB, SVM, SGD, Adaboost, GBM without extracting the features. Table 1 shows the result with train test split 80% - 20% without feature extraction. Table 1 shows the result with train test split 80% - 20% with feature extraction. Here features extraction is performed by combining MFCC, Chroma and Contrast features total 147 features are extracted. These extracted features are fed to the machine learning algorithms. The best accuracy was 81% when the Multilayer perceptron (MLP) was used to classify the sound. Figure 6 and Figure 7 shows the accuracy comparison graph with and without feature extraction technique.

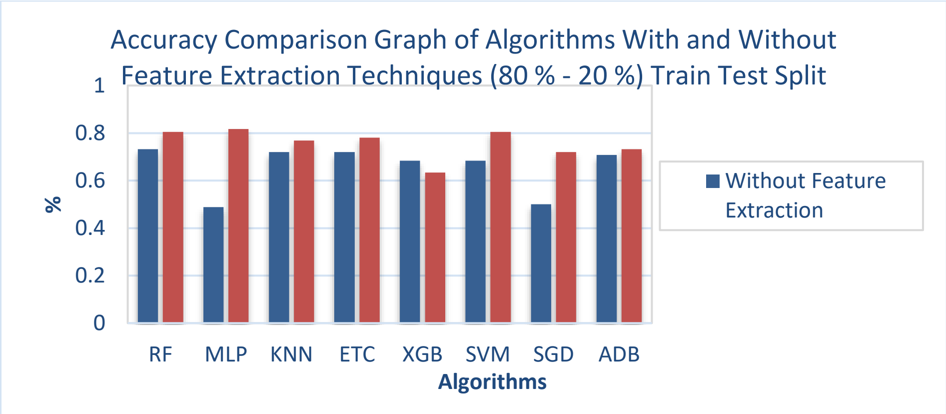

Fig. 6. Accuracy Comparison of Algorithms with and Without Feature Extraction Techniques (80 % - 20 %) Train Test Split

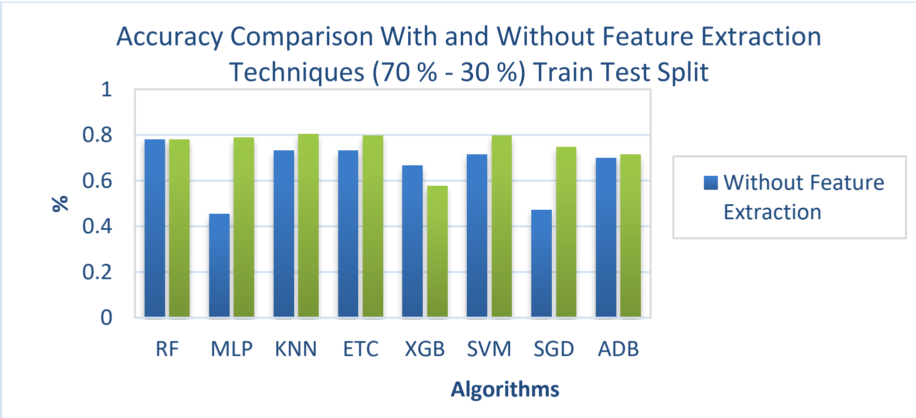

Fig.7. Accuracy Comparison of Algorithms with and Without Feature Extraction Techniques (70 % - 30 %) Train Test Split

Table 2 shows the performance comparison of proposed classification algorithms with exiting classification algorithms, the performance is compare with various validation split ratio, from table we can see that the Parallel HYB CNN + AE proposed model achieved the best accuracy of 86.0% for (80% - 20%) split ratio, 88.0 % for (70%-30%) split ratio and 84.0 % for (60%-40%) split ratio. The respective comparison graphs are shown in Figure 8, Figure 9 and Figure 10.

