Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 12, No. 01, March 2026.

MULTI-CLASS HEART DISEASE CLASSIFICATION USING MULTI-LEAD ECG FEATURES AND ENSEMBLE LEARNING

Bala Venkateswarlu Isunuri, Venkata Sravya Alapati, Sri Charan Marripudi, Gopala Krishna Parimi and Akshaya Valli Koganti

(Received: 28-Sep.-2025, Revised: 13-Dec.-2025, Accepted: 6-Jan.-2026)

ABSTRACT

Cardiovascular diseases (CVDs) are the leading causes of global mortality and require an early and precise diagnosis. This work presents an automated multi-class classifier for diagnosing cardiac disease from electrocardiogram (ECG) images through image processing and machine-learning techniques. The proposed framework consists of three steps, including pre-processing, feature extraction, and ensemble learning. Initially, the ECG image undergoes a comprehensive pre-processing pipeline that includes lead segmentation, grayscale conversion, Gaussian filtering, and Otsu thresholding. The contour-based features are extracted and then reduced by PCA to preserve discriminative information. Finally, multiple machine-learning models, including K-nearest neighbors (KNNs), Random Forest and support vector machines (SVMs), are ensembled using voting and stacking classifiers to improve the performance of the proposed framework. The proposed ensemble model is evaluated on a public dataset that consists of ECG images that are categorized into four classes: normal, abnormal, myocardial infarction (MI), and history of MI. The proposed ensemble model attained the highest classification accuracy of 98.06% and outperformed the existing pre-trained and state-of-the-art models.

KEYWORDS

Electrocardiogram, Ensemble learning, Multi-lead ECG features, Multi-class heart disease classification.

1. INTRODUCTION

Cardiovascular diseases (CVDs) are a major public-health issue and the leading cause of death worldwide. The World Health Organization (WHO) says that CVDs kill 17.9 million people each year, which is 31% of all deaths in the world [2]. CVDs are a category of diseases that include coronary artery disease, myocardial infarction (MI), heart failure, and arrhythmias. These diseases don't show up until later stages. For early medical treatment, lower healthcare costs, and a better quality of life, it is important to diagnose heart disorders correctly and early. Electrocardiography (ECG) is the most relevant and extensively employed non-invasive diagnostic modality for evaluating cardiac electrical activity. A conventional ECG provides important information about the heart's rhythm, conduction patterns, and how anomalies show up, which helps doctors figure out what kind of heart disease a person has [3]. The physical examination of an ECG incurs considerable time and practice, and the results may vary among individuals. This constraint makes it exceedingly challenging to identify by a cardiologist in rural or under-resourced regions. This issue has generated a demand for automated ECG analysis solutions utilizing artificial intelligence (AI) to enhance diagnostic accuracy and scalability for physicians. Recent breakthroughs in machine learning (ML) and deep learning (DL) have significantly revolutionized the domain of biological signal processing. Techniques, such as convolutional neural networks (CNN), recurrent neural networks (RNNs) and ensemble-learning models, have achieved unprecedented advancements in the classification of cardiac diseases utilizing ECG waveforms and pictures [4]-[5]. The models can discover valuable traits, discern subtle trends, and manage extensive data volumes with limited human involvement.

CNN-based structures have widely been applied to image-related ECG classification applications, as they have superior capabilities in spatial-learning features [6]. Some previous studies have successfully applied these methods, yielding reasonable outcomes. In [1], the authors employed CNN models like MobileNetV2 and VGG16, for four-class classification of ECG images into Normal, Abnormal, Myocardial Infarction (MI), and History of MI. The method employed real-time deployment by

