Optimizing Myocardial Infarction Prediction Performance Using Advanced Machine Learning Techniques

Shridevi K. Jamage[1,2,*], Ramesh Y. Mali[3], and Virendra V. Shete[1]
[1] MIT School of Engineering and Science (MIT SOES), MIT Art, Design and Technology University, Pune, India
[2] E&TC Department, SCTR's Pune Institute of Computer Technology, Pune, India
[3] Department of Electrical and Electronics Engineering, MIT School of Computing,
MIT Art, Design and Technology University, Pune, India
Email: shridevikjamage@gmail.com (S.K.J.), ramesh.mali@mituniversity.edu.in (R.Y.M.),
virendra.shete@mituniversity.edu.in (V.V.S.)

Manuscript received June 4, 2025; revised July 9, 2025; accepted August 18, 2025
*Corresponding author: shridevikjamage@gmail.com

***Abstract*—The rising prevalence of Myocardial Infarction (MI) and limited clinical resources highlight the need for accurate, automated diagnostic tools. This study presents a Machine Learning (ML) framework for early MI prediction using both structured health records and Electrocardiogram (ECG) data. Multiple ML algorithms—including ridge classifier, radius neighbor classifier, linear SVC, and extra trees classifier—are evaluated on two publicly available datasets and two clinical datasets collected from hospitals. The additional trees classifier achieves the highest training accuracy of 1.00, with consistent performance across datasets. For ECG-based diagnosis, a deep learning model combining Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) is developed using the ECG Heartbeat Categorization Dataset. It classifies five heartbeat types: normal, Fusion of Paced and Normal (FPAN), Fusion of Ventricular and Normal (FVAN), Atrial Premature Contractions (APC), and Premature Ventricular Contractions (PVC). The model achieves a testing accuracy of 0.98, supported by strong precision and recall across classes. The novelty of this study lies in its integration of public and real-world datasets, noise-augmented training to improve ECG robustness, and a multi-class CNN–RNN framework that enhances generalizability beyond conventional binary classifiers. The proposed approach contributes to more reliable and interpretable cardiovascular diagnostics, with strong potential for clinical deployment and improved patient outcomes.**

***Index Terms*—classification, Convolutional Neural Networks (CNN), electrocardiogram, heart disease, machine learning, myocardial infarction**

I. INTRODUCTION

Myocardial Infarction (MI), also recognized as a heart attack, is a condition caused by a disturbance in the blood supply to a segment of the heart [1]. Diagnosis is essential since MI carries a high mortality rate, especially in older age groups. The traditional methods of MI diagnosis through ECG (Electrocardiogram) signals require trained medical practitioners, prone to errors and bias in the interpretation process. Recent work has been in the direction of automation of MI diagnosis using techniques like Machine Learning (ML) and Deep Learning (DL). Traditional ML, as well as DL, addresses disease detection with their different approaches and benefits in heart disease detection [2]. ML is emerging as a revolutionary tool in medicine and, specifically, in forecasting, besides the analysis of cardiovascular ailments [3]. Through big volumes of health data utilizing complex ML algorithms, systems unravel relationships and intuitive patterns not easily found with traditional diagnostic methods. This, therefore, leads to the diseases' early recognition as well as forecast, thereby possibly improving patient results while reducing the scourge of illnesses on the world at large. The application of ML techniques in medicine highlights the promise these approaches carry toward solving complex health challenges [4]. For MI, specifically, ML enables actionable insights that could lead to early and beneficial diagnosis and treatment, which, therefore, translates into an added benefit toward improving public health. With these benefits, there has been a growing urgency to develop ML-based systems both on-site and through the Internet of Medical Things (IoMT) [5] for improved prediction and diagnosis of MI. Coupled with comprehensive medical data, ML-driven systems might embed technology and change the way that MI is managed to bring about improvements in healthcare outcomes. The process of ML algorithms starts with manual feature extraction from ECG signals that may be biased and labour-intensive and need domain-specific knowledge demanding high-dimensional instances, due to being linked with human judgment. Moreover, it requires substantial computational resources to calculate, manage, and assess many features, which makes it difficult for large-scale applications. Deep learning [6], however, it overcomes most of these challenges by learning how to automatically select features directly from raw ECG data, with no requirement for handcrafted feature engineering. CNNs (Convolutional Neural Networks) and RNNs (Recurrent Neural Networks) are some DL models that learn relevant features from big datasets through filter weights evolution during training. This reduces the chance of human bias, relies less on domain expertise, and generally consumes less computational power in extracting features.

Recent studies employed ML approaches to utilize data Electronic Health Record (EHR) to predict the risk of acute MI and subsequent mortality [7]. These approaches hold great promise but have limitations in their adaptation for clinical decision-making. Most of the ML models are complex to understand, interpret, and validate the model's predictions [8].

This study advances the existing literature by introducing a robust, hybrid CNN-RNN model for ECG signal classification, capable of capturing both spatial and temporal features of cardiac rhythms. Unlike many previous studies, which use only CNN or conventional ML models, our architecture leverages the strengths of CNNs in feature extraction and RNNs in temporal pattern recognition, offering improved classification performance across five distinct cardiac categories.

Furthermore, although the ML models are risk predictors, but may not resemble actual probabilities. For instance, a model may give a 90% mortality risk, yet the actual risk might have been quite different and would challenge experts to believe in and apply such a tool in clinical practice. Thus, the study aims to compare the performance of the ML model and improve the accuracy of prediction. The model is validated on different datasets, including patient health records. Furthermore, the study aims to investigate ECG data for MI prediction with improved accuracy.

The study has a significant contribution to the state-of-the-art in both ML and cardiovascular diagnostics, with an imperative contribution of a noise-augmented approach to the processing of ECG data. The study improves signal robustness, enhancing classification accuracy on different datasets. In addition, the integration of CNN and RNN architectures classifies the ECG signal with the temporal history of cardiac events generated regarding critical features for heart monitoring. Besides that, standard metrics with detailed analysis provide validation of the efficacy of the proposed model, highlighting its application in real clinical environments for constant cardiac health monitoring as well as timely diagnosis. Unlike many prior studies that rely on a single dataset or binary classification, this work advances the field by combining multiple structured datasets (both public and hospital-collected), applying noise-augmented training to enhance robustness, and employing a CNN–RNN hybrid architecture for detailed multi-class ECG classification. These elements collectively contribute to improving both the generalizability and clinical applicability of ML-based cardiac diagnosis.

The objectives of the study are

- Develop an ML-based model to predict heart disease using patient health data (EHR) in addition to a CNN-based model to categorize ECG signals in five cardiac categories.
- Evaluate the model's classification performance using accuracy, precision, F1 score, and recall metrics.
- Compare the outcomes of different ML algorithms on two different heart disease prediction datasets.
- Facilitate automation in cardiovascular diagnostics by creating a scalable model for enhanced heart disease detection.

II. Related Work

During the past decade, cardiovascular diseases, including heart disease and myocardial infarction, have attracted much attention from medical researchers and healthcare professionals as their impact on human health has grown increasingly greater. Meanwhile, ML techniques have been advanced as effective tools for the accurate estimation of the incidence of heart disease, despite the huge EHR data used. Some studies utilized various ML algorithms to predict heart disease with varying strengths in terms of accuracy. The research by Sudha and Kumar [9] contributed to advancements in healthcare by suggesting a hybrid CNN-LSTM (Long Short-Term Memory) framework for diagnosing heart disease. The method utilized CNN's robust feature extraction capabilities and LSTM's ability to process sequential data, which was key for time-series medical records. This combination improved classification performance and proved deep learning's potential in medical diagnosis.

Miah *et al.* [10] investigated the prediction of myocardial sickness, an important challenge for cardiovascular medicine. Through comparisons of six machine learning algorithms—logistic regression, Support Vector Machine (SVM), decision tree, bagging, XGBoost (eXtreme gradient boosting), and LightGBM (light gradient boosting machine), they sought to determine the best predictive method. Of these models, XGBoost's highest accuracy was 92.72%, which underlines its best performance. The research focused on highlighting the contribution of sophisticated machine learning methods towards better early diagnosis and enhancing proactive medical approaches to cardiovascular disease.

In further research [11] authors considered the classification of Cardiovascular Disease (CVD) employing different supervised machine learning models, which were experimented with the Sani Z-Alizadeh dataset of the UCI (University of California, Irvine) repository. The SMOTE (Synthetic Minority Oversampling Technique) was utilized owing to the class imbalance of the dataset, along with assessment using ten-fold cross-validation. The classifiers' performance was compared for Multilayer Perceptron (MLP), SVM, Random Forest (RF), Logistic Regression (LR), and Decision Tree (DT), where RF was the most accurate among all of the classifiers, while MLP had the highest precision for the resampled data. The work proved that these models are able to support early diagnosis of CVD and aid in better clinical decision-making.