Table 2. Performance comparison table of proposed classification algorithms

Model	**ACC**	**PRE**	**Sensitivity**	**F1**	**ACC**	**PRE**	**Sensitivity**	**F1**	**ACC**	**PRE**	**Sensitivity**	**F1**
	(80 % - 20 %) Train Test Split				*(70 % - 30 %) Train Test Split*				*(60 % - 40 %) Train Test Split*			
RF	0.8	0.79	0.69	0.71	0.78	0.72	0.68	0.69	0.78	0.71	0.68	0.69

MLP	0.81	0.77	0.8	0.78	0.78	0.74	0.77	0.75	0.77	0.7	0.68	0.69
KNN	0.76	0.71	0.69	0.7	0.8	0.75	0.74	0.74	0.76	0.69	0.69	0.69
ETC	0.78	0.73	0.67	0.69	0.79	0.74	0.69	0.71	0.79	0.74	0.69	0.71
XGB	0.63	0.26	0.29	0.27	0.57	0.27	0.26	0.27	0.63	0.27	0.28	0.28
SVC	0.8	0.75	0.75	0.75	0.79	0.74	0.75	0.74	0.78	0.72	0.75	0.73
SGD	0.71	0.35	0.5	0.41	0.74	0.69	0.71	0.69	0.74	0.87	0.51	0.44
Adaboost	0.73	0.67	0.68	0.67	0.71	0.63	0.63	0.63	0.77	0.7	0.68	0.69
GBM	0.74	0.67	0.64	0.65	0.77	0.7	0.7	0.7	0.81	0.75	0.72	0.74
Lightweight CNN	0.82	0.83	0.82	0.82	0.84	0.85	0.84	0.84	0.82	0.84	0.81	0.82
HYB CNN + AE	0.84	0.84	0.83	0.83	0.87	0.86	0.84	0.86	0.84	0.84	0.83	0.84
Parallel HYB CNN + AE	0.86	0.86	0.84	0.8	0.88	0.89	0.87	0.85	0.84	0.85	0.85	0.83

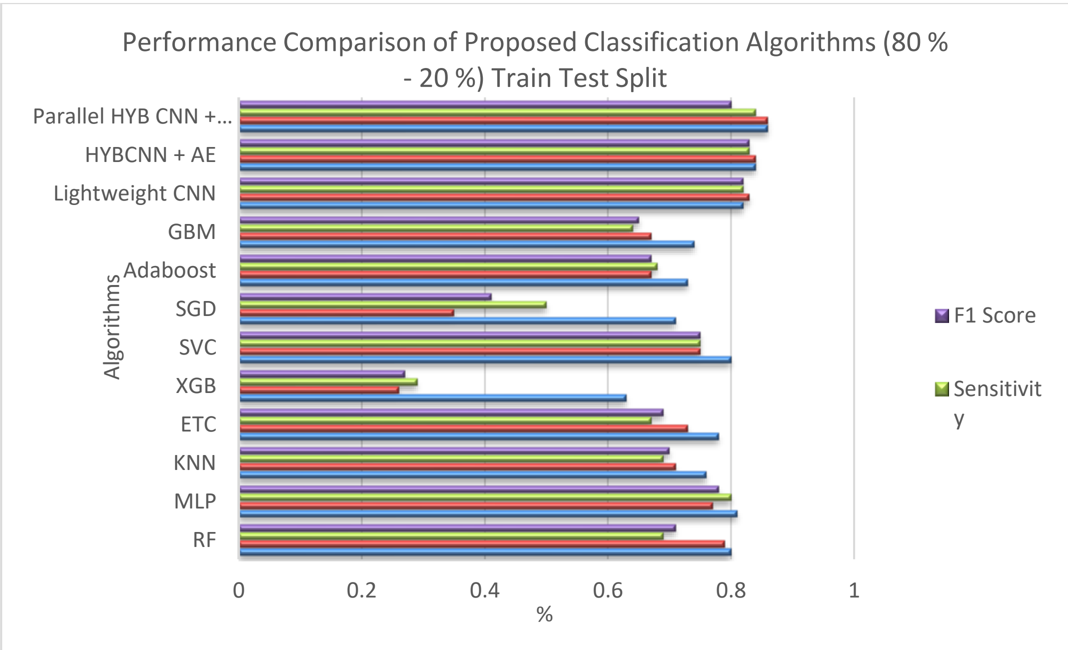

Fig. 8. Performance Comparison graph of Proposed Classification Algorithms (80 % - 20 %) Train Test Split

