MULTI-LAYERED ARCHITECTURE CONVOLUTION NEURAL NETWORKS FOR DIAGNOSING AND PREDICTING HEART DISEASES ON MULTI-MODAL DATA

Varsha [1*] ***, P Sardar Maran*** [2]

[1,2] Dept. of Computer Science and Engineering
Sathyabama Institute of Science and Technology, Chennai, India.

Emails: varsha.haridas90@gmail.com [1*], psmaran@sathyabama.ac.in[2]

ABSTRACT

Recently, heart diseases are increasing daily with the current lifestyle. Different types of cardiac diseases need to be diagnosed accurately for faster treatment to avoid death rates. Heart diseases are susceptible to attack a person speedily and cause sudden deaths. Medical experts are confused while diagnosing a person who comes in a medical emergency. In-time diagnosis, accurate prediction, and the proper treatment can save patients from heart diseases. Heart-related diseases cannot be predicted only by ECG data related to other medical information of a patient. This paper aims to develop an advanced artificial intelligence algorithm for analyzing and detecting heart diseases on multi-modal data. One of the advanced machine learning algorithms, Convolution Neural Network, is implemented with different architectural configurations for analyzing Body Mass Index (BMI), Electrocardiogram (ECG), and Physikalisch-Technische Bundesanstalt (PTB). Multiple CNN models are used for analyzing the multi-modal patient data to predict the presence of heart diseases. The CNN models are created by 80% of the training data and validated by 20% of the testing data. Different hyperparameters of the CNN models are tuned well to extract more data features to improve the prediction accuracy. The CNN models are developed in Python for processing and categorizing the feature variables as to whether heart disease is present or not. Medical data analytics require this model to get 98% of prediction accuracy.

Keywords: Artificial Intelligence; Machine Learning Model; Heart Disease; Prediction; ECG Data; BMI Data; Physikalisch-Technische Bundesanstalt (PTB).

1.0 INTRODUCTION

World Health Organization has published a report that heart disease is one of humans' most significant threatening diseases. The primary causes of heart disease are excessive cholesterol, high blood pressure, unhealthy eating habits, abnormal heart rhythms, obesity, and smoking. People affected by heart disease die suddenly due to inadequate diagnosis immediately at the earlier stages. Since it is the deadliest and leading cause of death globally, the medical industry has aimed to reduce the death rate due to heart failure and diseases. An efficient prediction model for heart disease using multi-modal data is needed. Most forecasting, classification, and prediction applications use efficient machine and deep learning algorithms. CHD (Congenital Heart Disease) forecasting in a patient might be a binary classification task. Under certain conditions, the neural network has proved to be an effective classifier. Recently, neural networks have been configured with several hidden layers under application usage, notably improved in various areas. The application area of the neural networks includes speech processing, image processing, and time series prediction. Large datasets are used in the training process to create various architectures. The function of the artificial neural network includes modification of input data over a hidden layer and calculation of error in the output layer. The gradient descent algorithm has used the errors in the output to update the layer weights. Various experiments and analyses have proposed multiple improvements in the gradient descent algorithm. By making the layer nonlinear, the training process has been scheduled.

One of the primary aims of the research is to improve the accuracy of predicting, which is critical in medical applications. Multiple algorithms have been proposed to resolve various issues in classification. For example, data augmentation is one technique that artificially populates. Small datasets lead to poor training and inaccurate type. Malicious prediction leads to a more significant penalty than image synthesis. Because of poor prediction strategies, patients with CHD have been left untreated, leading to wrong therapeutic medications. Hence, this research proposed CNN(convolutional neural network) to overcome these accuracy issues. CNN is an artificial neural network primarily utilized for image processing and recognition. It is also known as converts and Neural Networks. In CNN, dimensions are represented by different colors. Imagine there is an image that incorporates a cuboid that encompasses length, width, and height. CNN adds an extra layer called convolutional. With the help of deep learning, we can easily take a 3D image as an input compared to the preceding format of artificial neural networks, which could only take input vectors consisting of some features. It can detect and distinguish objects in an image. To analyze data, CNN uses perceptrons for supervised learning and applies them to NLP and other cognitive tasks. CNN is a basic example of deep understanding where sophisticated models push by offering systems that simulate different biological human brain activity. Convolution allows the amalgamation of two sets of information. The application of the convolution layer is to filter the input data and produce a feature map where the filter is called a kernel. The kernel does a matrix multiplication of elements one after another and goes over the input element. The feature notices the outcome of each field. Striding and padding operations are utilized in processing images accurately. Building an image by adding pixels when processing by the kernel of CNN, termed padding, is one of the processes in CNN. When we set padding to ZERO, every pixel value becomes zero. A more accurate analysis of the image is done in padding.

The convolution operation of the CNN algorithm uses padding, striding, and sliding operations to learn and extract the image features. The number of features that can be extracted depends on the size of the striding operation. The dimension of the image can also be reduced during the preprocessing task and the pooling layers in the CNN. Max pooling is a sample-based process that transfers from continuous to discrete functions. The main aim is to reduce input by downscaling its dimensionality and processing rejected features in the subregion. When the input matrix is divided, it downscales into a rectangular region and calculates the average value of each area. This paper proposed multi-configured CNN processes in three different datasets, BMI, ECG, and PTB, for predicting different types of heart disease. It is a hybrid model emphasizing the features obtained from different datasets for heart disease prediction. Thus, this paper makes the following contributions, they are:

- Process the PTB dataset and extract R waves from the ECG signals.
- Separate CNN models for processing BMI, ECG, and PTB models.
- Combining the results from different datasets and making an effective prediction.

With the healthcare sector's development, the generation of healthcare data is enormous. Additionally, for the efficient treatment of patients, various parameters need to be analyzed to conclude. Using AI and ML algorithms helps in the faster prediction of diseases. These models can be trained to provide better accuracy. They need vast amounts of data to train to make efficient predictions of diseases. The ML and DL algorithms can be trained effectively, and faster prediction of various heart diseases can be made using such a large amount of data. Among the different deep learning algorithms, CNN is widely used in medical image processing and ECG signal analysis. Prediction of heart diseases can be easily made by processing the ECG signals.