In another study [12], the author added the advanced ML methods for the forecast of heart disease, viz., gradient boosting, logistic regression, as well as SVM algorithms, in addition to a grouping method using a voting classifier on the Cleveland heart disease dataset. The authors concluded that the ensemble voting classifier model, which utilized chosen critical attributes from the dataset, achieved a staggering accuracy of 97.9%. This mixed model not only exhibited high accuracy but also performed better compared to single algorithms in accuracy and precision.

In a heart disease prediction study [13], authors Ogundepo and Yahya used the Cleveland and Statlog datasets, where there are multiple ML algorithms, like decision tree, logistic regression, and random forest, with KNN (K-Nearest Neighbours), ANN (Artificial Neural Network), SVM, naive Bayes, extreme gradient tree, linear discriminant analysis, and conditional random forests. The experiments showed that SVM produced worthy predictive results with the highest accuracy of 85%. Hybrid models have also shown improvements in prediction accuracy. They combine multiple ML algorithms into a hybrid model [14], which achieves maximum accuracy, but more experimentation is needed to understand the generalization of hybrid models across different tasks and different datasets.

Finally, Natarajan *et al.* [15] analyzed the data for recognizing prominent features in the dataset for heart disease and improving prediction accuracy using ensemble methods. They used stacking and voting approaches on the Z-Alizadeh Sani dataset of the UCI machine learning repository. The study demonstrated that stacking with the features selected by the firefly algorithm yielded the highest accuracy of 86.79%. The findings indicate that ensemble learning and metaheuristic feature selection can enhance heart disease prediction models.

This comparison underlines the importance of choosing an appropriate data mining tool and ML algorithm that is apt for the specific purpose of duty at hand. For instance, predicting cardiovascular diseases. These studies generally illustrate the diversity of the various ML techniques applicable to heart disease prediction. Continuous improvement of these models presents a great prospect for improving early detection and resultant outcomes in cardiovascular health.

Furthermore, the existing studies on ECG-based detection show various innovative tactics for the identification of cardiac anomalies. Acharya *et al.* [16] have recommended a model using the basis of contourlet and shearlet transforms on scalograms for entropy and statistical feature extraction using DT (Decision Tree) and KNN classifiers. The studies [17, 18] on MI collectively contributes towards the establishment of a robust tool to recognize several cardiac abnormalities concurrently using 12-lead ECG data that could be very easily accommodated in smart wearables, with energy efficiency [19]. In early related research work, authors Raghukumar and Naveen [20] achieved 88.79% using a gradient boosting classifier with the extracted features from the ECG signal. It enhances diagnosis accuracy as well as efficiency in terms of heart attack prediction. Lee *et al.* [21] obtained a remarkable 97.8 % accuracy with a CNN-based predictive framework using the MITDB dataset, including a 2D BSM representation (two-dimensional beat-score-map) from the ECG input. Feng *et al.* [22] have further shown that the DL techniques are viable by using a one-dimensional UNet model for RR interval segmentation with an accuracy of 97% and a sensitivity of 95.5%.

Despite these advancements, several limitations remain. Many of the existing approaches focus on binary classification or use standard feature extraction methods, which can overlook complex spatial and temporal patterns in ECG signals. A significant number of studies also rely on single, benchmark datasets such as Cleveland or Z-Alizadeh Sani, limiting their generalizability across diverse populations and real-world scenarios. Moreover, the black-box nature of many high-performing models reduces clinical interpretability, which is critical for adoption in healthcare settings. Lastly, relatively few models explore robust performance under noisy or imbalanced conditions, nor do they adequately address the need for multi-class ECG classification that reflects the complexity of cardiac abnormalities.

The present study addresses these gaps by evaluating ML models on multiple EHR-based datasets and introducing a hybrid CNN-RNN architecture for classifying ECG signals across five distinct cardiac classes. This integrated approach improves generalizability, robustness, and diagnostic depth while supporting real-world applicability in automated cardiovascular monitoring systems.

III. Methodology

A. ML-Based Prediction

The study focuses on utilizing ML algorithms to predict heart disease from patient Electronic Health Records (EHR) with enhanced performance.

It uses two datasets to compare different ML algorithms like the ridge classifier, radius neighbors classifier, linear SVC (Support Vector Classifier), and extra trees classifier. The dataset is divided into 70% for training in addition to 30% for testing, following data pre-processing and feature selection. ML algorithms are applied to classify the data, indicating the presence or absence of Myocardial Infarction (MI). These algorithms are then assessed on the test set to assess their prediction performance. Both datasets are tested on the proposed ML model, and the algorithms' predictive abilities are compared. In addition, we have utilized two datasets collected from the hospitals for validation. Fig. 1 displays the design of the anticipated ML system.

1) Data collection

Two datasets available online are utilized in this study, with two datasets collected from the hospitals for validation.

Dataset 1 [23]: The dataset of heart disease prediction consists of 324 entries with 14 variables representing both medical and demographic details about patients. Significant features in the dataset are gender, age, chest pain type, cholesterol levels, resting blood pressure, and maximum heart rate. Other variables cover fasting blood sugar, exercise-induced angina, and thalassemia status, among others. The dataset includes the occurrence of heart disease, with over half of the entries showing a positive diagnosis.

Dataset 2: This heart disease prediction dataset [24], derived from a cardiovascular study conducted in Framingham, Massachusetts, containing 4,240 entries and

16 attributes covering demographic, behavioural, and medical factors. The primary objective is to predict each patient's ten-year risk of CHD (Coronary Heart Disease), identifying individuals at a heightened risk. The dataset includes a combination of demographic details such as gender and age, behavioural factors like smoking status and average daily cigarette use, and medical history indicators including prior strokes, hypertension, diabetes, and blood pressure medication use. Current medical data features measurements for total cholesterol, diastolic as well and systolic blood pressure, Body Mass Index (BMI), glucose levels, and heart rate. The dataset's target variable reflects the presence or absence of CHD risk within a ten-year timeframe. This comprehensive dataset, accessible on Kaggle, provides essential insights into various factors contributing to heart disease risk.

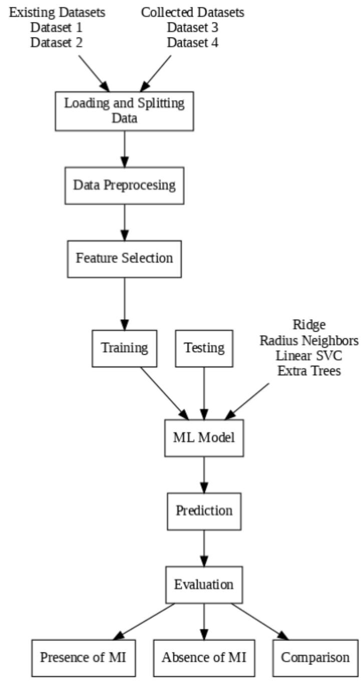


Fig. 1. Proposed ML-based MI prediction model.

Dataset 3: This dataset was collected by the authors from Bharati Vidyapeeth (Deemed to be University), Medical College and Hospital, Sangli – Miraj Road. It contains 1,012 rows and 12 columns and approximately occupies 95 KB of memory. It includes demographic, health, and lifestyle information, mainly on factors associated with cardiovascular health. The dataset consists of both integer and categorical data types and covers a range of attributes that are relevant to the analysis. The Age column records the ages of individuals, ranging from 20 to 79 years, with a mean of 50.58 years and a standard deviation of 17.29 years. The rest of the columns are categorical, with attributes like gender (male, female), diabetes, hypertension, and biomarkers like serum creatinine and LDL cholesterol labelled as either "Positive/Negative" or "High/Normal". Lifestyle-related variables include BMI, smoking duration, smoking frequency, and tobacco use, which capture smoking habits and physical health. Moreover, alcohol consumption is classified into "None", "Low", "Moderate", and "High" levels, while the target column, myocardial infarction, shows whether a patient has had a heart attack. One of the notable features of this dataset is that some categories are imbalanced, such as a preponderance of "Positive" cases for diabetes. Most of the categorical features are binary or have a small number of unique values, which makes them ideal for classification problems. This dataset provides a good mix of demographic, medical, and lifestyle features to study the risk factors for cardiovascular diseases and predict the probability of myocardial infarction.

Dataset 4: This dataset was collected from Aryan Heart Care, Miraj, Maharashtra, under institutional approval. It consists of 1,372 rows and 17 columns and uses approximately 182 KB of memory. It includes both numerical and categorical variables, all of which are intended to offer an elaborate view of demographics, health conditions, and cardiovascular-related metrics. The age column, which is of integer type, ranges between 20 and 79 years with a mean of 51.15 years and a standard deviation of 17.17. Other variables, like gender (male or female), diabetes, and hypertension, are binary and indicate the presence of conditions. Additional clinical indicators include serum creatinine and LDL cholesterol. BMI is categorized as "High" or "Normal", while smoking-related features include smoking duration, tobacco use, etc. Other cardiovascular indicators include resting blood pressure, ECG categories, maximum heart rate (ranging from 60 to 200 bpm, average 128.82, SD 40.24), exercise-induced angina (oldpeak), slope of the ST segment, number of major vessels, and thalassemia status. The target variable indicates myocardial infarction, with balanced classes ("Yes" and "No"). However, some features, such as diabetes, are heavily imbalanced (1,313 out of 1,372 cases are "Positive"). This dataset provides a rich mix of health, lifestyle, and clinical features, making it suitable for predictive modeling and cardiovascular disease analysis.