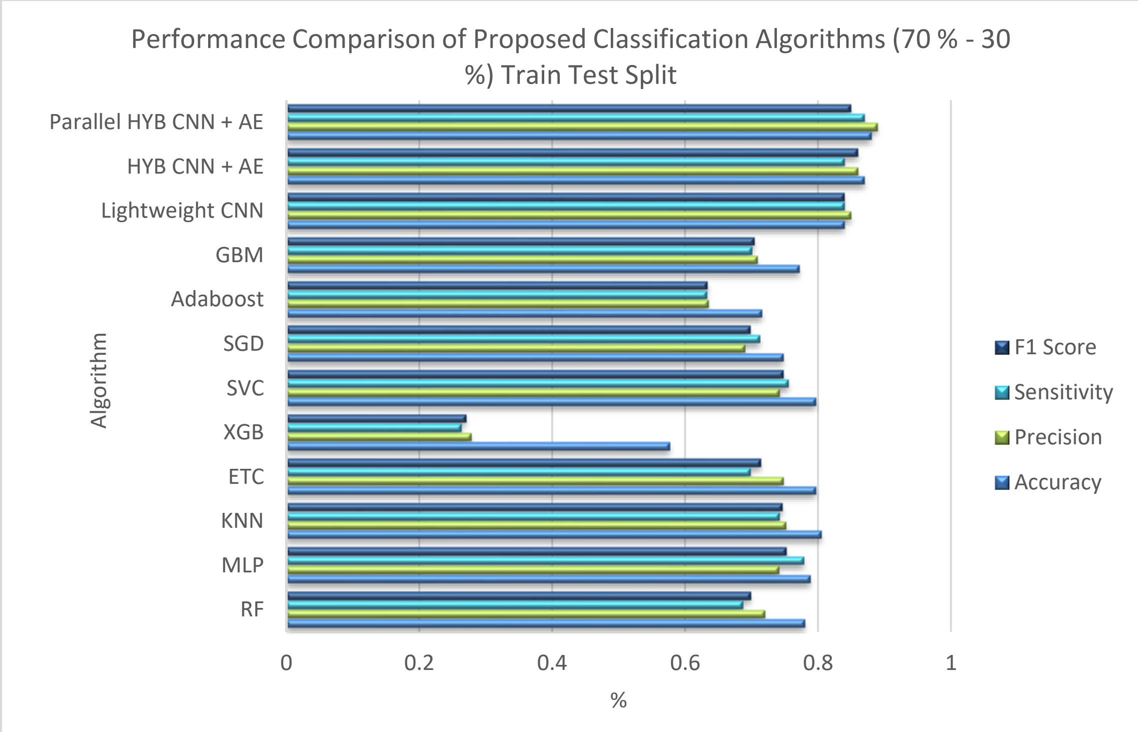

Fig.9. Performance Comparison graph of Proposed Classification Algorithms (70 % - 30 %) Train Test Split

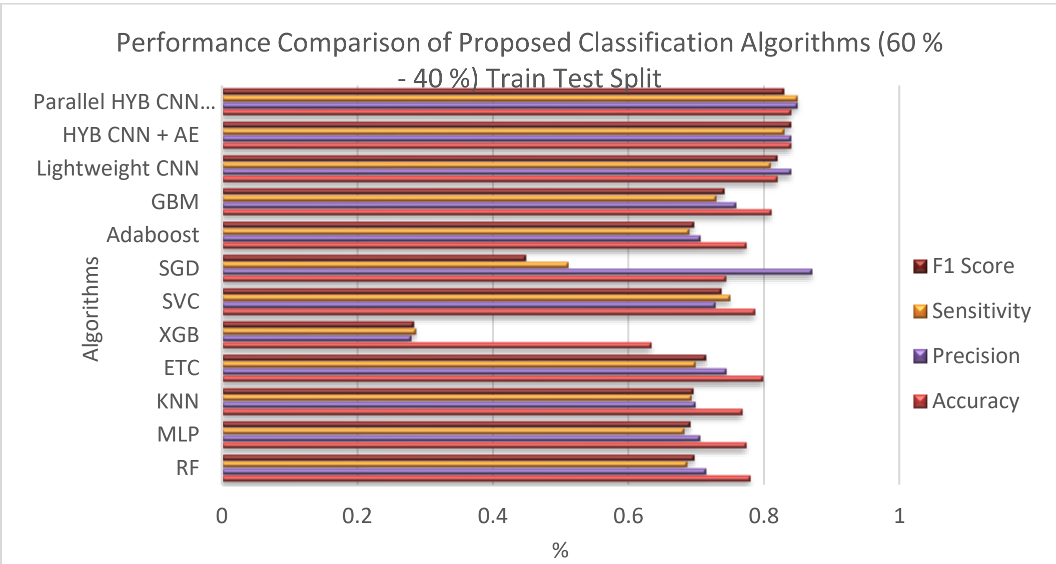

Fig. 10. Performance Comparison graph of Proposed Classification Algorithms (60 % - 40 %) Train Test Split