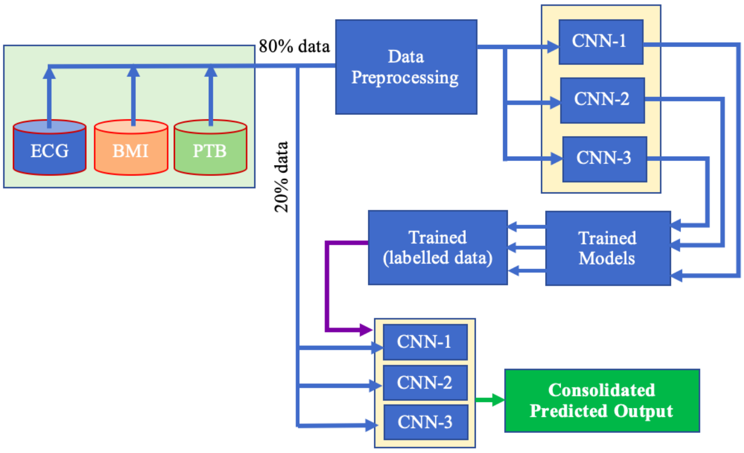

Fig. 1: Proposed Methodology

This paper aims to identify and confirm whether heart disease is present in the dataset. However, most research works that use the CNN algorithm only concentrate on any dataset for their prediction process, leaving them to predict any specific disease. Hence, a configurable CNN is built and trained with multiple datasets like BMI, ECG, and PTB to accurately predict different heart disease types. The proposed model is compared with similar research works, and their efficiency is evaluated and shown in Fig. 1.

2.0 LITERATURE SURVEY

Before designing the proposed research methodology, understanding the issues and challenges of the existing methods is essential. It helps to create a new roadmap for creating a better solution. For example, [1] have focused on predicting the possibility of Heart Disease by using affordable medical tests that can be administered in local clinics. They have concluded that this technique enhances the accuracy and trustworthiness of physicians' diagnoses. The author proposed a majority voting ensemble technique that makes a prediction based on the prediction score obtained from various ML methods and got 90% accuracy. [2] have outlined a technique that utilizes the random forest algorithm to detect an instance of Heart Disease. By analyzing clinical data and test results from multiple patients, this approach accurately determines the presence of heart disease in individuals. Finally, this could provide valuable insight for medical professionals when deciding upon patient treatment options. [3] have discussed a system predicting heart disease in real-time. It uses a robust computer framework called Apache Spark that can quickly handle lots of data and calculations. The system has two main parts: one that analyzes incoming data and predicts heart disease and another that stores and displays the data. Finally, the system keeps the data using a unique Apache Cassandra tool because much information is generated. [4] have suggested combining decision trees and artificial neural network classifiers to improve heart disease prediction. They used WEKA software to test their idea on a dataset of heart disease patients. They checked how accurate and reliable their method was by running different tests. Finally, they analyzed and classified the performance metrics.

[5] used deep learning methods to detect heart disease by listening to the sounds of the heart. They have used a computer program that can tell if the heart sounds normal or if there is a problem. They concluded that our test showed that this method works well in real life. [6] examined how a mother's obesity and diabetes can increase the chances of her child having heart disease. They focus on understanding how this happens and how we can detect and prevent it before the baby is born. They have concluded that by learning more about the risks for mothers, we can help identify at-risk pregnancies and provide better care to prevent heart disease in babies. [7] have studied to enhance the efficiency of diagnosing congestive heart failure (CHF) more accurately. They have used a computer model called DenseNet and created a diverse CHF database to test it. They have also devised a way to measure how well the program worked. Finally, our results showed that our program was very good at detecting CHF, with an accuracy of 94.97%. [8] have introduced an IoT-enabled ECG system to monitor and analyze ECG results. They used a small device that can monitor your heart's electrical signals, which is called an ECG. The device can tell if your heart rate is normal or abnormal. It calculates different signal features, like how fast your heart beats and how regular the beats are. This information can help doctors diagnose heart problems. It is easy to use and can be used at home, so you do not always have to go to the hospital.

[9] has introduced a new method to check for heart diseases using a smartwatch and a heart monitor meter. It uses a computer program called a Modified Deep Convolutional Neural Network (MDCNN) to analyze the data from the monitor and determine whether the person has a healthy or unhealthy heart. Finally, it has been proved that the new method is better than other ways of checking for heart disease and can correctly identify a person's heart health with 98.2% accuracy. A new method has been proposed to predict a patient's presence probability. Still, the experimental result also shows that the proposed method outperforms and achieves an accuracy rate of 97%. [10] have discussed the effect of obesity and diabetes on the structure and function of the heart muscle. They have also demonstrated how advanced imaging techniques can help to understand the underlying mechanisms of myocardial dysfunction, predict long-term clinical outcomes, and inform treatment options. [11] stated that heart disease is severe in developing countries and is currently the primary cause of death in the world. Hence, detecting heart disease in the early stage helps the patient to prevent danger but also helps the practitioner learn the significant cause of the heart attack and help avoid the danger before the actual occurrence inpatient. [12] [13] stated that diagnosing heart disease using an end-to-end deep learning method using the ECG signal channel has been presented in the study. Due to the popularity of deep learning algorithms for time classification, series based on the one-dimensional CNN, unlike the traditional CNN model-based classification, have been proposed. Various cardiac arrhythmias have been recognized automatically in the new multi-level wavelet CNN. The proposed method attains a maximum accuracy of 99.57%.

[14][15] stated that a deep learning system was developed to assist the cardiologist with heart arrhythmia disease classification. The classification attains an accuracy rate of 99.37%, which shows that heart-related complications lead to health-threatening conditions. Hence, automatic detection of arrhythmia heart disease in an early stage is of the highest interest, and the mortality rate of cardiac disease patients level would be reduced. A fully connected layer architecture has been proposed using different network optimizers. [16] stated that heart disease is one of the fatal diseases that take the most lives compared to other diseases worldwide. Hence, early detection of the disease helps to save many valuable lives. Heart disease is detected using ECG, heart sound, and computed tomography

images, but the accuracy of heart disease detection is low and imbalanced. The GAN model performs the best way to detect disease potentiality. The detection methodology is applied to other diseases and healthcare problems. In [15] stated that Congestive heart failure associated with other high prevalence is a severe psychophysical condition, hence demanding efficient methods for detecting heart disease. The current methodology has advanced in CHF detection and caters to clinical practitioners. The recent method is focused on advanced signal processing and machine learning. The potential process applies a convolutional neural network for automatically detecting CHF, which needs to be addressed so far. The model has been tested on available ECG datasets for 100% detection accuracy.

3.0 MATERIALS AND METHODS

Heart disease detection and prediction is one of the most essential processes in the medical industry. ECG plays a vital role in detecting heart diseases. It can provide the severity level of the root causes of heart disease. Confirmation of heart disease can not be confirmed by only analyzing the ECG data, which needs more physical and logical information about the patient. Thus, the CNN model analyzes various types of patient data (ECG, BMI, PTB) to predict heart diseases and their symptoms. The heart disease datasets used in this paper are taken from various internet sources that are freely available for research purposes, like the UCI repository, Kaggle dataset, and Physionet websites.