2) Data pre-processing and feature selection

To prepare and optimize datasets for heart disease prediction, data pre-processing and feature selection are conducted to ensure that the models can learn from the most relevant and clean data. This improves the models' ability to classify heart disease risk effectively, focusing learning on the most informative features. Descriptive statistics are generated to summarize the dataset with exploratory data analysis, showing the shape and identifying missing values. The distribution of the target variable, heart disease presence or absence, is visualized. For pre-processing, label encoding converts the target variable's categorical format to numerical values suitable for model input. Additionally, an interaction term between 'age' and 'thalach' (maximum heart rate achieved) is created to capture any combined effect on heart disease prediction. The 'exang' (exercise-induced angina) column, which is deemed less significant, is dropped to reduce noise and improve efficiency. In feature selection, cross-validation is employed to identify and retain the most important features for the predictive models. This process iteratively removes less significant features while cross-validating performance, allowing for the selection of the most impactful attributes that will enhance the accuracy of classifier models. While no explicit class balancing method, such as SMOTE, was applied to the EHR datasets, care was taken during model evaluation to monitor metrics like recall and F1-score that reflect the model's ability to handle potential class imbalance. Future work may explore resampling techniques or cost-sensitive learning to further improve classification performance for underrepresented classes.

3) ML algorithms

To predict heart disease using EHR, several ML algorithms are utilized, each bringing unique strengths.

- Ridge classifier, a linear model with L2 regularization, effectively manages high-dimensional data by penalizing large coefficients to reduce overfitting, which can be beneficial in handling complex medical datasets.
- Radius neighbors classifier is a non-parametric technique that assigns class labels according to the majority class within a specified radius, helping to capture local data patterns that may indicate disease.
- Linear SVC maximizes the margin between data points by identifying the optimal hyperplane, providing robust classification for distinguishing heart disease risk groups.
- The extra trees classifier is an ensemble method to construct manifold randomized decision trees and average their outputs, enhancing prediction accuracy and controlling overfitting, making it especially suitable for datasets with diverse features in EHRs.

B. ECG Data Classification

Furthermore, this study implements a CNN-RNN model for classifying ECG signals utilizing the ECG Heartbeat Categorization dataset. It is trained on ECG signals representing normal heartbeats, as well as signals associated with myocardial infarction and arrhythmias. These ECG signals are segmented after pre-processing, where each segment captures a single heartbeat. Different heartbeat types are included in the dataset—Normal, Fusion of Paced and Normal (FPAN), Fusion of Ventricular and Normal (FVAN), Premature Ventricular Contractions (PVC), as well as Atrial Premature Contractions (APC). To illustrate the distinctions among heartbeat types, the study plots one ECG sample per class, highlighting the unique signal characteristics of each. Gaussian noise is added to the signals as a data augmentation technique, creating augmented ECG signals that balance the dataset by dropping the prevalence of the majority class (Normal) to counterpart the sample count of minority classes.

The model structure (Fig. 2) combines convolutional and recurrent layers, specifically tailored for time-series data processing like ECG signals. The proposed CNN-RNN model uses an Adam optimizer, including definite cross-entropy as the loss function. Also, the overfitting is bypassed using early stopping and built-in validation loss, which saves the top-performing model. Following training on the balanced data, the performance is evaluated in terms of a confusion matrix and results of classification on test data, including F1 scores, recall, and precision for these classes. The evaluation framework delivers an ample assessment of the model's accuracy and robustness in categorizing different ECG signal types.

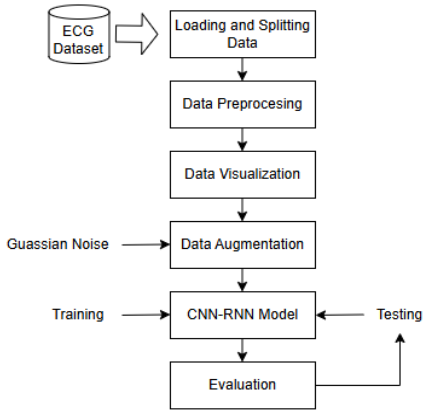


Fig. 2. Proposed ECG classification model.

a) Data collection

The study applies a dataset that integrates heartbeat signals from the widely utilized ECG heartbeat categorization dataset [25]. The significant number of samples robustly establishes a deep CNN model trained over the dataset for the classification of heartbeats with potential classification capabilities, as well as the effects of employing ECG data. Each signal is the shape of a heartbeat in the ECG, capturing both the normal heartbeat and those relating to arrhythmias or myocardial infarction. It ensures that every segment relates to one heartbeat by pre-processing and segmenting the signals. The entire dataset is then separated into training and testing sets, including 70,043 samples and 17,511 samples, respectively. The proposed architecture ensures that training and testing of the phases are evenly balanced to maximize the generalization ability and accuracy of classification for the heartbeat types.

b) Data pre-processing

Basic Exploratory Data Analysis (EDA) in this study comprises exploring the distribution of heartbeat class labels in the training data. The class frequencies reveal an imbalance where certain classes have a higher number of samples, like the Normal class, than other classes. This imbalance can potentially bias the model's predictions, potentially skewing them toward the majority class. Thus, pre-processing includes data augmentation and downsampling to achieve a balanced dataset and improve model performance across all classes. This dual strategy ensured that the classifier did not become biased toward the overrepresented "Normal" class and was able to learn meaningful patterns across all heartbeat types equally, leading to improved recall in minority classes.

c) Data visualization

Data visualization in this study includes drawing plots that show one ECG sample for every class of the database, with signal patterns of each heartbeat type. Five different plots were made for the classes involved, which are Normal, FPAN, PVC, APC, and FVAN. These visualizations (Fig. 3) are pretty informative and clearly show distinctions between classes in the ECG signal, which reflect the intrinsic characteristics of each class; they help build models and interpret classification results.

d) Data augmentation

Gaussian noise enhances ECG data, in addition to the robustness of DL models to the dataset. It includes the mean as well as standard deviation set to 0 and 0.01, which control the intensity of the noise. Gaussian noise is created over the original ECG signal data, with values sampled from a given Gaussian distribution defined with mean and std. Therefore, the same shape for the original data is ensured, and added noise may not drastically alter the ECG signal. It is essential to train a model for real-world perturbations and inaccuracies in the original ECG. The model is then trained using the augmented ECG signal, including both the original signal and added noise. This will help the model learn to recognize patterns even in slight signal disturbances and play a key part in data augmentation and augmenting the variability of the dataset, thus improving the ability of the model to classify unseen and new data, as shown in Fig. 4.