---

B. V. Isunuri (Corresponding Author), V. S. Alapati, S. C. Marripudi, G. K. Parimi and A. V. Koganti are with the Department of Computer Science and Engineering, SRM, University AP, Amaravati, Andhra Pradesh, India. Emails: {`bala.v`, `venkatasravya_a`, `sricharan_m`, `gopalakrishna_p` and `akshayavalli_k`}`@sramp.edu.in`

executing the model on Raspberry Pi platforms and obtained classification accuracy rates of up to 95%. The authors largely trained the model using end-to-end learning without proper signal pre-processing and lead-wise segmentation, potentially limiting its interpretability and adaptability. Other works followed the traditional machine-learning approach by employing hand-designed pre-processing pipelines with ML classifiers. For instance, works in [9] and [10] pre-processed ECG images by segmenting them into 12 leads, computed statistics and shape-based features, and employed classifiers, like Support Vector Machines (SVMs), K-Nearest Neighbors (KNNs), and Logistic Regression. Such pipelines employed Principal-component Analysis (PCA) as a feature reducer and utilized voting classifiers for class accuracy improvement, with the best accuracy of 94.2%. They employed only 12 leads and did not follow strict validation regimes, like K-fold cross-validation. Hence, they are less robust and less generalizable.

We present a new and complete ECG-based cardiac-disease classification system that takes the best of these previous works and extends them. Our contribution is 13 -lead ECG segmentation that records a wider and more nuanced spatial coverage of the heart's electrical activity than conventional 12-lead systems. Each lead is treated separately, enabling our system to learn local features from alternative heart views. This approach alone makes our model structurally and physiologically different, enabling better representation of subtle ECG waveform abnormalities. Further, our ensemble technique aggregates the predictions of multiple base classifiers, including SVMs, KNNs, logistic regression, and XGBoost, to achieve maximum diversity and performance. The contributions of this paper are multifold:

1. Adding 13-lead segmentation to improve spatial resolution.
2. A pre-processing pipeline with grayscale, Gaussian filtering, and Otsu thresholding.
3. The system integrates multiple conventional classifiers using ensemble-learning methods.
4. Cross-validation is utilized to obtain a precise performance estimate.

These contributions together offer a scalable, interpretable, and accurate solution for automated ECG interpretation with high potential for application in real-world clinical decision-support systems and portable diagnostic devices. The paper is organized as follows: Section 1 introduces multi-class cardiovascular-disease classification. The literature review is presented in Section 2. Details of the proposed ensemble framework are explored in Section 3. Section 4 includes an analysis of the results and a discussion, while Section 5 concludes the findings.

2. LITERATURE REVIEW

In recent years, the application of artificial-intelligence (AI) methods in automated electrocardiogram (ECG) data processing for the identification and classification of cardiovascular diseases has significantly increased. Conventional methods rely heavily on clinical expertise and human judgment, resulting in subjectivity or delays. Researchers have sought to utilize both conventional machine-learning (ML) and deep-learning (DL) techniques to improve diagnostic accuracy and scalability. Various studies have examined deep-learning models for the classification of ECG images. Lightweight CNN architectures, such as MobileNetV2 and VGG16, have been employed to extract discriminative features from ECG images, facilitating real-time implementation on embedded systems [1]. The models demonstrated consistent effectiveness in detecting cardiac anomalies, including myocardial infarction and arrhythmias, with an accuracy of up to 95%. Convolutional Neural Networks (CNNs) require extensive datasets and GPU-based training and may lack interpretability-factors that can hinder their clinical implementation in resource-constrained settings [3]-[4]. Nevertheless, conventional machine-learning techniques have gained prominence due to their cost-effectiveness and comprehensibility. Effective approaches for ECG classification include support vector machines (SVMs), K-Nearest Neighbors (KNNs), logistic regression, and XGBoost. These methods are accurate when utilized alongside well-crafted features [2][5][7]. Although the models demonstrate robust performance, they may exhibit heightened sensitivity to noise and extraneous features, particularly when handling high-dimensional data. Ensemble approaches, like voting and stacking, have been proposed to tackle these issues. These are predicated on the robustness of a mixture of multiple base models to enhance their durability and generalization capabilities. Voting classifiers employ majority voting to combine the judgments, whereas stacking uses a meta-model to select the ideal combination of basic outputs to reach improved accuracy [6], [9]. For example, [9] employed ensemble models optimized by GridSearch, attaining a peak accuracy of 92.4%. Pre-processing is an essential component in rendering feature extraction significant. Gray-scale conversion, Gaussian filtering, and Otsu's thresholding are methods commonly employed to enhance the clarity of the ECG waveform [2][5][10]. Principal-component Analysis (PCA) is employed to diminish feature dimensionality while preserving the most pertinent information with decreased computational effort. The latest research is confined to standard 12-lead ECG pictures.