Accuracy and Loss Comparison Graph of Proposed Improved Models Figure depicts the accuracy and loss of Proposed models. Below figure shows the training and validation accuracy and loss curve for proposed algorithms, Lightweight CNN, Hybrid CNN – Autoencoder, Parallel Hybrid CNN – Autoencoder

Fig. 11 (A) Lightweight CNN Accuracy Comparison Graph, (B) Lightweight CNN Loss Comparison Graph (C) Hybrid CNN – Autoencoder Accuracy Comparison Graph. (D) Hybrid CNN – Autoencoder Loss Comparison Graph (E) Parallel Hybrid CNN – Autoencoder Accuracy Comparison Graph. (F) Parallel Hybrid CNN – Autoencoder Loss Comparison Graph

Confusion Matrix: Following figure shows the confusion matrix of best performing algorithm (Parallel Hybrid CNN – Autoencoder Classifier)

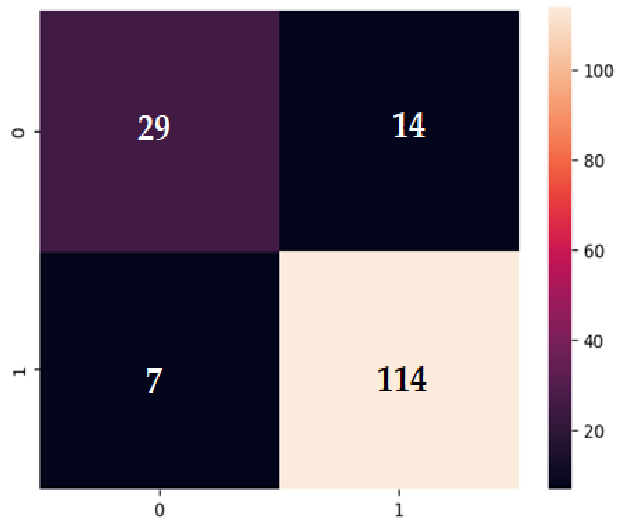

Fig. 12. Confusion Matrix of Heart failure detection using Parallel Hybrid CNN – Autoencoder Classifier

The Confusion matrix serves as a foundation for various performance metrics like accuracy, precision, recall, and F1-score, enabling a comprehensive assessment of a model's ability to classify data accurately and its capacity to avoid false positives and false negatives, which can be critical in decision-making and problem-solving scenarios.

4.4 Result Findings

- By examining Figures 9, 10 and 11, it becomes evident that the proposed improved deep learning classification algorithms consistently outperform the traditional machine learning algorithm across multiple performance metrics.

- Additionally, it is apparent that the 70-30% train-test split ratio yields superior results compared to other split ratios.
- Notably, among all the algorithms evaluated, the Parallel Hybrid CNN-autoencoder approach emerges as the top-performing method, demonstrating its effectiveness and robustness in heart failure detection

5. Conclusion

Congestive heart failure (CHF) remains a critical and life-threatening chronic condition, underscoring the need for efficient diagnostic tools and models. This study has successfully contributed to the field by establishing an effective correlation for CHF identification in patients. Through a comprehensive experimental results analysis of machine learning and deep learning and improved deep learning techniques applied to CHF detection models, this work has shed light on the diverse methodologies available for accurate diagnosis. Furthermore, we delved into the realm of heart sound classification, focusing on widely-used datasets such PhysioNet. Our findings emphasize that the feature extraction technique such as MFCC, Chroma, Spectral Contrast features performs better compare to others feature extraction techniques. These feature extraction techniques serve as critical tools in identifying specific patterns and characteristics within heart sounds indicative of chronic heart failure. The primary objective of this study was to thoroughly investigate and analyse existing models for heart failure prediction, particularly those involving heart sound data (PCG). Our performance analysis has yielded insightful results, indicating that the proposed lightweight model, hybrid model and parallel hybrid CNN-autoencoder architectures outperformed other models. This underscores the potential of these advanced approaches in enhancing the accuracy and efficiency of CHF detection, ultimately contributing to more effective patient care and early intervention.