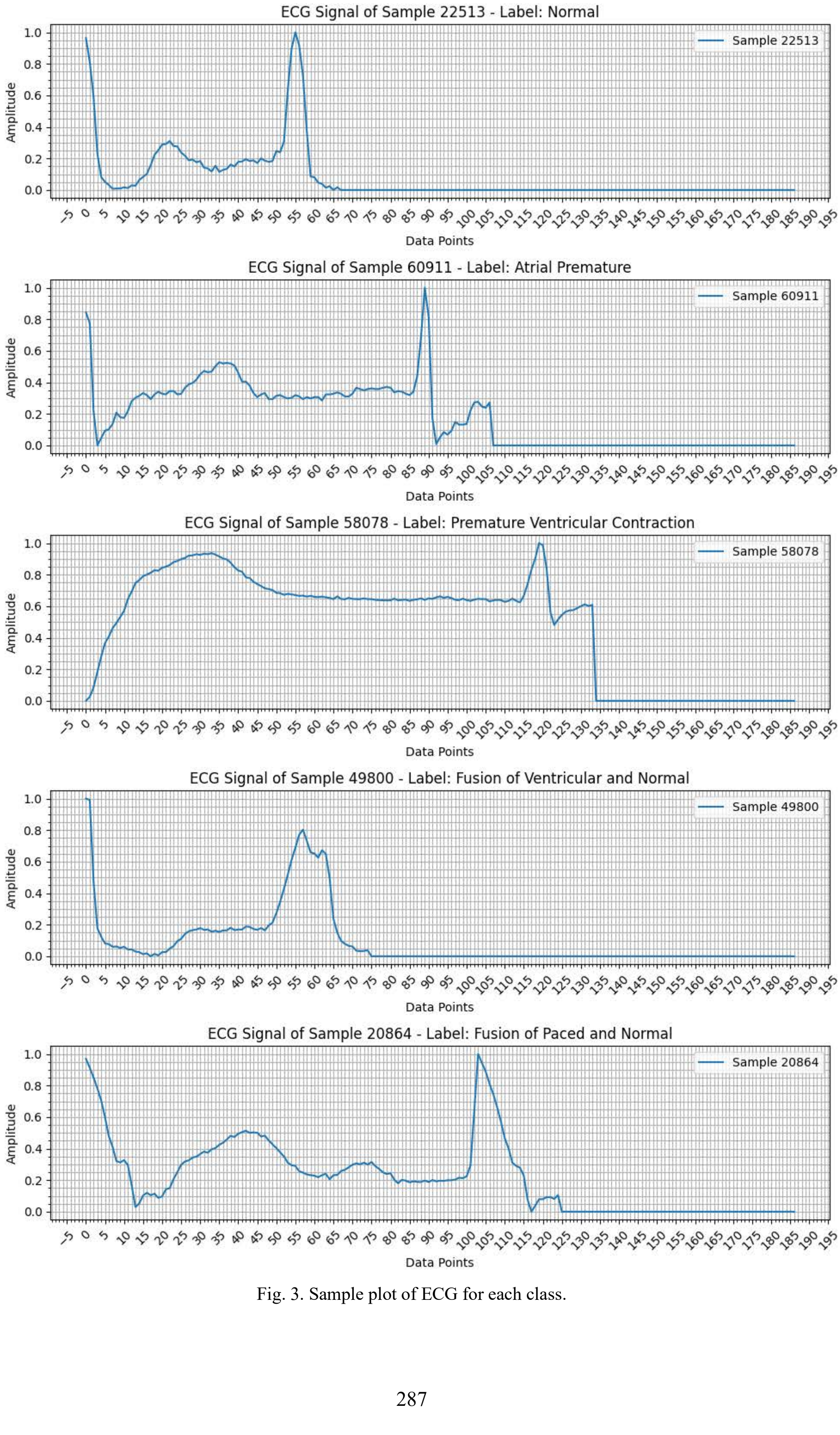


Fig. 3. Sample plot of ECG for each class.

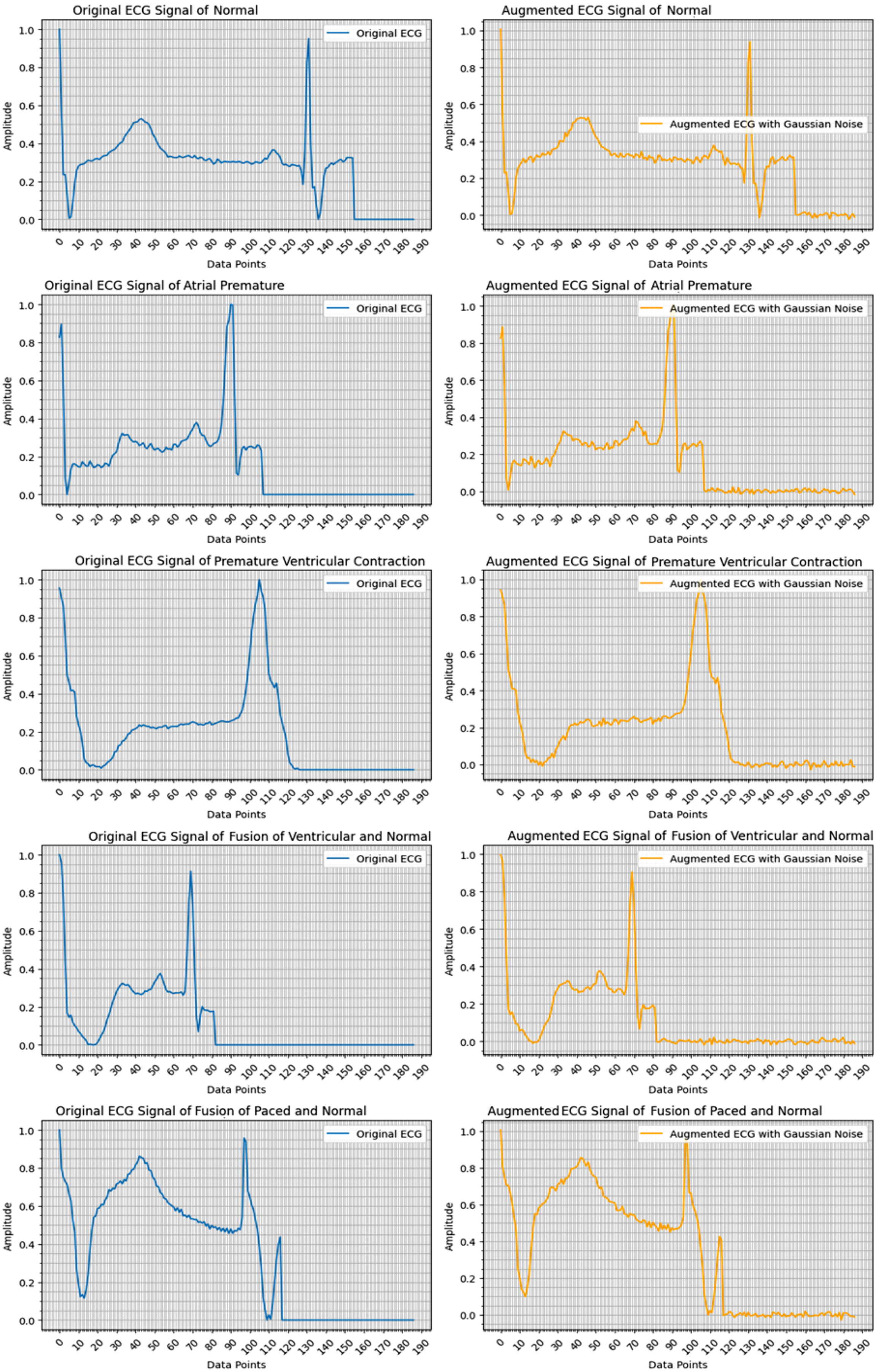


Fig. 4. Original and augmented ECG using Gaussian noise for each class.

TABLE I: SAMPLES IN EACH CLASS AFTER BALANCING

Class	Samples	Downsampling
Normal	57977	8128
FPAN	5145	8128
PVC	4630	8128
APC	1778	8128
FVAN	513	8128

Furthermore, down-sampling balances the dataset by the majority class (Normal) to equal the quantity of samples in the minority classes. The new class distribution balances the dataset with an identical amount of 8128 samples for every class, as shown in Table I below. The labels are one-hot encoded again after balancing.

e) Training and testing

This model is trained with an augmented ECG signal dataset for the prediction of five different classes, where CNN is combined with RNN. The training of the model uses a sequential pattern to categorize the ECG signals into five classes. This is initiated by a one-dimensional convolutional layer with a kernel size of 3, 32 filters, and an activation function like ReLU (rectified linear units). The last layer is the final dense layer using softmax activation for creating class predictions. The CNN extracts feature maps through several convolutional layers, automatically capturing the key characteristics of signals by using filters. The obtained feature maps are then entered into the RNN layers that model temporal dependencies and analyse sequential data to detect any pattern over time, from changes in heartbeat. This dual approach effectively extracts spatial and temporal information latent within the ECG signal with a more accurate prediction of cardiac anomalies. The Adam optimizer in the proposed model is used with categorical cross-entropy loss to accurately classify heart conditions from ECG signals, achieving better accuracy and AUC (area under the curve).

The CNN-RNN model was trained for 25 epochs using a batch size of 32 and the Adam optimizer with a learning rate of 0.001. Early stopping was applied with a patience of 5 epochs by monitoring validation loss to prevent overfitting. The ECG dataset was divided into 70,043 training samples, 17,511 validation samples, and 21,892 testing samples, each with 188-time steps. The model architecture includes a 1D convolutional layer (Conv1D) with 64 filters and a kernel size of 3, followed by a dropout layer with a rate of 0.5. This is followed by two GRU layers with 64 and 32 units, respectively, to capture temporal dependencies in heartbeat sequences. The output is flattened and passed through a dense layer with 64 units using ReLU activation, followed by another dropout layer. The final dense layer uses softmax activation to predict one of five heartbeat classes: Normal, Atrial Premature, Premature Ventricular Contraction (PVC), Fusion of Ventricular and Normal, and Fusion of Paced and Normal. This architecture enables the model to learn both spatial and sequential patterns in ECG signals, improving classification accuracy and robustness across multiple cardiac conditions.

f) Evaluation metrics

The proposed model is gauged utilizing various metrics, accounting for correctly predicted positive (TP) and negative (TN) classes, as well as incorrect predictions of these classes, including false positives and negatives (FP, FN).

Accuracy: The model's accuracy calculates the number of correct results, including true positives and negatives, out of the overall cases analyzed. It is calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Precision: The ratio of appropriately predicted positives to the total number of positive predictions is referred to as precision.