Table 1: Dataset Used in The Experiment

Dataset	Total number of data	Number of Features
ECG	109446, 5 categories	5 classes, 234757 trained parameters
PTB	21837	71 features
DMI	21837	9 features

A PTB-XL dataset from the Kaggle open-source database is used in this paper. This dataset consists of 21837 ECG signals from 12 clinics, collected from 18885 patients. Each ECG signal is about 10 seconds long. The data collected is corrected and annotated by two cardiologists. They have assigned ECG statements for every single ECG. The dataset is split into training and testing datasets. The waveform data underlying the PTB-XL ECG dataset was collected with devices from Schiller AG over nearly seven years between October 1989 and June 1996. The records were curated and converted into a structured database within a long-term Physikalisch-Technische Bundesanstalt (PTB) project. With the acquisition of the original database from Schiller AG, the total usage rights were transferred to the PTB. Along with the annotations, it helps better train and test the ML and DL models. Various interpretation algorithms are proposed to process ECG signals, and they help medical professionals plan for the treatment quickly and implement it in automatic diagnosis applications. However, the prediction model must be trained with large datasets to develop such algorithms. This aspect is not covered in any of the freely available ECG datasets.

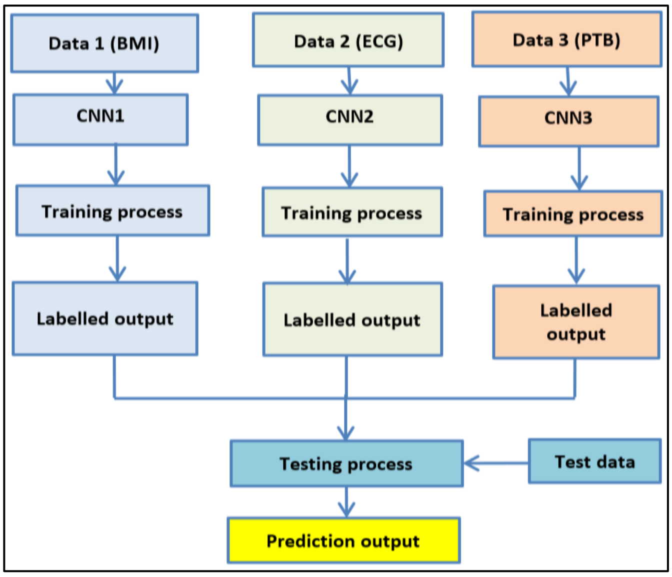

Fig. 2: CNN Model For Multi-Modal Data Analysis

3.1 Methods

This paper uses three CNN models with different layer architectures to analyze ECG, PTB, and BMI datasets. CNN1 is used to analyze the BMI dataset, CNN2 is used to analyze the ECG dataset, and CNN3 is used to analyze the PTB dataset. From the three datasets, 80% of the data is taken for creating the corresponding CNN model by the training process. The trained data is annotated with the heart disease class, also called data labeling. The trained data is compared with the output obtained from the test process.

3.2 ECG Processing or Heart Disease Prediction

Analyzing the ECG signals recorded on a time series can quickly evaluate the heart's state and the heart disease's impact. The ECG signals are represented in the form of floating point values that are plotted on time series to obtain the wave format. These wave formats correspond to the functioning of the patient's heart. In the ECG data analysis, various time series analysis methods are used to help find the relative deflection of the data, which gives the state of the heart. The abnormal deflections in ECG or those without deflections correspond to heart disease or dysfunction. The ECG raw signals considered in this paper consist of different peaks like R-peaks, QRS complex, and entropy features. From the annotations present in the dataset, the R-peaks can be easily detected. The entropy features of the QRS complex can also be easily found. After extracting the entropy features, the dataset is split for training, testing, and validation. The proposed model is first trained and tested, and the results are discussed in detail. The raw signals are recorded from the patient's body in a compressed format. Each signal has 12 leads with various electrodes on the patient's right arm. The Metadata of each signal is collected with SCP-ECG statements from cardiologists and used as an annotation. From the annotation, ha, is-1, and is-2 are verified, and each signal's final label is assigned.

3.3 R Wave Detection

Some essential components of ECG signals are the R-wave, T-wave, and QRS complex. The QRS complex is a crucial ECG feature that helps find cardiac abnormalities. The initial stage of that prediction process is the R-peak prediction from the QRS complex. Though the given dataset has the same time, the beats are not at the same time. The waves are cut into 10 parts, with each piece having a single R-wave. It is not difficult to find the R wave within the ECG Waveform. Various detectors can predict R-wave, like the Hamilton, Two average, Stationary Wavelet Transform, Christov, Pan-Tompkins, and Engzee with modification. The following methods provide the R wave's position in the ECG signal. The dataset consists of 12 lead records that help in detecting the position of the R-wave accurately. After finding the R-waves from the ECG signals, the total number of R-waves in one signal is calculated. The average number of waves in an ECG signal is calculated. Let the record be n_R, and the BPM for each record is calculated. It can be defined in the following manner,

$$\boldsymbol{f_1, \ldots., f_n : R^{5000} \rightarrow r_1, \ldots., r_n}$$

$$\boldsymbol{F = \{f_i \vee i \epsilon N^{+} i \leq n\}}$$

$$\boldsymbol{C_i = \{|F_j(X_i)|; j \epsilon N^{+} j \leq |F|\}}$$

$$\boldsymbol{\mu_{i,1/2} = C_{i,(|C_i|+1/2)}}$$

Where the ith ECG signal is represented as $\boldsymbol{X_i}$, and functions that are involved in processing the R-wave are represented as $\boldsymbol{f_1, \ldots., f_n}$**.** These functions are obtained from 5000 samples modified into points representing the R-wave center. $\boldsymbol{F}$ is defined as the function used for extracting the R-waves. The cardinalities present in the R-wave are represented as $\boldsymbol{C_i}$. The average cardinality of the R-waves is $\boldsymbol{\mu_{i,1/2}}$. The total number of functions is $\boldsymbol{N}$, and the positive natural numbers are represented as $\boldsymbol{N^{+}}$. The position of the R-waves detected from the detector helps to identify the R-waves in the ECG signal. The position of the R-wave is clustered, and the total number of R-peaks is to be considered **K.** Then, the position where most R-waves are clustered around **K** is considered the R-wave location.

3.4 Entropy-Based Features

These features are very important for the prediction of cardiac diseases. The accuracy of the proposed model increases with the inclusion of entropy-based features. In this paper, the ECG signals are considered for the extraction from ECG signals, where some of the widely considered entropy-based features are Shannon, Approximate, Sample, Permutation, Spectral, SVD, Renyi, Tsallis, and Extropy. Based on the research conducted by Granelo-Belinchon et al. [60], various non-stationary signals having short periods can be measured through information theory. However, it does not apply to the ECG signals, as they are dynamic. However, when they were split into single-second from 10-second clips, 89.5% of the clips were stationary. The records are split for training, testing, and validation. Some essential features extracted from the signals are raw signals and QRS from 12 leads, entropies for raw signal and QRS, and class and subclass from the records.