References

[1] Li, F., Liu, M., Zhao, Y. et al. “Feature extraction and classification of heart sound using 1D convolutional neural networks”, EURASIP Journal on Advances in Signal Processing 2019, 59 (2019) https://doi.org/10.1186/s13634-019-0651-3

[2] Chen, Qianru, Wu, Zhifeng, Zhong, Qinghua Li, Zhiwei, “Heart Sound Classification Based on Mel-Frequency Cepstrum Coefficient Features and Multi-Scale Residual Recurrent Neural Networks” , Journal of Nanoelectronics and Optoelectronics, Volume 17, Number 8, August 2022, pp. 1144-1153(10) https://doi.org/10.1166/jno.2022.3305

[3] Li F, Zhang Z, Wang L, Liu W. ,”Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning”. Front Physiol. 2022 Dec 22;13: 1084420. https://doi.org/10.3389/fphys.2022.1084420

[4] Er M. B., “Heart sounds classification using convolutional neural network with 1d-local binary pattern and 1d-local ternary pattern features”, Appl. Acoustics, Vol 180, 108152, https://doi.org/10.1016/j.apacoust.2021.108152

[5] Li F., Tang H., Shang S., Mathiak K., Cong F.,” Classification of heart sounds using convolutional neural network”, Appl. Sci. 2020 Vol. 10, Issue 11, 3956 https://doi.org/10.3390/app10113956

[6] Li T., Yin Y., Ma K., Zhang S., Liu M. Lightweight end-to-end neural network model for automatic heart sound classification” ,Information 2021,Vol. 12, Issue 54 , https://doi.org/10.3390/info12020054

[7] schannen M., Kramer T., Marti G., Heinzmann M., Wiatowski T. (2016). “Heart sound classification using deep structured features,” in 2016 computing in Cardiology conference (CinC) (IEEE), Vol.43, 565–568.

[8] Tariq, Z.; Shah, S.K.; Lee, Y.,” Feature-Based Fusion Using CNN for Lung and Heart Sound Classification”, Sensors 2022, Vol. 22, Issue 4, 1521 https://doi.org/10.3390/s22041521

[9] T. S. Roy, J. K. Roy and N. Mandal, “Early Screening of Valvular Heart Disease Prediction using CNN-based Mobile Network,” 2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE), Kolkata, India, 2023, pp. 1-8, 10.1109/ICCECE51049.2023.10085513

[10] Abeer Ali Alnuaim, Mohammed Zakariah, Chitra Shashidhar, Wesam Atef Hatamleh, Hussam Tarazi, Prashant Kumar Shukla, Rajnish Ratna, "Speaker Gender Recognition Based on Deep Neural Networks and ResNet50", Wireless Comm. and Mobile Computing, vol. 2022, Article ID 4444388, 13 pages, Vol. 2022, https://doi.org/10.1155/2022/4444388

[11] Yang, W., Xu, J., Xiang, J. et al., “Diagnosis of cardiac abnormalities based on phonocardiogram using a novel fuzzy matching feature extraction method”, BMC Med Inform Decis Mak 22, Vol.230 (2022), https://doi.org/10.1186/s12911-022-01976-6

[12] Alkhodari M, Fraiwan L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. Comput Methods Programs Biomed. 2021 Mar; 200:105940. Epub 2021 Jan 17. PMID: 33494031 , https://doi.org/10.1016/j.cmpb.2021.105940.

[13] 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," in IEEE Access, vol. 9, pp. 36955-36967, 2021, https://doi.org/10.1109/ACCESS.2021.3063129.

[14] Al-Issa, Y., Alqudah, A.M.,"A lightweight hybrid deep learning system for cardiac valvular disease classification", Sci Rep 12, 14297(2022) https://doi.org/ 10.1038/s41598-022-18293-7

[15] Liu, Chengyu & Springer, David & Li, Qiao & Moody, Benjamin & Abad, Ricardo & Chorro, Francisco & Castells Ramon, Francisco & Roig, Jos & Silva, Ikaro & Johnson, Alistair & Syed, Zeeshan & Schmidt, Samuel & Papadaniil, Chrysa & Hadjileontiadis, Leontios & Naseri, Hosein & Moukadem, Ali & Dieterlen, Alain & Brandt, Christian & Tang, Hong & Clifford, Gari. "An open access database for the evaluation of heart sound algorithms", Physiological Measurement 2016,Vol. 37,Issue 12, 2181-2213, https://doi.org/10.1088/0967-3334/37/12/2181