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

Recall: It measures the capability of the model to discover all the pertinent cases. The proportion of correct positive estimates to the entire actual observations is acknowledged as recall.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F-score: It is a single metric that measures the harmonic mean of Recall and Precision, balancing them.

$$F - Score = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \tag{4}$$

IV. RESULTS AND DISCUSSION

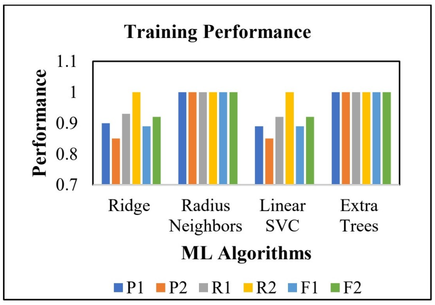
A. Performance of ML Algorithms on Existing Datasets

Different ML algorithms, including the ridge classifier, radius neighbors classifier, linear SVC, and extra trees classifier, were applied to Dataset 1 and Dataset 2. The Performance assessed during both the training and testing phases of the prediction module gives promising results. The proposed model demonstrated significant performance gains, particularly with the radius neighbors classifier and extra trees classifier, achieving the highest precision values (P_1, P_2) of 1.00, identical to the F-score (F_1, F_2) and recall (R_1, R_2) for both datasets. Meanwhile, the ridge classifier slightly outperformed linear SVC, showing marginally higher precision and recall on Dataset 1. Table II indicates evaluation metrics for Dataset 1 (P_1, R_1, F_1) and Dataset 2 (P_2, R_2, F_2).

The results indicate that the training phase for Dataset 1 yields higher precision (P_1) values compared to Dataset 2, suggesting that Dataset 1 has fewer false positives. Conversely, Dataset 2 performs better in terms of recall (R_2) and F-score (F_2) than Dataset 1, indicating that it captures positive instances more effectively and balances precision and recall better.

TABLE II: COMPARISON OF PERFORMANCE OF TWO DATASETS

Model	Training Result						Testing Result					
	Dataset 1			Dataset 2			Dataset 1			Dataset 2		
	P_1	R_1	F_1	P_2	R_2	F_2	P_1	R_1	F_1	P_2	R_2	F_2
Ridge Classifier	0.90	0.93	0.89	0.85	1.00	0.92	0.87	0.91	0.88	0.84	1.00	0.91
Radius Neighbors Classifier	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.04	0.07	1.00	0.04	0.07
Linear SVC	0.89	0.92	0.89	0.85	1.00	0.92	0.87	0.91	0.88	1.00	1.00	0.91
Extra Trees Classifier	1.00	1.00	1.00	1.00	1.00	1.00	0.86	0.91	0.87	0.85	1.00	0.92



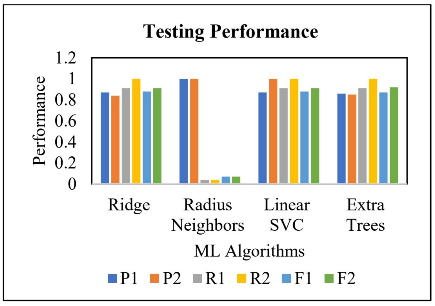


Fig. 5. Comparing training and testing results.

The Fig. 5 shows graphical representations of the results obtained.

In contrast, the testing results highlight the performance differences among the models on each dataset, where classifiers demonstrate strengths in precision, recall, or F-score on the datasets. The balanced performance of the ridge classifier gives relatively high precision and recall values on both datasets. On Dataset 2, it attained an R_2 of 1.00 with all the positive instances captured and achieved a higher F_2. Also, the radius neighbors classifier has an extremely high precision of 1.00, meaning it does not give many false positives, but has a low recall of 0.04, making its F-score very low in both datasets. The linear SVC model is equally good as the ridge classifier on Dataset 1. For Dataset 2, it achieves P_2 and R_2 both equal to 1.00 for all correctly classified instances. It, therefore, achieves a high F_2 and makes this classifier an extremely good candidate for Dataset 2. The extra trees classifier is well consistent on both datasets, and it achieves a high R_2 of 1.00 on Dataset 2, and F_2 is quite close to those of the ridge classifier and linear SVC, so they are good choices for both data sets.

The ridge classifier and extra trees classifier perform strongly and in balance across both datasets. The linear SVC performs exceptionally well on Dataset 2, scoring perfectly on all of them. Meanwhile, the radius neighbors classifier shows a high precision but very poor recall. Thus, it is not as useful for this purpose as it captures very few positive cases, as reflected in the low F_1 scores. This analysis suggests that the linear SVC and extra trees classifier might be the most reliable options, especially when both precision and recall are highly regarded in applications. In clinical scenarios such as myocardial infarction detection, recall is particularly critical, as it reflects the model's ability to correctly identify true positive cases. A low recall, such as that observed with the Radius Neighbors Classifier, implies that many actual MI cases could be missed. This could lead to underdiagnosis and delayed treatment, posing serious risks to patient health. Thus, despite its high precision, a model with low recall may not be suitable for real-world clinical deployment, where missing positive cases can be more harmful than false positives.

Table III shows the highest accuracy of 1 for the extra tree classifier in training on both datasets, while linear SVC slightly enhanced its performance (0.8661, 0.8533, 0.866, 0.8427) than the ridge classifier (0.8661, 0.8526, 0.8557, 0.84) during training as well as testing. On the other hand, the radius neighbors classifier is better in training performance (0.8661, 0.8526) while lacking in testing performance (0.0206) on dataset 1. However, its testing performance for dataset 2 is efficient (0.84), as ridge classifier (0.84). Overall, the linear SVC shows the most consistent performance, indicating reliability without significant overfitting and stability across both datasets. It may generalize better and is likely the optimal model for predicting heart ailments.

TABLE III: ACCURACY OF PROPOSED MODEL

Models	Train Acc.		Test Acc.	
	Dataset 1	Dataset 2	Dataset 1	Dataset 2
Ridge Classifier	0.8661	0.8526	0.8557	0.84
Radius Neighbors Classifier	1	0.9993	0.0206	0.84
Linear SVC	0.8661	0.8533	0.866	0.8427
Extra Tree Classifier	1	1	0.8454	0.8453

B. Datasets 3 and 4 (Collected from the Hospitals)

Table IV reports the training and testing performance of four classifiers, namely the ridge classifier, radius neighbors classifier, linear SVC, and extra trees classifier, on two datasets. The performance is reported in terms of precision (P), recall (R), and F1-score (F), and thus the reader gets an idea about the performance of each model on both datasets.

The ridge classifier shows consistent performance on both datasets, with only minor fluctuations in the training and testing results. For Dataset 3, P_1, R_1, and F_1 are at 94%, 96%, and 95% in training, respectively; meanwhile, the testing metrics were lower, with precision at 96%, recall at 92%, and F1-score at 94%. On Dataset 4, the training results are similar, having P_2 = 96%, R_2 = 93%, and F_2 = 94%; the testing results show similar high accuracy with P_2 = 92%, R_2 = 95%, and F_2 = 93%. This classifier is both stable and reliable, showing similar generalizations across different datasets.

TABLE IV: RESULTS OF THE CLASSIFIER OF THE 2 DATASETS COLLECTED FROM THE HOSPITALS

Model	Training Result						Testing Result					
	Dataset 3			Dataset 4			Dataset 3			Dataset 4		
	P_1	R_1	F_1	P_2	R_2	F_2	P_1	R_1	F_1	P_2	R_2	F_2
Ridge Classifier	0.94	0.96	0.95	0.96	0.93	0.94	0.96	0.92	0.94	0.92	0.95	0.93
Radius Neighbors Classifier	0.8	0.95	0.87	0.92	1	0.96	0.93	0.71	0.8	1	0.9	0.95
Linear SVC	0.95	0.95	0.95	0.97	0.96	0.96	0.94	0.94	0.94	0.95	0.96	0.96
Extra Trees Classifier	0.95	0.94	0.94	0.99	0.98	0.99	0.93	0.95	0.94	0.98	0.99	0.98

The radius neighbors classifier shows more variability between datasets and between the training and testing phases. For Dataset 3, the training metrics are very strong: P_1 = 80%, R_1 = 95%, and F_1 = 87%, and so are the testing metrics: P_1 = 93%, R_1 = 71%, F_1 = 80%. For Dataset 4, the training results are significantly improved (P_2 = 92%, R_2 = 100%, F_2 = 96%), indicating that it can fully fit the data. Testing results for Dataset 4 are still high (P_2 = 100%, R_2 = 90%, F_2 = 95%), but there is a significant drop in recall from the training phase, indicating some overfitting.