3.5 BMI Analysis for Heart Disease Prediction

Body mass index is a collection of data that measures physiological information about a human. It is expressed in **kg/m^2**. As the BMI value is obtained from the weight and height of the patients, the BMI value is obtained by dividing the patient's weight by the square of the height in meters. The BMI data is used to evaluate the primary health status of the patient and, when combined with other results like ECG and BMI, helps in the faster prediction of heart disease. Along with weight and height, the gender of the patients also has an impact. It helps for a comparative analysis of the patient.

3.6 PTB Analysis for Heart Disease Prediction

As the ECG analysis shows the state of heart disease and its impact on the heart, it is challenging to find different diseases from ECG through normal analysis. Large datasets are needed to train the model to detect patterns from the ECG signals accurately. With data availability and open-source models, various pre-trained and preprocessed ECG datasets are available. These data can be used to evaluate heart diseases. It is achieved using ECG interpretation algorithms that detect special signal patterns, and these patterns are fed into diagnosis support systems that make efficient predictions from the available data. The steps involved in the PTB extraction process are given in Figure-3. Each signal consists of varying numbers of QRS complexes, and if smaller numbers are taken as solutions, there might be a significant information loss. This variance can be minimized using divisible sub-networks: encoding blocks, max pooling, average pooling, and fully connected layers. Let the input QRS signals be in the form of,

$$\boldsymbol{X_i = QRS_1, QRS_2, \ldots, QRS_n; n\epsilon N^+}$$

The encoding function takes 100 samples simultaneously through the 12-channel signal, and the output is a 24-dimensional vector.

$$\boldsymbol{g: R^{12x100} \rightarrow R^{24}}$$

It is used to extract the QRS code in the input signals.

$$\boldsymbol{Z_i = \{g(X_i, j) \vee j\epsilon N^+ \cap j < |X_i|\}}$$

Fig. 3: Pattern Extraction from ECG signals for PTB dataset

The Max and average pooling processes are used to process the QRS patterns in the ECG signals. The 24-dimensional variable length vectors are obtained as $\boldsymbol{Z_i}$. The maximum values are selected during the maximum pooling process, and the maximum value is chosen in each dimension of the vector set.

$$\boldsymbol{Z_{max_i}} = \left[\boldsymbol{max}\left(\left\{\boldsymbol{Z_{i,j,1}} \middle| \boldsymbol{j \epsilon N^{+}} \cap \boldsymbol{j} < |\boldsymbol{X_i}|\right\}\right), \dots, \boldsymbol{max}\left(\left\{\boldsymbol{Z_{i,j,24}} \middle| \boldsymbol{j \epsilon N^{+}} \cap \boldsymbol{j} < |\boldsymbol{X_i}|\right\}\right)\right]$$

The average pooling process finds the average value in the matrix of vectors, which can be seen in the following equation,

$$\boldsymbol{Z_{avg_i}} = \left[\frac{\boldsymbol{1}}{|\boldsymbol{Z_i}|}\sum_{\boldsymbol{j=1}}^{|\boldsymbol{Z_i}|} \boldsymbol{Z_{i,j,1}}, \dots, \frac{\boldsymbol{1}}{|\boldsymbol{Z_i}|}\sum_{\boldsymbol{j=1}}^{|\boldsymbol{Z_i}|} \boldsymbol{Z_{i,j,24}}\right]$$

The maximum and average values from the matrix are then selected, and the fully connected layers that pass 48 units of data with 20 neurons, and the final output is obtained in the form of a 20-dimensional final vector,

$$\boldsymbol{Zall} = [\boldsymbol{Zmax_i}, \boldsymbol{Zavg_i}]$$

$$\boldsymbol{Zfinall_i} = \boldsymbol{f}(\boldsymbol{Zall}); \boldsymbol{f}: \boldsymbol{R^{48}} \rightarrow \boldsymbol{R^{20}}$$

The encoding and decoding are done to process the QRS complex, and the CNN algorithm helps. The ReLU function is used as an activation function, and the negative slope coefficient is represented as $\boldsymbol{\alpha}$, and maintained at 0.01. The final output is flatted to a 24-dimensional layer.

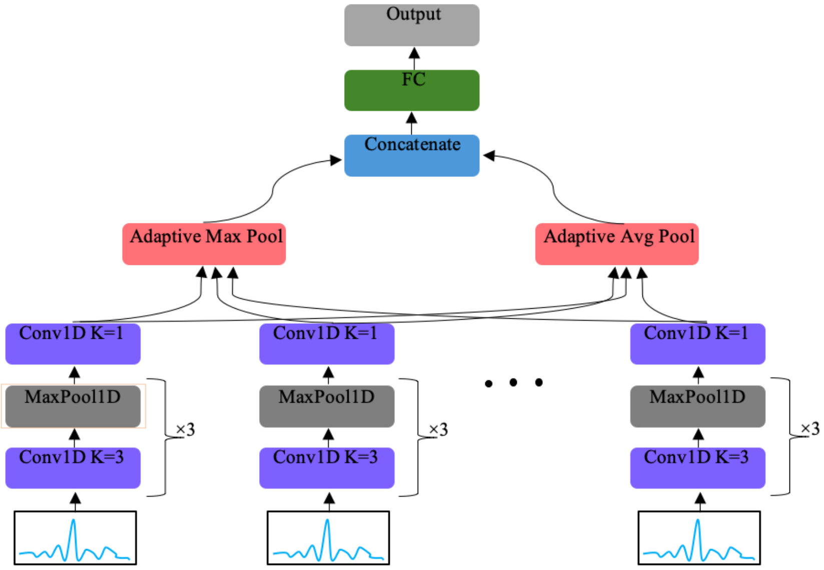

Fig. 4: Hybrid CNN architecture

3.7 CNN Architecture

CNN works on a single input, single output fashion experimentation for binary classification. The patients with the presence of CHD would be classified as **1**, and other CHD-absent patients are classified as **0** using a majority voting algorithm. Active CHD attributes have been derived. Let us assume the number of training as **N**. Hence, the Input layer has been noted as **R^N*14**. 14 CHD attributes have been selected, so the variables chosen by voting are merged in the dense, fully connected layer with 64 neurons. Multiple layers are effectively normalized in the dense layer before the nonlinear transformation, which is done by rectifying linear units or relying on them. The dropout layer conducts the function of dropout by 15 in standard cases. The dense, fully connected layer followed the convolution layer with a cascaded set. There are two filters with no padding and 4 kernels, and stride with the size of 2 are introduced in the first convolution layer with average pooling where the pooling regains the average of features. In contrast, it smoothens the image while keeping the feature's importance in an image. Externally, no padding is added in this convolution layer. Different experiments with various strategies are conducted in the pooling layer to find the average pooling size under all constraints. Output from the fully connected layer is convened into a tensor in the first convolution layer. The convened tensor flow is further implicated in various manipulation activities like nonlinear transformation. Batch normalization and average pooling are to generate an output fully connected to tensor dimensions.