Finally, the literature encompasses a broad spectrum of CVD detection from ECGs. However, there is still a gap in performance optimization that does not compromise interpretability or increase computational complexity. Our contribution bridges these gaps by putting forward an ensemble learning-based classification pipeline with improved performance without deep architectures. In the proposed framework, we incorporate 13-lead segmentation, providing greater spatial resolution of the cardiac activity. We incorporate ensemble approaches, including voting and stacking with cross-validation, for robust and balanced assessment. These additions make our model more capable of performing better than conventional ML pipelines [9]-[10] as well as deep CNN-based models, like MobileNetV2 [1].

3. METHODOLOGY

Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of input ECG images, where every $x_i \in \mathbb{R}^{2213\times1572\times3}$. Similarly, $Y = \{y_1, y_2, \dots, y_n\}$ represents class labels of each x_i, where $y_i \in \{0,1,2,3\}$. Table 1 lists the class description for each class. The proposed framework comprises three stages for multi-class heart-disease classification. It accepts the input ECG image x_i and produces the class label y_i, as shown in Figure 1. The details of each stage are as follows.

Table 1. Description of each class label.

Class label	Description
0	HB (Abnormal Heart Beat)
1	MI (Myocardial Infraction)
2	Normal
3	HMI (History of MI)

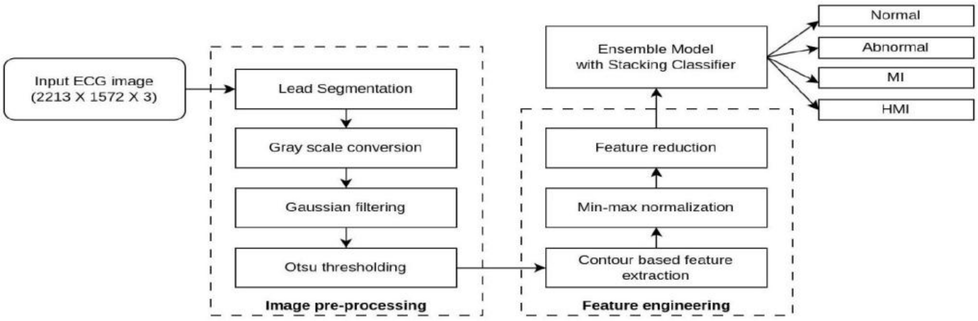
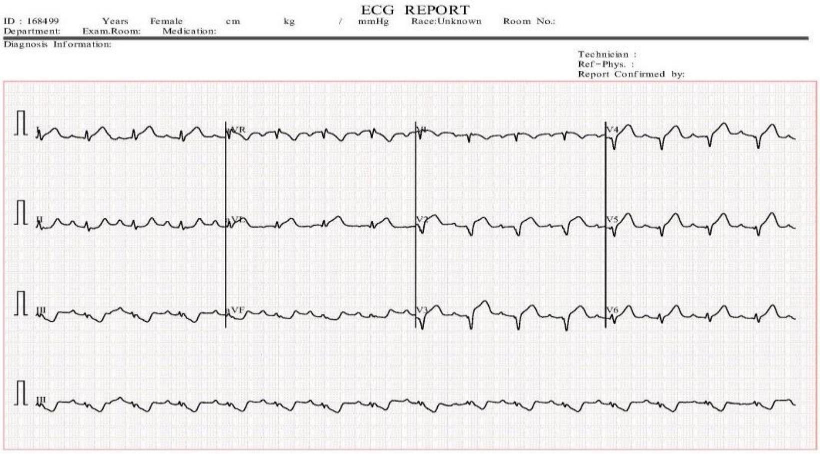


Figure 1. The proposed multi-lead features and ensemble learning framework.