The linear SVC has a balanced performance with minimal variation between the training and testing results on both datasets. For Dataset 3, the training metrics are P_1 = 95%, R_1 = 95%, and F_1 = 95%, whereas the testing metrics are P_1 = 94%, R_1 = 94%, and F_1 = 94%. On Dataset 4, training precision, recall, and F1-score (P_2 = 97%, R_2 = 96%, F_2 = 96%) are similar to the testing metrics (P_2 = 95%, R_2 = 96%, F_2 = 96%). This suggests that linear SVC generalizes well and does not overfit compared to other models.

The extra trees classifier achieves the highest metrics across both datasets, particularly for Dataset 4, where training metrics are near-perfect (P_2 = 99%, R_2 = 98%, F_2 = 99%). For Dataset 3, the training results are also impressive (P_1 = 95%, R_1 = 94%, F_1 = 94%), with strong testing performance (P_1 = 93%, R_1 = 95%, F_1 = 94%). The testing results on Dataset 4 are exceptional (P_2 = 98%, R_2 = 99%, F_2 = 98%), reflecting the classifier's ability to effectively generalize across datasets while maintaining high accuracy and reliability.

Fig. 6 illustrates the training and testing performance comparison of the classifiers. Here, the stability and strong performance of the extra trees classifier and linear SVC are visible, as well as variability in the recall of the radius neighbors classifier. The results of the ridge classifier remain stable and hence are reliable for applications requiring balanced metrics. This analysis underlines the need for classifiers that classify well during training and generalize as well as possible to unseen testing data across a range of datasets.

Table V is a comparison of the performance of four different classifiers: Ridge classifier, radius neighbors classifier, linear SVC, and Extra Trees classifier, on two different datasets. Key metrics include training accuracy, testing accuracy, training time, and memory usage for each model.

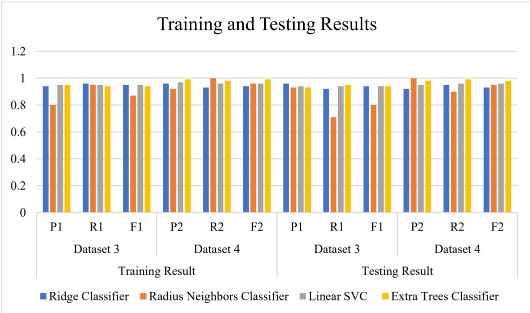


Fig. 6. Training and testing results of classifiers.

TABLE V: COMPARISON OF ACCURACY OF DATASETS 3 AND 4

Model	Dataset 3		Dataset 4	
	Train Accuracy	Test Accuracy	Train Accuracy	Test Accuracy
Ridge Classifier	0.9456	0.9458	0.9389	0.9382
Radius Neighbors	0.9852	0.8424	1	0.9527
Linear SVC	0.9666	0.9409	0.9517	0.96
Extra Trees	0.9975	0.9409	1	0.9855

The ridge classifier performed well across the two data sets. It had training accuracies of 94.56% on Dataset 3 and 93.89% on Dataset 4. Its testing accuracy was also stable, performing at 94.58% on Dataset 3 and 93.82% on Dataset 4. Training times were relatively low, being at 0.0198 seconds on Dataset 3 and 0.0266 seconds on Dataset 4. The memory usage did increase for Dataset 4, showing how the complexity of the data somewhat affects the amount of resources needed.

The radius neighbors classifier exhibits considerable performance differences between the datasets. Here, in Dataset 3, the training accuracy is exceptional at 98.52% but drops to 84.24% during testing, pointing to a possibility of overfitting. On the contrary, for Dataset 4, both training and testing accuracies are perfect: 100% and 95.27%, respectively; thus, it indicates improved generalization capability. But, on the other hand, training time is much higher for Dataset 4 (0.8813 seconds) than for Dataset 3 (0.0079 seconds), and also increases memory usage.

The linear SVC performs well and equally well for both datasets. For Dataset 3, it shows that training accuracy is 96.66% and testing accuracy is 94.09% whereas for

Dataset 4, a training accuracy of 95.17% is found to be accompanied by a testing accuracy of 96%. Training times are very small, 0.0149 seconds for Dataset 3 and 0.0210 seconds for Dataset 4. Memory usage is reasonable compared to other models to show its computational efficiency.

The extra trees classifier shows the best training accuracy among all the classifiers at 99.75% for Dataset 3 and a perfect 100% for Dataset 4. The testing accuracies are equally good at 94.09% for Dataset 3 and 98.55% for Dataset 4. However, the training time and memory usage are very high as compared to other classifiers, at 2.8888 seconds for Dataset 3 and 3.0607 seconds for Dataset 4. The memory usage is quite high, especially for Dataset 4, at 1,562,172 bytes.

Fig. 7 illustrates that the radius neighbors classifier and the extra trees classifier achieve much higher accuracy, while the ridge classifier and the linear SVC have balanced performance with very low computational costs.

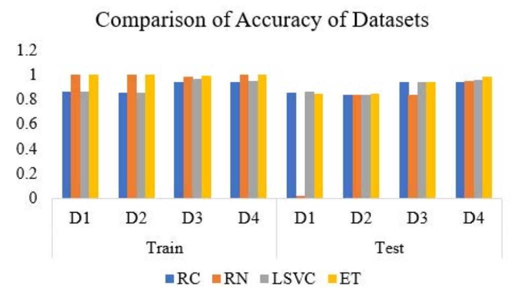


Fig. 7. Comparison of accuracy.

The choice of a classifier depends on the given requirements of the task, that is, computational resources, accuracy needs, or dataset complexity.

A comparison of model performance across Datasets 1–4 reveals important insights into generalizability. The classifiers generally achieved higher training and testing accuracy on Datasets 3 and 4, which were collected from real-world hospital settings, compared to the publicly available Datasets 1 and 2. For example, the Extra Trees classifier achieved nearly perfect training accuracy and over 98% testing accuracy on Dataset 4, outperforming its performance on Datasets 1 and 2. This suggests that the model may adapt well to real-world, clinically diverse data, likely due to more informative features and realistic variance. However, it also highlights that real-world datasets, when well-prepared and curated, can support high-performing ML models that generalize well to unseen clinical scenarios.

The trends observed across datasets also highlight the reliability and generalization behavior of the classifiers. For instance, models like Linear SVC and Extra Trees consistently maintained high precision and recall across both controlled (Datasets 1 and 2) and real-world clinical datasets (Datasets 3 and 4), indicating their robustness in varied data environments. In contrast, classifiers such as Radius Neighbors showed significant performance fluctuations, suggesting lower reliability in generalizing beyond the training data. These findings emphasize the importance of evaluating models on real-world data, not just benchmark datasets, to ensure their clinical applicability and deployment readiness.

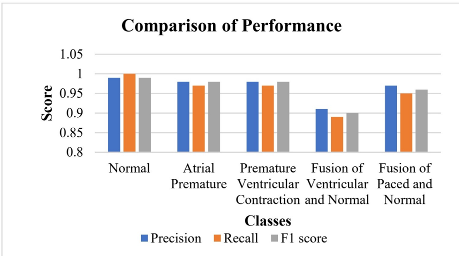
C. ECG Dataset Results

The CNN-RNN model is trained to predict different classes of ECG abnormalities and is validated using a validation dataset. Throughout the training, the model achieves a validation accuracy that closely matches the training accuracy. It indicates well well-generalized model without significant overfitting. In addition, the testing accuracy is high with low loss. The proposed system generates predictions on the test data, presenting results through a confusion matrix and metrics like F1-score, recall, and precision for each class. The best performance of the system is in the normal class, and the lowest ECG classification performance is in the FVAN class (Table VI). Both the classes, atrial premature and premature ventricular contraction, show similar performance, but less than that of the normal class. The FPAN class also represents notable performance.

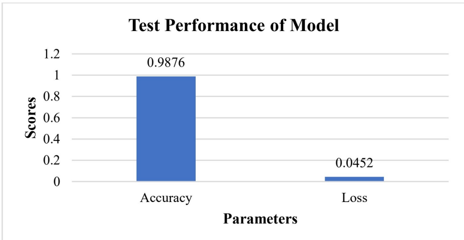
TABLE VI: PERFORMANCE OF PROPOSED MODEL

Classes	Precision	Recall	F1 score
Normal	0.99	1.00	0.99
APC	0.98	0.97	0.98
PVC	0.98	0.97	0.98
FVAN	0.91	0.89	0.90
FPAN	0.97	0.95	0.96

The results of the proposed model depict a precision of 0.99 for the normal class, slightly lower for atrial premature, and premature ventricular contraction at 0.98, with a recall of 0.97, besides an F1 score of 0.98. For the FPAN class, precision is slightly lower at 0.97, but the recall and F1 scores are reduced to 0.95 and 0.96, respectively. The lowest precision achieved is 0.91, with a recall of 0.89 and a 0.90 F1 score, for forecasting the ECG class as FVAN. The overall results indicate high performance, with a comparison of performance for all classes (Fig. 8 (a)). Furthermore, the projected model attained a significantly higher testing accuracy of 0.989 (Fig. 8 (b)). Thus, a very low loss (0.045) is observed, indicating strong performance of the model.