The second convolution layer consists of 4 filters with no external padding added to it. It is handled with six kernels and a stride value of 2 and optimized to generate the output tensor of convolution layer 2. The output from the tensor is then delivered to the average pooling of the convolution layer 2. The output from the average pooling is then forwarded into a dense layer and later into average pooling. The softmax layer is where we can note the categorical production of the loss function, which is set as a class-cross-entropy loss. Initialization bias with some random numbers with N as the number of connections to the layers coming from the previous layer, some random number has been drawn from a normal distribution. Adam optimizer has been used with some specific value. A model has been proposed with various hyperparameters to obtain accuracy with classification consistency. In the training phase, varying epochs provide the results with particular values. Each dense layer has been manipulated except for the last one. During the training phase, the number of filters also varies.

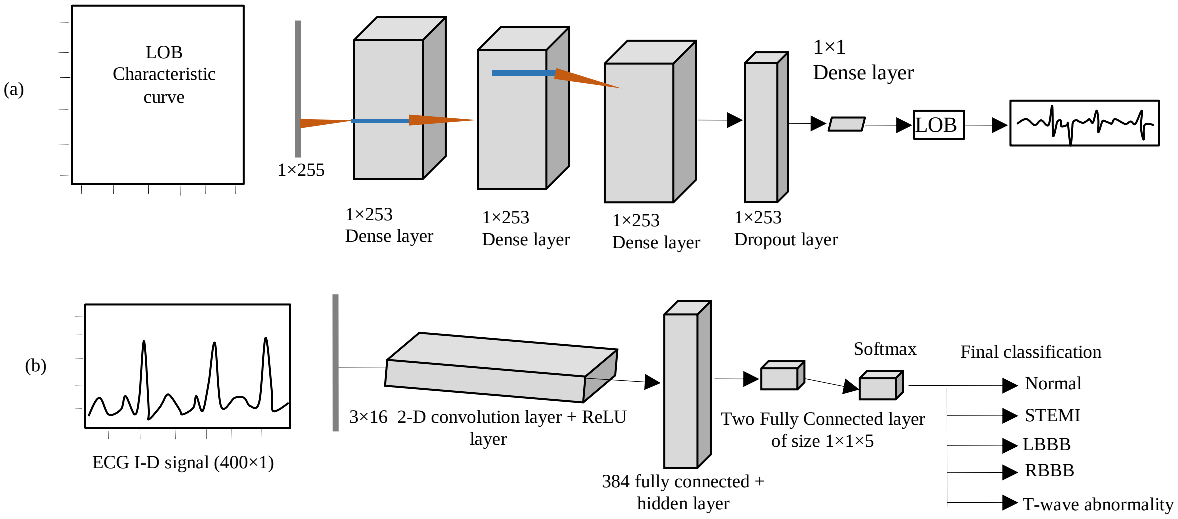

Fig. 5: Architecture of CNN1 and CNN2

Among the ECG samples, 12 leads are converted into single leads, and the possible threshold values are found. All the signals are already available as a 1-D characteristic curve. These curves are then converted into 1D-characteristic curves, which are processed through the CNN model. They are plotted, the R-waveform and the QRS signals are detected, and the LBBB, RBBB, T-wave abnormality, and STEMI can be detected. The architecture of CNN1 and CNN2 can be seen in Figure-5.

Table 2: CNN1 architecture

Layer (type)	**Output Shape**	**Param #**
conv1d (Conv1D)	(None, 182, 32)	192
conv1d_1 (Conv1D)	(None, 182, 32)	5152
conv1d_2 (Conv1D)	(None, 182, 32)	5152
max_pooling1d (MaxPooling1D)	(None, 89, 32)	0
conv1d_3 (Conv1D)	(None, 89, 32)	5152
conv1d_4 (Conv1D)	(None, 89, 32)	5152
max_pooling1d_1 (MaxPooling 1D)	(None, 43, 32)	0
conv1d_5 (Conv1D)	(None, 43, 32)	5152
conv1d_6 (Conv1D)	(None, 43, 32)	5152
max_pooling1d_2 (MaxPooling 1D)	(None, 20, 32)	0

conv1d_7 (Conv1D)	(None, 20, 32)	5152
conv1d_8 (Conv1D)	(None, 20, 32)	5152
max_pooling1d_3 (MaxPooling 1D)	(None, 8, 32)	0
conv1d_9 (Conv1D)	(None, 8, 32)	5152
conv1d_10 (Conv1D)	(None, 8, 32)	5152
max_pooling1d_4 (MaxPooling 1D)	(None, 2, 32)	0
flatten (Flatten)	(None, 64)	0
dense (Dense)	(None, 512)	33280
batch_normalization (BatchNormalization)	(None, 512)	2048
dense_1 (Dense)	(None, 256)	131328
batch_normalization_1 (BatchNormalization)	(None, 256)	1024
dense_2 (Dense)	(None, 64)	16448
batch_normalization_2 (BatchNormalization)	(None, 64)	256
dense_3 (Dense)	(None, 5)	325
Total params: 236,421 Trainable params: 234,757 Non-trainable params: 1,664		

3.8 PTB Prediction

The PTB-XL classification involves data classification in a multi-labeled format. This CNN model is used for processing the ECG signals. However, a separate CNN model is used to process the PTB dataset and predict heart diseases. Another one is a 1D convolutional network trained with ECG signals. It is similar to CNN, which processes raw ECG signals, but is a different model in which signals are tuned explicitly for PTB through R-wave detection and QRS analysis. The CNN model is trained with three datasets: **X, Y**, and **Z**. The layers are initially created, trained with separate datasets, and combined to form the overall model. Initially, a fully connected dense network, consisting of 2 dense layers and a dropout regularization layer, was used. These layers are trained with the X dataset. This network is concluded with a regularisation and dense layers. The sigmoid activation function is used for feature extraction and prediction. This model is trained once and then followed by the next model, which is trained with Dataset Y. This CNN model consists of 1D convolutional layers followed by batch normalization and activation. A Maxpooling layer of **1** dimension separates these layers. The data size increases as the Convolutional layers increase the number of features. So, pooling and dropout layers are the final layers for the prediction process. This model is trained with Dataset Y and once trained. Finally, all the trained layers are combined in the final model. Both the ECG curves and Patient data are incorporated in the third model. The pre-trained models are used to predict heart diseases accurately in this model.