2020-08-03 10:36:08 PM

Figure 2. Sample ECG image with 13 leads.

3.1 Image Pre-processing

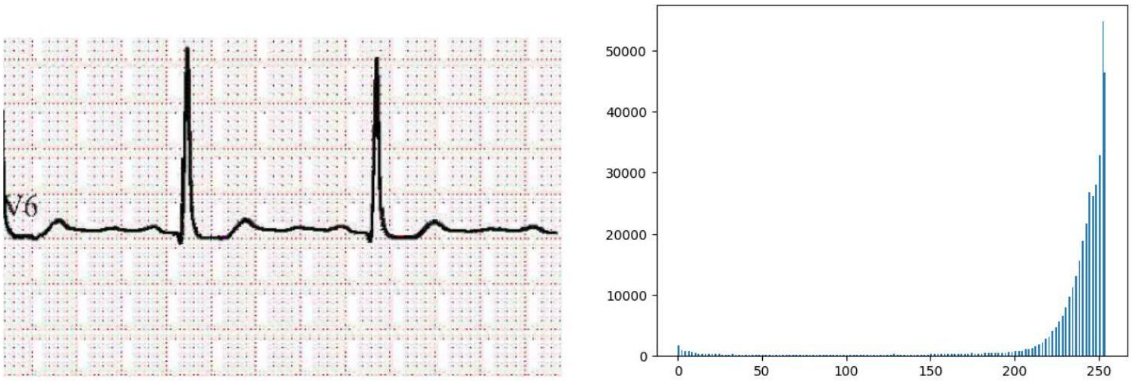
The proposed framework utilizes a specialized pre-processing pipeline to enhance ECG-image features. The pre-processing of images involves four steps, and the details are as follows.

1) Lead Segmentation

The literature reveals that existing models utilize 12 ECG leads for the classification of heart diseases [22]. Madias [19] presented the importance of lead 13 of ECG in heart diseases. This motivated us to involve the 13th lead for the proposed framework. Thus, each ECG image x_i is divided into 13 independent leads, each of which corresponds to different electrical activities captured from different parts of the heart. The dataset shows that each lead L_j has fixed coordinates. Thus, coordinate slicing across standard spatial positions for all images was used to mark the 13 leads $L = \{L_1, L_2, \dots L_{13}\}$, where every L_j is a single lead image. Figure 2 represents a sample ECG image acquired from the dataset [20]. The leads are obtained from three horizontal bands and one bottom strip. Leads 1-4 are taken from the top section (rows 300-600), leads 5-8 from the middle band (rows $600 - 900$), and leads $9 - 12$ from the bottom band (rows $900 - 1200$). Lead 13 spans rows 1250-1480 and spans the entire horizontal range, which is known as an extended rhythm strip. This partitioning allows for uniform spatial extraction of waveforms for all samples. The derived leads are harvested based on actual coordinate regions as follows:

$$\begin{gathered}
L_1 \leftarrow x_i[300{:}\,600{,}150{:}\,643] \\
L_2 \leftarrow x_i[300{:}\,600{,}646{:}\,1135] \\
L_3 \leftarrow x_i[300{:}\,600{,}1140{:}\,1625] \\
L_4 \leftarrow x_i[300{:}\,600{,}1630{:}\,2125] \\
L_5 \leftarrow x_i[600{:}\,900{,}150{:}\,643] \\
L_6 \leftarrow x_i[600{:}\,900{,}646{:}\,1135] \\
L_7 \leftarrow x_i[600{:}\,900{,}1140{:}\,1625] \\
L_8 \leftarrow x_i[600{:}\,900{,}1630{:}\,2125] \\
L_9 \leftarrow x_i[900{:}\,1200{,}150{:}\,643] \\
L_{10} \leftarrow x_i[900{:}\,1200{,}646{:}\,1135] \\
L_{11} \leftarrow x_i[900{:}\,1200{,}1140{:}\,1625] \\
L_{12} \leftarrow x_i[900{:}\,1200{,}1630{:}\,2125] \\
L_{13} \leftarrow x_i[1250{:}\,1480{,}150{:}\,2125]
\end{gathered} \tag{1}$$