[16] Hussein HI, Mohammed AO, Hassan MM, Mstafa RJ., "Lightweight deep CNN-based models for early detection of COVID-19 patients from chest X-ray images" Expert Syst Appl. 2023 Aug 1; Vol. 223: 119900. Epub 2023 Mar 18. PMID: 36969370; PMCID: PMC10023206, https://doi.org/10.1016/j.eswa.2023.119900

[17] Khozeimeh, F., Sharifrazi, D., Izadi, N.H. et al.," Combining a convolutional neural network with autoencoders to predict the survival chance of COVID-19 patients" , Sci Rep Vol. 11, 15343 (2021), https://doi.org/10.1038/s41598-021-93543-8

[18] Nahiduzzaman M, Goni MOF, Hassan R, Islam MR, Syfullah MK, Shahriar SM, Anower MS, Ahsan M, Haider J, Kowalski M. ," Parallel CNN-ELM: A multiclass classification of chest X-ray images to identify seventeen lung diseases including COVID-19", Expert Syst Appl. 2023 Nov 1;229: 120528. Epub 2023 May 27. PMID: 37274610; PMCID: PMC10223636, https://doi.org/10.1016/j.eswa.2023.120528

[19] Krishnan, Palani & Balasubramanian, Parvathavarthini & umapathy, Snekhalatha. "Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network", Physical and Engineering Sciences in Medicine 2020 Vol.43, Pages 505-515, https://doi.org/10.1007/s13246-020-00851-w

[20] Yadav, Anjali & Singh, Anushikha & Dutta, Malay & Travieso, Carlos. (2020). Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. Neural Computing and Applications.Vol. 32, Pages 17843-17856, https://doi.org/10.1007/s00521-019-04547-5.

[21] A. I. Humayun, S. Ghaffarzadegan, M. I. Ansari, Z. Feng and T. Hasan, "Towards Domain Invariant Heart Sound Abnormality Detection Using Learnable Filterbanks," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 8, pp. 2189-2198, Aug. 2020. https://doi.org/10.48550/arXiv.1910.00498

[22] Li, Fan & Tang, Hong & Shang, Shang & Mathiak, Klaus & Cong, Fengyu. (2020). Classification of Heart Sounds Using Convolutional Neural Network. Applied Sciences, Volume 10 ,Issue 11 , https://doi.org/10.3390/app10113956

[23] T. Mehta, O. Manolkar, S. Multaikar, S. Patil and N. Gawande, "Intensification of Agriculture using Deep Learning and Machine Learning: A Survey," 2021 10th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON), Jaipur, India, 2021, pp. 01-06, https://doi.org/10.1109/IEMECON53809.2021.9689196.

[24] Narváez, Pedro & Gutierrez, Steven & Percybrooks, Winston. (2020). Automatic Segmentation and Classification of Heart Sounds Using Modified Empirical Wavelet Transform and Power Features. Applied Sciences. 10. 4791. https://doi.org/10.3390/app10144791.

[25] N Gawande, D Goyal, "Empirical review on heart failure detection techniques using heart sound", AIP Conf. Proc. 2782, 020037 (2023), https://doi.org/10.1063/5.0154522

[26] N. Ahire, A. Dachawar, P. Disale, S. Galande and N. Gawande, "A Comparative Study on Facial Detection for Criminal Probe," 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-7, https://doi.org/10.1109/ASIANCON58793.2023.10270759

[27] Mehta, T., Multaikar, S., Patil, S., Gawande, N. (2024). Quality Prediction of a Stack Overflow Question Using Machine Learning. In: Sharma, H., Chakravorty, A., Hussain, S., Kumari, R. (eds) Artificial Intelligence: Theory and Applications. AITA 2023. Lecture Notes in Networks and Systems,

vol 844. Springer, Singapore. https://doi.org/10.1007/978-981-99-8479-4_6

[28] T. Mehta, S. Multaikar, S. Patil, O. Manolkar and N. Gawande, "A Comparative Study on Approaches for Text Quality Prediction using Machine Learning and Natural Language Processing," 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2022, pp. 1-5, https://doi.org/10.1109/SMARTGENCON56628.2022.10083911