Table 3: CNN1 And CNN2 Architecture

Layer (type)	**Output Shape**	**Param #**	**Connected to**
Y_inputs (InputLayer)	[(None, 800, 12)]	0	
Y_conv_1 (Conv1D)	(None, 800, 64)	5440	Y_inputs[0][0]
Y_norm_1 (BatchNormalization)	(None, 800, 64)	256	Y_conv_1[0][0]
Y_relu_1 (ReLU)	(None, 800, 64)	0	Y_norm_1[0][0]
Y_pool_1 (MaxPooling1D)	(None, 400, 64)	0	Y_relu_1[0][0]
Y_conv_2 (Conv1D)	(None, 400, 128)	24704	Y_pool_1[0][0]
Y_norm_2 (BatchNormalization)	(None, 400, 128)	512	Y_conv_2[0][0]
Y_relu_2 (ReLU)	(None, 400, 128)	0	Y_norm_2[0][0]
Y_pool_2 (MaxPooling1D)	(None, 200, 128)	0	Y_relu_2[0][0]
X_inputs (InputLayer)	[(None, 7)]	0	
Y_conv_3 (Conv1D)	(None, 200, 256)	98560	Y_pool_2[0][0]
X_dense_1 (Dense)	(None, 32)	256	X_inputs[0][0]
Y_norm_3 (BatchNormalization)	(None, 200, 256)	1024	Y_conv_3[0][0]
X_drop_1 (Dropout)	(None, 32)	0	X_dense_1[0][0]
Y_relu_3 (ReLU)	(None, 200, 256)	0	Y_norm_3[0][0]

X_dense_2 (Dense)	(None, 32)	1056	X_drop_1[0][0]
Y_aver (GlobalAveragePooling1D)	(None, 256)	0	Y_relu_3[0][0]
X_drop_2 (Dropout)	(None, 32)	0	X_dense_2[0][0]
Y_drop (Dropout)	(None, 256)	0	Y_aver[0][0]
Z_concat (Concatenate)	(None, 288)	0	X_drop_2[0][0]
Z_dense_1 (Dense)	(None, 64)	18496	Z_concat[0][0]
Z_dense_2 (Dense)	(None, 64)	4160	Z_dense_1[0][0]
Z_drop_1 (Dropout)	(None, 64)	0	Z_dense_2[0][0]
Z_outputs (Dense)	(None, 5)	325	Z_drop_1[0][0]
Total params: 154,789			
Trainable params: 153,893			
Non-trainable params: 896			

The X and Y datasets representing the BMI and ECG datasets are combined and given as input, and Z is obtained as the output. The features obtained from the Metadata and the curves are combined. Two fully connected dense layers achieve this combination. A single dropout and dense layer then succeed it. The activation function used in the dense layer is the sigmoid function. Again, the generated model is fitted, and the callbacks from the model are checked.

4.0 RESULT AND DISCUSSION

In this paper, the input data are collected from three datasets, including every patient's electronic reports of ECG, PBT, and BMI. Applying the preprocessing techniques removes the overall artifacts in the input datasets. Once the anomalies are removed from the input data, they are classified into three phases: training, testing, and validation. For this classification, from the overall dataset, 20% of data from each dataset is used for testing, and the remaining 80% of data in every three datasets is used for training.

The final classification results, such as patients with and without heart diseases, are predicted accurately from the training and testing results. These results are also more helpful in improving the efficiency of the proposed approach.

4.1 Performance Metrics

In the deep learning algorithm, the efficiency of the proposed model is evaluated using various performance metrics such as precision, recall, f1-score, and accuracy. These metrics more accurately detect the classification efficiency of the proposed CNN model. The following section briefs the function of each performance metric.

Precision- This metric is calculated to detect the number of positive predicted values from the total positive inputs. It is performed using the equation (1).

$$\boldsymbol{Precision = \frac{TP}{TP+FP}} \qquad \textbf{(1)}$$

Recall- This is used to predict the correctly predicted value from the total number of inputs. It is evaluated by dividing the TP value by the sum of TP and FN values. It is expressed in the equation (2).

$$\boldsymbol{Recall = \frac{TP}{TP+FN}} \qquad \textbf{(2)}$$

F1-score- The resultant value of this metric is evaluated by combining the precision and recall values. Calculating when precision or recall values are low or high is complicated. Using the equation-(3) F1-score is evaluated.

$$\boldsymbol{F1 - score = 2 \times \frac{(precision \times recall)}{(Precision + recall)}} \qquad \textbf{(3)}$$

Accuracy- This is one of the significant metrics to define the overall efficiency of the deep learning algorithm in predicting the final classification result. It is clearly expressed in the equation (4). It is evaluated by the ratio between correctly predicted values and the total number of prediction values.

$$\boldsymbol{Accuracy = \frac{TP+TN}{TP+TN+FP+FN}} \qquad \textbf{(4)}$$

In the above equations (1), (2), and (4), TP, TN, FP, and FN represent the total number of true positive, true negative, false positive, and false negative values of the proposed model rn, respectively.

4.2 Experimental Results and Discussion

In this section, the simulation result of the proposed approach is discussed in detail. In Figure 6, the total number of data classified for training and testing is illustrated and defined from the overall datasets; 80% is used for training, and 20% is used for testing.

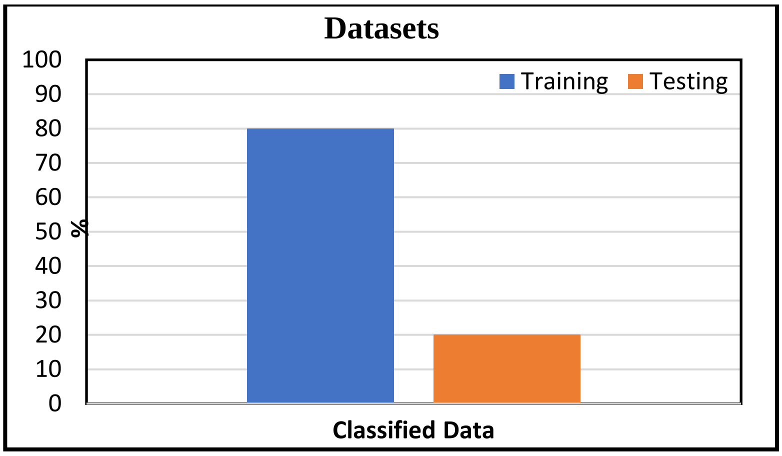

Fig. 6: Classified Input Data

After classifying the input datasets into two phases, the confusion metrics value of the proposed model is evaluated to define its performance, which includes the proposed approach's TP, TN, FP, and FN values. The actual number of negative and positive predicted class values is clearly defined in Figure 7.