In general, ECG images consist of red-colored grid-like patterns on which black-colored ECG signals are printed. Figure 3(a) and Figure 3(b) depict the 12th lead segment color image and corresponding histogram of the sample ECG image shown in Figure 2. From the histogram, it can be observed that there is a huge number of pixels (around 50000) with high intensity above 250. Simple thresholding will cause loss of ECG signal information. Thus, the proposed framework utilizes a sequence of pre-processing steps to preserve the ECG wave pattern.



(a) (b)

Figure 3. Color image of 12th lead segment along with its histogram.

2) Grayscale Conversion

Each color lead image (L_j) is transformed to grayscale to reduce dimensionality and emphasize signal morphology over color. The grayscale conversion is defined by Equation 2.

$$L_{\text{gray}}(x, y) = 0.29 \cdot R(x, y) + 0.58 \cdot G(x, y) + 0.11 \cdot B(x, y) \tag{2}$$

The red, green, and blue intensities of pixel (x, y) are represented by R, G, and B . The value of $L_{\text{gray}}(x, y)$ represents the resulting grayscale intensity for pixel (x, y). This conversion maximizes downstream processing and excludes color data that is irrelevant to the analysis of ECG waveforms. Figure 4(a) and Figure 4(b) depict grayscale of 12th lead segment color image and corresponding histogram, respectively. From the histogram, it can be observed that the peak number of pixels has reduced to 10000 due to single channel.

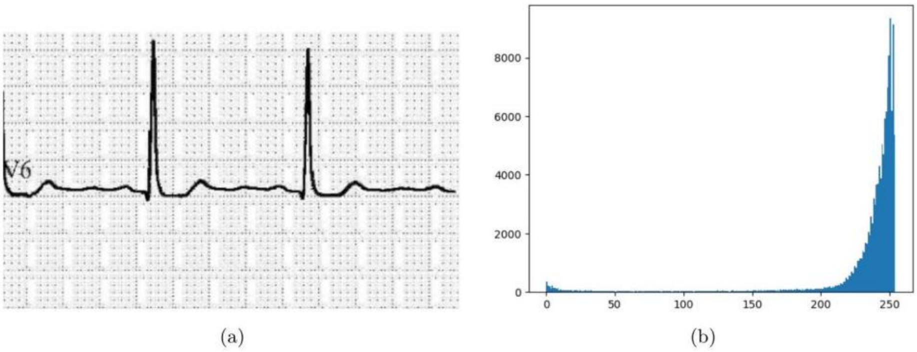


Figure 4. Grayscale image of the 12th lead segment along with its histogram.

3) Gaussian Filtering

The grayscale image was filtered with a Gaussian filter to cut down on high-frequency noise and blur it without losing the structural outlines [21]. The two-dimensional Gaussian kernel looks as follows:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{3}$$

In this case, σ stands for the standard deviation of the Gaussian filter. It tells the quantity of smoothing, and the blurring effect of the Gaussian filter can be controlled using the σ value. The ECG images need more smoothing to suppress the effect of the background grid pattern. Thus, we have considered high sigma as $\sigma = 0.7$. Similarly, kernel size is another parameter that influences the feature quality. Thus, we have conducted an ablation study with different kernel sizes to identify a better kernel size. The results of the ablation study are reported in the results section. The filtered image is the result of the convolution of the grayscale image with the Gaussian kernel generated from Equation 3. The process smooths the ECG trace lines by eliminating background variance and minor graphical artifacts. Figure 5(a) and Figure 5(b) depict the smoothened image of the 12th lead segment image and the corresponding histogram, respectively. From the histogram, it can be observed that the peak number of pixels has further reduced to 7000 due to suppression of background pixels.