(a)



(b)

Fig. 8. Performance: (a) Comparison of all classes and (b) test performance.

Overall, the results show the highest performance in ECG-based prediction of myocardial infarction and heart diseases. When the proposed model of ECG-based prediction is further compared to existing methods, it exhibits enhanced performance as revealed in Table VII below. The proposed system attained an enhanced accuracy of 0.987 as compared to the model proposed by Moghadam and Asl [26] (0.965). While Cao *et al*. [27] illustrate the lowest performance (0.908) among all models. Alkurdi and Abdulazeez [28] also achieved better accuracy (0.960), but precision is low, 0.93, with slightly reduced recall (0.95). Although Riek *et al*. [29] performed well with an accuracy of 0.947, the recall is considerably low (0.88) among all models. Moreover, the proposed model achieves higher precision as well as recall at 0.98 and 0.97, respectively, as compared to Moghadam and Asl [26] which achieved at 0.96 and 0.97.

TABLE VII: COMPARING EXISTING MODELS WITH PROPOSED MODEL

Existing Methods	Accuracy	Precision	Recall
Riek *et al*. [29]	0.947	0.96	0.88
Cao *et al*. [27]	0.908	0.95	0.95
Alkurdi and Abdulazeez [28]	0.960	0.93	0.95
Moghadam and Asl [26]	0.965	0.96	0.97
Proposed method	0.987	0.98	0.97

Fig. 9 represents the performance of all the above models, indicating a significant increase in evaluation parameters for the proposed model as compared to the existing model.

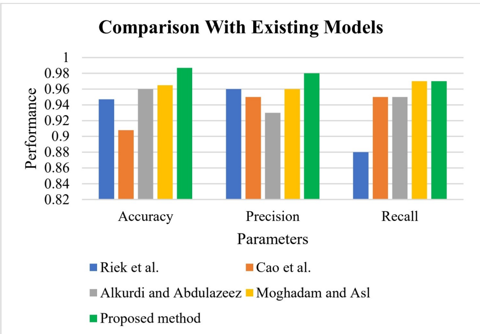


Fig. 9. Comparing the proposed model with existing models.

The strength of the proposed framework lies not in architectural novelty alone, but in its integration of both public and real-world clinical datasets for validation, the use of noise-augmented training to improve robustness, and the successful implementation of a multi-class ECG classification model using a hybrid CNN-RNN approach. This combination enhances the model's generalizability and relevance across varied clinical conditions. Furthermore, to improve clinical interpretability, future work may explore the integration of explainable AI techniques such as Grad-CAM or attention-based mechanisms, which would help clinicians better understand and trust the model's decisions. From the results, even though noise is added in the Gaussian form to augment the ECG signals, the overall achieved accuracy is quite high, at 0.987. Most classes have scores greater than 0.95, which indicates the robust performance of the model. However, the class "Fusion of Ventricular and Normal" shows lower performance to some extent. The result offers evidence of how the integration of the CNN and RNN outcomes leads to increased accuracy in the prediction of ECG models that can distinguish between MI and other conditions. In addition, the performance in achieving optimal scores of accuracies and loss at the time of testing shows that training and validation accuracy have been around 0.95, with a maximum accuracy of 0.98, meaning the model is effective in classifying the heart conditions using ECG signals. Further validation of the proposed method is attained through the comparison with existing studies. It is observed that the model performs well on all evaluation metrics. As a result, this study outshines previous methods by applying an integrated ML technique for predicting myocardial infarction (MI) using different datasets, which shows the great impact of ML in heart disease classification and MI diagnosis.

The proposed prediction model is a new advancement to diagnose myocardial infarction (MI) to improve classification accuracy. The model captures meritoriously both spatial as well as temporal features of ECG signals with notable improvements in performance metrics such as an accuracy value of 0.987 and very high precision and recall for most classes. This research therefore also highlights the successful integration of architectures comprising CNN and RNN while contributing to the production of robust and accurate early-detecting diagnostic tools, further contributing to saving thousands of lives through diagnosis based on medical signals that the current conventional methods often could not provide signals that the current conventional methods often could not provide. While the reported improvements in accuracy and recall may appear marginal in numerical terms, even small performance gains can be clinically significant, especially in early MI detection, where timely diagnosis can reduce mortality risk. In practical deployment, this model could be integrated into ECG monitoring systems in emergency departments or primary care clinics to assist physicians in the early identification of myocardial infarction. Additionally, its robustness across both standard and hospital datasets suggests potential for real-world use in diverse clinical settings, though further testing on broader external datasets is essential. Future deployment could also involve integration into wearable devices or cloud-based diagnostic platforms under the Internet of Medical Things (IoMT) framework.

V. CONCLUSION

Heart diseases, particularly Myocardial Infarction (MI), remain a major public health concern, where early and accurate diagnosis is critical to improving patient outcomes. Machine Learning (ML) has transformed the ability to analyze complex patient data rapidly and reliably, facilitating earlier detection and intervention. This study presents an ML-based approach for heart disease prediction using both EHR and ECG data. A comparative analysis across four datasets—two public and two real-world hospital datasets—demonstrated consistent and robust performance, with linear SVC showing the most stable results and the Extra Trees classifier achieving the highest accuracy in training.

The proposed CNN-RNN-based ECG classification model achieved 98.7% accuracy in identifying five heartbeat classes, confirming its potential in clinical applications. The application of Gaussian noise augmentation contributed to improved robustness and generalizability, especially for minority classes. Comparisons with existing studies further validate the strength of the proposed model in outperforming several established methods across multiple metrics.

Importantly, improved MI prediction using such automated models could directly benefit patient care by enabling faster diagnosis in emergency settings, supporting clinical decision-making, and reducing the likelihood of missed or delayed treatment. In real-world deployment, this could lead to earlier interventions, reduced hospitalization time, and ultimately, better survival rates.

Despite strong results, certain heartbeat classes—such as "Fusion of Ventricular and Normal"—show relatively lower performance, indicating areas for further model refinement. Future research should explore the use of larger and more demographically diverse datasets to enhance generalizability across populations. In addition, integrating techniques such as transfer learning, attention mechanisms, or hybrid ensemble models could improve performance for rare or complex ECG patterns. Incorporating explainable AI components (e.g., Grad-CAM or SHAP) would also enhance model transparency, which is essential for clinical acceptance. Finally, deployment in real-time monitoring systems and wearable devices under the Internet of Medical Things (IoMT) framework could further expand the practical utility of the proposed model in remote and continuous cardiac health monitoring.

Data Availability and Ethical Considerations

Datasets 1 and 2 used in this study are publicly available. Dataset 1 was sourced from an open-access heart disease prediction repository, and Dataset 2 is derived from the Framingham Heart Study dataset available on Kaggle. Both datasets are anonymized and contain no personally identifiable information.

Dataset 3 and Dataset 4 were collected from Bharati Vidyapeeth (Deemed to be University), Medical College and Hospital, Sangli – Miraj Road, and Aryan Heart Care, Miraj, Maharashtra. These datasets were collected under institutional approval by the authors from Bharati Vidyapeeth Medical College and Aryan Heart Care, Miraj. The data is not publicly available but may be shared by the corresponding author upon reasonable request, subject to ethical compliance. No identifiable patient information was used at any stage. As the data was retrospective and anonymized, the requirement for informed consent was waived by the respective institutional ethics committees. All procedures followed ethical guidelines for research involving human data.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Shridevi K. Jamage conducted the research; Ramesh Y. Mali and Virendra V. Shete analyzed the data and provided guidance in research methods and contributed to data analysis; Shridevi K. Jamage wrote the paper; all authors had approved the final version.

Acknowledgment

The author thanks the colleagues from SCTR's Pune Institute of Computer Technology, Pune (MS), who provided insight and expertise that greatly assisted the research. F. A. The author thanks Dr. Ramesh Y. Mali and Dr. Virendra V. Shete, MIT ADT University, Pune (MS), for their assistance, which greatly improved the manuscript.

References

[1] S. Mandala, S. S. Amini, Adiwijaya *et al.*, "Enhanced myocardial infarction identification in phonocardiogram signals using segmented feature extraction and transfer learning-based classification," *IEEE Access*, vol. 11, pp. 136654–136665, 2023. doi: 10.1109/ACCESS.2023.3338853

[2] J. Mishra and M. Tiwari, "IoT-enabled ECG-based heart disease prediction using three-layer deep learning and meta-heuristic approach," *Signal, Image Video Process.*, vol. 18, no. 1, pp. 361–367, Feb. 2024. doi: 10.1007/s11760-023-02743-4

[3] D. B. Olawade, N. Aderinto, G. Olatunji, E. Kokori, A. C. David-Olawade, and M. Hadi, "Advancements and applications of Artificial Intelligence in cardiology: Current trends and prospects," *J. Med. Surgery, Public Heal.*, vol. 3, 100109, Aug. 2024. doi: 10.1016/j.glmedi.2024.100109

[4] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine learning-based predictive models for detection of cardiovascular diseases," *Diagnostics*, vol. 14, no. 2, p. 144, Jan. 2024. doi: 10.3390/diagnostics14020144

[5] S. Bebortta, S. S. Tripathy, S. Basheer, and C. L. Chowdhary, "FedEHR: A federated learning approach towards the prediction of heart diseases in IoT-based electronic health records," *Diagnostics*, vol. 13, no. 20, 3166, Oct. 2023.