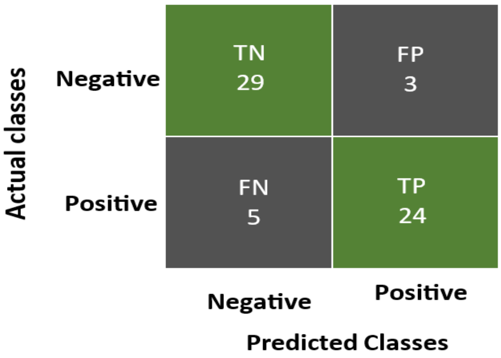

Fig. 7: Confusion Metrics of The Proposed Approach

Table-1 defines the proposed approach's various performance metrics values such as precision, recall, f1-score, and accuracy. The experimental result shows that the proposed model has achieved 95.71%, 90.02%, 85.73%, and 87.92% accuracy, precision, recall, and F1-score values, respectively.

Table 4: Performance Calculation

Performance Metrics	Accuracy (%)
Accuracy	95.71
Precision	90.02
Recall	85.73
F1 score	87.92

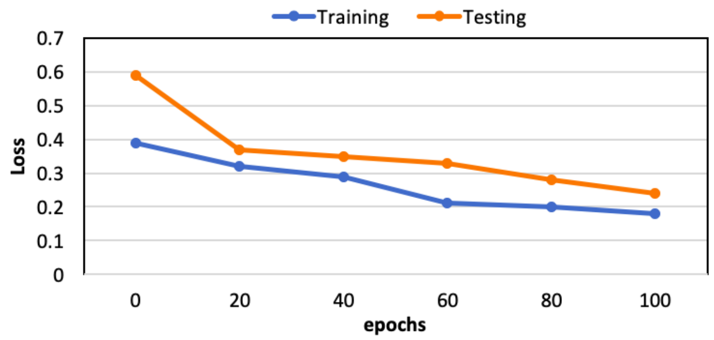

Fig. 8: Lose Vs. Epochs

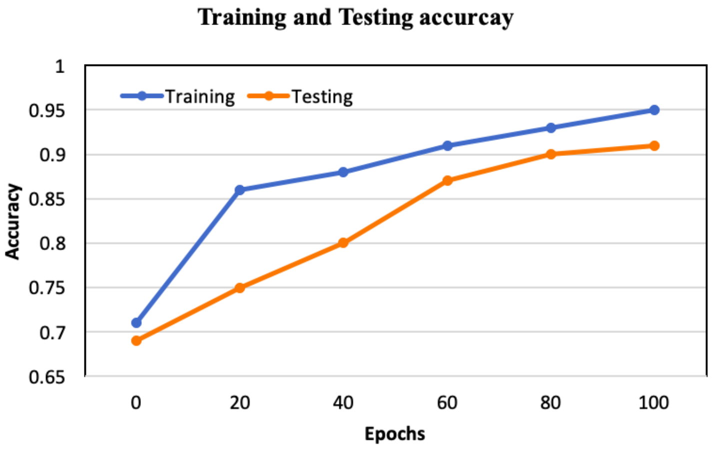

Fig. 9: Accuracy Vs. Epochs

Figure-8 illustrates the training and testing loss value of the input datasets. The simulation result of the proposed approach indicates that when the number of epochs increases, the loss rate of the trained data decreases. The simulation result demonstrates that the proposed method is more efficient in classifying the input data observed from three different datasets with less loss rate. Figure-9 depicts the training and testing accuracy value of the proposed three CNN models on detecting and classifying heart diseases from the input datasets. The experimental result of the proposed approach indicates that when the number of epochs increases, the accuracy value of the trained data also increases. This practical result clearly defines that all three proposed CNN-based disease detection techniques are more helpful in classifying heart diseases from the input BMI, ECG, and PBT datasets.

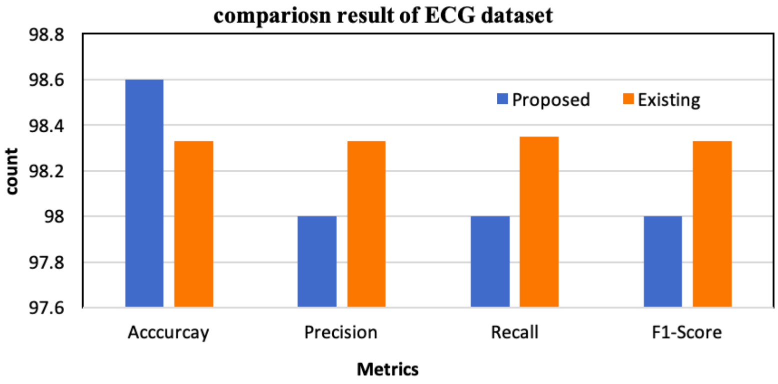

Fig. 10: Comparison result of the ECG datasets

Figure-10 illustrates the efficiency of the proposed and existing approach [11] in detecting ECG from the input electronic signal. It is achieved by evaluating the performance metrics values of the current and existing approaches' precision, recall, F1-score, and accuracy values. The result of the evaluation indicates that the proposed model performed well.

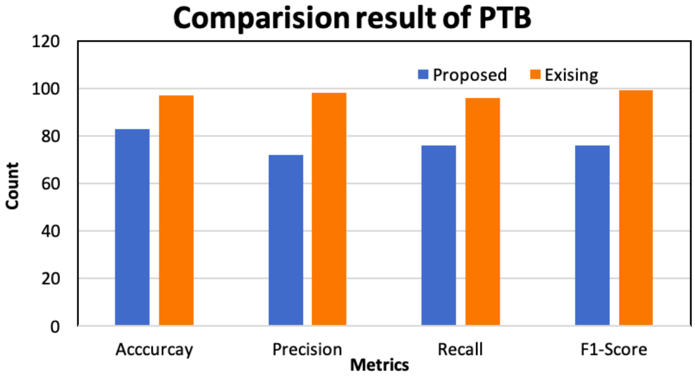

Fig. 11: Performance comparison of proposed and existing PBT [13] prediction

Table 5: Comparison of Performance Factors

α	Class	TPR	TNR	FPR	FNR	F1 Score
1	BMI	97.0%	96.1%	2.6%	1%	98.2%
2	ECG	100%	99.2%	1.5%	1%	95.59%
3	PBT	95.0%	96.96%	1.07%	3%	96.5%
Mean Accuracy	95.71%	95.71%	95.71%	1.65%	1.66%	98.33%

Similar to Figure 10, Figure11 indicates the performance of the current and existing [12,13] approaches to detecting heart diseases from the input BMI data and PBT physiological signals, respectively. The performance of the models is effectively analyzed using various performance metrics. The comparison result shows that the proposed three CNN-based approaches better accurately classified heart disease. Table 5 defines the different confusion metrics such as TP, TN, FP, FN, and f1-score values of the proposed three CNN models for predicting heart diseases from BMI, ECG, and PBT input data. The analysis shows that the proposed model classifies the input BMI datasets with 97.09%, 96.1%, 2.6%, 1%, and 98.2% TP, TN, FP, FN, and f1-score values, respectively. Likewise, the TP, TN, FP, FN, and f1-score values of the input ECG and PBT datasets are classified into (100%, 99.2%, 1.5%, 1%, 95.9%) and (95.09%, 96.65%, 1.07%, 3%, 96.5%) respectively.