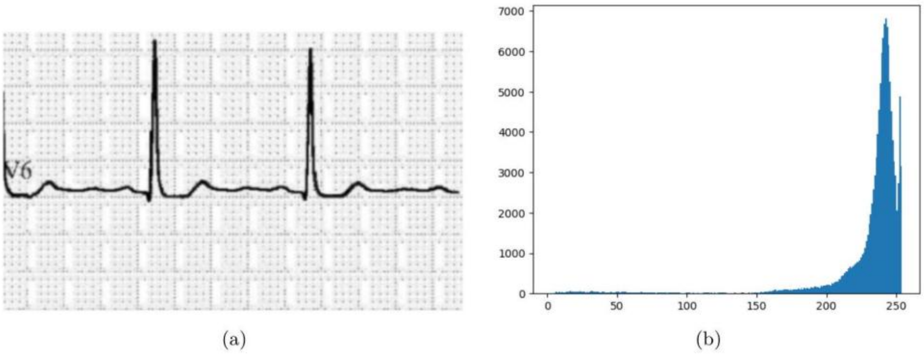


Figure 5. Smoothened image of the 12th lead segment along with its histogram.

4) Otsu's Thresholding

We apply Otsu's method to turn the smoothened image into binary and make the ECG signal stand out as a binary image. Otsu's thresholding finds the best threshold that maximizes the difference between background and foreground pixels [21]. The optimal threshold t is the one that maximizes $\sigma_b^2(t)$.

$$\sigma_b^2(t) = \omega_0(t) \cdot \omega_1(t) \cdot [\mu_0(t) - \mu_1(t)]^2 \tag{4}$$

In this context, the symbols $\omega_0(t)$ and $\omega_1(t)$ represent the probabilities (weights) associated with the background and foreground classes, respectively. Similarly, $\mu_0(t)$ and $\mu_1(t)$ are the mean intensities of background and foreground classes. $\sigma_b^2(t)$ is between-class variance at threshold t. Finally, the Otsu thresholded binarized image is resized to 300×450 to reduce computational complexity. Figure 6 depicts the Otsu thresholded 12th lead segment, which removes the grid pattern of the ECG color image.

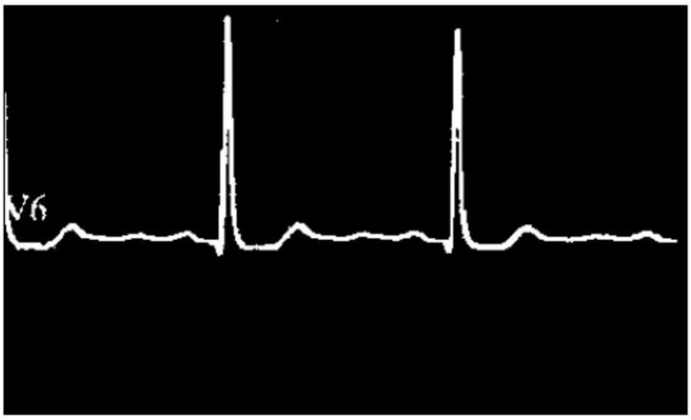


Figure 6. Otsu thresholded binarized image of 12th lead segment.

This phase transforms input image x_i to 13 binarized lead images $BL = \{BL_1, BL_2, \dots BL_{13}\}$ with improved contrast between the ECG waveform and background. This pre-processing sequence guarantees that the signal contours are maintained and evidently separable, allowing for efficient feature extraction in the subsequent phase.

3.2 Feature Engineering

Contour-based features are employed in the proposed framework. Therefore, the feature-engineering process consists of three sub-tasks as outlined below.

1) Contour-based Feature Extraction

After Otsu thresholding, contour detection is used to identify contours and boundaries of the shapes within the segmented leads. From the contours, morphological characteristics are obtained to describe significant structural characteristics of the segmented leads. The characteristics may be shape, size, and orientation. Other morphological features may include perimeter, aspect ratio, convexity, …etc. Figure 7 depicts the contour plot of the 13th lead segment. From that contour plot, only 255 points from each lead will be selected at equal internals.

2) Min-Max Normalization

The range of contour-based features depends on the lead number, which leads to varying scales and distributions. Therefore, normalization of the features extracted from each lead is necessary. We have considered min-max normalization to normalize the features using the following equation.