[6] P. Pabitha, R. Praveen, K. C. J. Chandana, S. Ponlibarnaa, and A. S. Aparnaa, "A comparative study of deep learning models for ECG Signal-based user classification," in *Proc. 2023 12th International Conference on Advanced Computing (ICoAC)*, Aug. 2023, pp. 1–8. doi: 10.1109/ICoAC59537.2023.10249402

[7] R. Khera, J. Haimovich, N. C. Hurley *et al.*, "Use of machine learning models to predict death after acute myocardial infarction," *JAMA Cardiol.*, vol. 6, no. 6, pp. 633–641, Jun. 2021.

[8] E. Nasarian, R. Alizadehsani, U. R. Acharya, and K.-L. Tsui, "Designing interpretable ML system to enhance trust in healthcare: A systematic review to proposed responsible clinician-AI-collaboration framework," *Inf. Fusion*, vol. 108, 102412, Aug. 2024. doi: 10.1016/j.inffus.2024.102412

[9] V. K. Sudha and D. Kumar, "Hybrid CNN and LSTM network for heart disease prediction," *SN Comput. Sci.*, vol. 4, no. 2, p. 172, Jan. 2023. doi: 10.1007/s42979-022-01598-9

[10] J. Miah, D. M. Ca, M. A. Sayed, E. R. Lipu, F. Mahmud, and S. M. Y. Arafat, "Improving cardiovascular disease prediction through comparative analysis of machine learning models: A case study on myocardial infarction," in *Proc. 2023 15th International Conference on Innovations in Information Technology (IIT)*, Nov. 2023, pp. 49–54. doi: 10.1109/IIT59782.2023.10366476

[11] N. Hazzaa and O. Yıldız, "Medical diagnosis support system for cardiovascular disease prediction based on machine learning," *J. Millimeterwave Commun. Optim. Model.*, vol. 4, no. 2, pp. 59–63, 2024.

[12] N. S. Suryawanshi, "Accurate prediction of heart disease using machine learning: A case study on the cleveland dataset," *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 7, pp. 1042–1049, 2024.

[13] E. A. Ogundepo and W. B. Yahya, "Performance analysis of

supervised classification models on heart disease prediction," *Innov. Syst. Softw. Eng.*, vol. 19, no. 1, pp. 129–144, Mar. 2023.
[14] K. B. Sk, D. Roja, S. S. Priya, L. Dalavi, S. S. Vellela, and V. Reddy, "Coronary heart disease prediction and classification using hybrid machine learning algorithms," in *Proc. 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, Mar. 2023, pp. 1–7. doi: 10.1109/ICIDCA56705.2023.10099579
[15] K. Natarajan, V. V. Kumar, T. R. Mahesh *et al.*, "Efficient heart disease classification through stacked ensemble with optimized firefly feature selection," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 174, Jul. 2024. doi: 10.1007/s44196-024-00538-0
[16] U. R. Acharya, H. Fujita, V. K. Sudarshan *et al.*, "Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal," *Knowledge-Based Syst.*, vol. 132, pp. 156–166, Sep. 2017.
[17] H. Wu, "Multiscale entropy with electrocardiograph, electromyography, electroencephalography, and photoplethysmography signals in healthcare: A twelve-year systematic review," *Biomed. Signal Process. Control*, vol. 93, 106124, Jul. 2024. doi: 10.1016/j.bspc.2024.106124
[18] J. Karhade, S. K. Ghosh, P. Gajbhiye, R. K. Tripathy, and U. R. Acharya, "Multichannel multiscale two-stage convolutional neural network for the detection and localization of myocardial infarction using vectorcardiogram signal," *Appl. Sci.*, vol. 11, no. 17, 7920, Aug. 2021. doi: 10.3390/app11177920
[19] X. Bai, X. Dong, Y. Li, R. Liu, and H. Zhang, "A hybrid deep learning network for automatic diagnosis of cardiac arrhythmia based on 12-lead ECG," *Sci. Rep.*, vol. 14, no. 1, 24441, Oct. 2024. doi: 10.1038/s41598-024-75531-w
[20] B. S. Raghukumar and B. Naveen, "Enhancing myocardial infarction diagnosis: insights from ECG image analysis and machine learning," *SN Comput. Sci.*, vol. 5, no. 5, p. 448, Apr. 2024. doi: 10.1007/s42979-024-02827-z
[21] J. Lee and M. Shin, "Cross-database learning framework for electrocardiogram arrhythmia classification using two-dimensional beat-score-map representation," *Appl. Sci.*, vol. 15, no. 10, p. 5535, May 2025. doi: 10.3390/app15105535
[22] X. Feng, F. Lin, Q. Li, and P. Zhang, "A segmentation-based deep learning method for automatic atrial fibrillation detection using RR intervals," in *Proc. the 2024 9th International Conference on Biomedical Signal and Image Processing*, New York, USA, Aug. 2024, pp. 23–28. doi: 10.1145/3691521.3691536
[23] D. Lapp. (Jul. 08, 2025). Heart Disease Dataset. Kaggle. [Online]. Available: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset
[24] Dataset. Logistic regression To predict heart disease. Kaggle. [Online]. Available: https://www.kaggle.com/datasets/dileep070/heart-disease-prediction-using-logistic-regression
[25] S. Fazeli. ECG Heartbeat Categorization Dataset. kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/shayanfazeli/heartbeat/data
[26] S. Ramezani Moghadam and B. M. Asl, "Automatic diagnosis and localization of myocardial infarction using morphological features of ECG signal," *Biomed. Signal Process. Control*, vol. 83, 104671, May 2023. doi: 10.1016/j.bspc.2023.104671
[27] M. Cao, T. Zhao, Y. Li, W. Zhang, P. Benharash, and R. Ramezani, "ECG heartbeat classification using deep transfer learning with convolutional neural network and STFT technique," *J. Phys. Conf. Ser.*, vol. 2547, no. 1, 012031, Jul. 2023.
[28] A. Alkurdi and A. M. Abdulazeez, "Comprehensive classification of fetal health using cardiotocogram data based on machine learning," *Indones. J. Comput. Sci.*, vol. 13, no. 1, pp. 277–300, 2024.
[29] N. T. Riek, M. Akcakaya, Z. Bouzid *et al.*, "ECG-SMART-NET: A deep learning architecture for precise ECG diagnosis of occlusion myocardial infarction," Jun. 2025. doi: https://doi.org/10.1109/TBME.2025.3573581

Ms. Shridevi K. Jamage is a research scholar in the Department of Electronics and Communication Engineering (ECE), MIT School of Engineering and Science, Pune. She holds a master of engineering (M.E.) degree in electronics and telecommunication engineering (E&TC) with a specialization in Microwave Engineering from Savitribai Phule Pune University (SPPU). With over 11 years of academic and teaching experience, she is currently serving as an Assistant Professor in the Department of E&TC at SCTR's Pune Institute of Computer Technology (PICT), Pune, India. Her research contributions include the publication of nine papers in reputed international journals and conferences. She is also the author of the textbook Basic Electronics Engineering (SPPU, 2022), published by Fortflag Books and Media Private Limited. Her teaching and research interests include microwave engineering, analog circuit design, artificial intelligence, and machine learning.

Ramesh Y. Mali is the professor and head of the Department of Electrical and Electronics Engineering (EEE) at the MIT School of Computing, MIT-ADT University, Pune. He has over 23 years of teaching experience and holds professional memberships in ISTE (Life Member), IETE, and IEEE. He also serves as a Board of Studies (BOS) Member at MIT Art, Design & Technology University, Pune. His academic contributions include more than 15 research/journal papers, over 10 conference papers, and three granted patents. He is an International Reviewer for reputed journals. Throughout his career, he has supervised three Ph.D. scholars, guided 17 postgraduate students, and mentored 28 undergraduate project groups.

Virendra Virbhadra Shete serves as the director of MIT School of Engineering and Science, Pune and Chairman at Institute of Electronics & Telecommunication Engineers, Pune Center, Chair, IEEE Education Society, Pune Section. He is a highly motivated educator and researcher with 28 years of a successful track record spanning education, teaching, and research. He possesses a solid foundational education, punctuality, and reliability coupled with a robust aptitude for organization and administration.