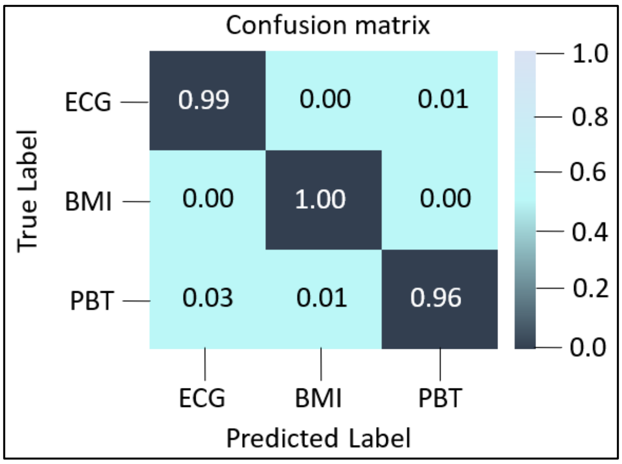

Fig. 12: Confusion metrics

Figure-12 defines the proposed model's overall confusion metrics value on classifying the input BMI, ECG, and PBT datasets. The result of this confusion metrics value illustrated that the proposed mode is more suitable for diagnosing the effects of normal and abnormal heart disease patients. To evaluate the performance of the proposed model explained here, the output in terms of accuracy is compared with the existing methods [18, 19, 20]. The comparison of accuracy and the related error value of the existing methods is given in Table-6. The comparison shows that the proposed hybrid CNN model provides more accuracy than the existing one.

Table 6: Comparison with Existing Methods

Method	Number of samples	Accuracy (%)	RMS error (%)
Ravichandran et al. [18]	53	–	12
Shi et al. [19]	105	99	
S.Misra et al. [20]	3200	97	0.034
Hybrid CNN	7000	99.87	0.13

5.0 CONCLUSION

This paper helps confirm the availability of heart disease in multi-modal data, such as BMI, ECG, and PTB datasets. Analyzing and predicting heart diseases from ECG datasets cannot provide high accuracy and confirmation of availability. Thus, this paper aimed to analyze multiple datasets, like ECG, BMI, and PTB, where heart diseases can also be predicted from these datasets. Three different CNN models with varying architectures of layers are used to predict and confirm heart disease present in a patient's body for analyzing different datasets. All three CNN models are implemented in Python software and experimented with different datasets. The performance of the proposed CNN models is evaluated by computing various performance factors, such as TPR, FPR, TPR, TNR, precision, recall, and accuracy. It is also compared with the existing methods, and the proposed CNN model outperforms the other models by obtaining 99.87% accuracy. Hence, it is concluded that the proposed hybrid CNN model is highly suitable for heart disease prediction on any dataset.

REFERENCE

[1] Atallah, Rahma, and Amjed Al-Mousa. "Heart disease detection using machine learning majority voting ensemble method." *2019 2nd international conference on new trends in computing sciences (ICTCS)*. IEEE, 2019.

[2] Alom, Zulfikar, et al. "Early stage detection of heart failure using machine learning techniques." *Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021*. Springer Singapore, 2022.

[3] Ed-Daoudy, Abderrahmane, and Khalil Maalmi. "Real-time machine learning for early detection of heart disease using big data approach." *2019 international conference on wireless technologies, embedded and intelligent systems (WITS)*. IEEE, 2019.

[4] Maji, Srabanti, and Srishti Arora. "Decision tree algorithms for prediction of heart disease." *Information and Communication Technology for Competitive Strategies: Proceedings of Third International Conference on ICTCS 2017*. Springer Singapore, 2019.

[5] Brunese, Luca, et al. "Deep learning for heart disease detection through cardiac sounds." *Procedia Computer Science*, Vol. 176, 2020, pp. 2202-2211.

[6] Helle, Emmi, and James R. Priest. "Maternal obesity and diabetes mellitus as risk factors for congenital heart disease in the offspring." *Journal of the American Heart Association*, Vol. 9, No. 8, 2020, pp. e011541.

[7] Devi, R. Lakshmi, and V. Kalaivani. "Machine learning and IoT-based cardiac arrhythmia diagnosis using statistical and dynamic features of ECG." *The journal of supercomputing*, Vol. 76, No. 9, 2020, pp. 6533-6544.

[8] Ahmed, Syed Thouheed, et al. "A generalized study on data mining and clustering algorithms." *New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India*, 2020, pp. 1121-1129.

[9] Khan, Mohammad Ayoub. "An IoT framework for heart disease prediction based on MDCNN classifier." *IEEE Access,* Vol. 8, 2020, pp. 34717-34727.

[10] Ng, Arnold CT, et al. "Diabesity: the combined burden of obesity and diabetes on heart disease and the role of imaging." *Nature Reviews Cardiology* Vol. 18, No. 4, 2021, pp. 291-304.

[11] Mehmood, Awais, et al. "Prediction of heart disease using deep convolutional neural networks." *Arabian Journal for Science and Engineering,* Vol. 46, No. 4, 2021, pp. 3409-3422.

[12] Ahmed, Syed Thouheed, et al. "Neural network based mental depression identification and sentiments classification technique from speech signals: A COVID-19 Focused Pandemic Study." *Frontiers in public health*, Vol. 9, 2021, pp. 781827.

[13] Khalil, Mohammed, and Abdellah Adib. "An end-to-end multi-level wavelet convolutional neural networks for heart diseases diagnosis." *Neurocomputing*, Vol. 417, 2020, pp. 187-201.

[14] Ahsan, Md Manjurul, and Zahed Siddique. "Machine learning-based heart disease diagnosis: A systematic literature review." *Artificial Intelligence in Medicine,* Vol. 128, 2022, pp. 102289.

[15] LK, Sowmya Sundari, et al. "COVID-19 outbreak based coronary heart diseases (CHD) prediction using SVM and risk factor validation." *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*. IEEE, 2021.

[16] Xie, Yilin, et al. "Computer-aided system for the detection of multicategory pulmonary tuberculosis in radiographs." *Journal of Healthcare Engineering,* 2020.