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{5}$$

where X is the original feature value. $X_{\min}$ and $X_{\max}$ are the lower and upper limits of the feature, respectively. X' is the normalized feature value; now the values of each lead image are in the range $[0,1]$. After normalization, each lead vector has 255 values. Thus, this stage generates a feature vector F with $13 \times 255 = 3315$ features for each ECG image.

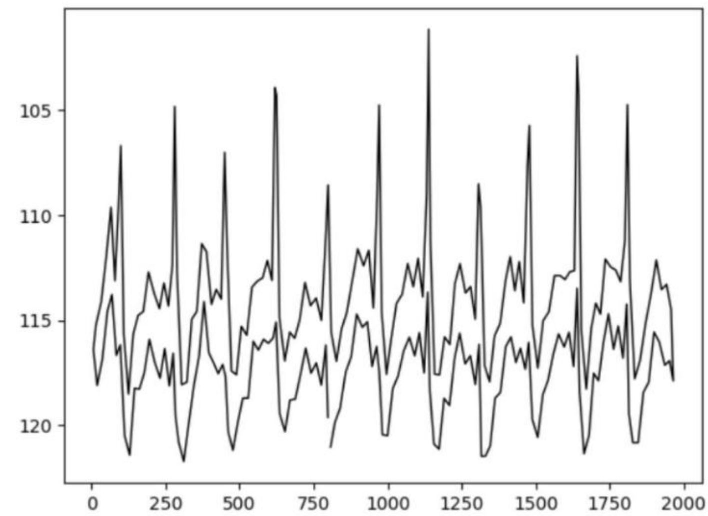


Figure 7. Contour plot of 13th lead segment.

3) Feature Reduction

To make ensemble learning less complicated, we used feature reduction on the feature vector F obtained from the previous phase. Principal-component Analysis (PCA) is a way to reduce the number of dimensions of features by mapping them onto a new collection of orthogonal components (principal components). PCA seeks to get rid of redundant information while keeping the most useful parts. The PCA implementation finds the eigenvalues and eigenvectors of the covariance matrix and uses them to move the data into a new sub-space, which makes it smaller. The output of the previous step produces a huge number of features (3315 features). It needs a feature reduction to optimize the computational cost of the proposed model. However, the number of features is significant for how well ensemble learning works. Therefore, we conducted an ablation study with varying numbers of features to achieve the optimal performance. We discovered in our ablation study that a feature vector with 512 features performs the best of the options that we looked at. This stage generates a feature vector with 512 normalized contour-based features.

3.3 Ensemble Learning

We used machine learning algorithms, including Support Vector Machines (SVMs), K-Nearest Neighbors (KNNs), and Logistic Regression (LR) to create the proposed ensemble framework. SVM builds a maximum-margin hyperplane to separate classes in a high-dimensional space. SVM is able to handle nonlinear data by employing kernel tricks and generalizes extremely well. SVM is particularly effective when applied to small- to medium-sized biomedical datasets. KNN is a very basic, non-parametric classifier that computes the majority vote of the K-nearest neighbors. Relevant features and high-dimensional data adversely affect KNN's ability to maintain local patterns. LR is class membership probability as a logistic function. LR is computationally efficient, interpretable, and a useful linear baseline for multi-class classification using the one-vs.-rest strategy. XGBoost is an outstanding ensemble method that is gradient-boosted decision tree-based. XGBoost performs best with complex patterns, restricts overfitting, and offers high accuracy as well as efficient operation. The individual training and testing of all models on the PCA-reduced feature vectors are performed to create the performance baselines. Two ensemble techniques were employed to obtain the highest individual model and lowest predictive consistency.

- Voting classifier: This ensemble method uses soft voting to combine the probability output of chosen base classifiers. Classification was performed by averaging predicted probabilities and selecting the best combined score label.
- Stacking classifier: The base-level classifiers' predictions are used as input to a meta-classifier.

The second-level learner learns the base models' dependencies, which enabled improved decision boundaries and classification outcomes.

Ensemble models are created to take advantage of the complementary strengths of each algorithm and to boost the overall robustness. The proposed ensemble model utilized the grid search-based cross validation method with five folds for hyper-parameter tuning. We achieved the best performance with an ensemble model that consists of SVM with $C = 1$, SVM with Gamma=0.01, KNN with 5 neighbors and Random Forest with 300 trees.

4. RESULTS

This section presents a detailed performance analysis of the proposed ensemble model along with an ablation study. Model performance was approximated using common classification metrics, including accuracy, precision, recall, and F1-score. The proposed model has been evaluated on the Mendeley Dataset [20]. The dataset holds 928 ECG color images with a resolution of 2213×1572. Each image belongs to one of the four cardiac conditions: normal, abnormal, myocardial infarction (MI), and history of MI (HMI).

4.1 Ablation Study

Gaussian filtering and feature reduction primarily influence the performance of the proposed model. Therefore, we conducted an ablation study focusing on these two aspects.

- **With different Gaussian filter sizes:**

  Firstly, we have considered 100 features and experimented with different Gaussian filter sizes in the pre-processing step, including 3×3,5×5 and 7×7. Table 2 represents a comparison of performance with different Gaussian filter sizes. This table reveals that the logistic-regression model exhibits the least performance, and the XGBoost model attains the highest performance among the basic machine-learning models. However, the proposed stacking-based ensemble model outperforms basic models with 94.5%, 96.2%, and 93.4% with 3×3,5×5, and 7×7 filter sizes, respectively. It also indicates that the proposed stacking-based ensemble model degrades its performance with a 7×7 filter size. The proposed stacking classifier exhibits competing performance with Gaussian filter having 3×3 and 5×5 filter sizes. However, the proposed model with Gaussian filter having 5×5 filter size exhibits best performance.

Table 2. Ablation study with different Gaussian filter sizes.

	Filter size		
Model	3 × 3	5 × 5	**7** × 7
KNN	80.5	84.8	81.7
Logistic Regression	76.6	79.3	81.1
SVM	91.2	92.9	91.5
XGBoost	92.2	93.5	91.8
Voting Classifier	91.3	94.2	92.0
Stacking Classifier	**94.5**	**96.2**	**93.4**

- **With different number of features:**

  In general, the number of features plays a vital role in the performance of a model. Thus, we experimented with different numbers of features, including 100, 400, and 512, in the feature-reduction process. We analyzed the proposed model with Gaussian filter sizes 5×5 and 3×3 to understand how it behaves. Table 3 represents a performance comparison with different numbers of features having a Gaussian filter size of 5×5. In this case, the proposed stacking-based ensemble model exhibits the best performance with 400 features. Similarly, Table 4 represents a performance comparison with different numbers of features having a Gaussian filter size of 3×3. In this case, the proposed ensemble model exhibits the best performance at 98% with 400 and 512 features. Tables 3 and 4 reveal that the proposed stacking-based ensemble model exhibits the best performance with a Gaussian filter size of 3×3 with 400 and 512 features. The computational cost of the proposed model with 400 features is less than that with 512 features. Further, the proposed model exhibits similar performance in both cases. Hence, we have considered 400 features to optimize the proposed stacking-based ensemble model.

Table 3. Ablation study with different numbers of features having Gaussian filter size of 5×5.

	Number of features		
Model	**100**	**400**	**512**
KNN	84.8	78.4	77.2
Logistic Regression	79.3	89.3	89.6
SVM	92.9	94.2	94.1
XGBoost	93.5	93.2	96.3
Voting Classifier	94.2	97.3	**97.7**
Stacking Classifier	**96.2**	**97.7**	95.7

Table 4. Ablation study with different numbers of features having a Gaussian filter size of 3×3.

	Number of features		
Model	**100**	**400**	**512**
KNN	80.5	77.4	76.8
Logistic Regression	76.6	89.0	88.9
SVM	91.2	93.0	93.3
XGBoost	92.2	93.4	97.1
Voting Classifier	91.3	95.9	96.7
Stacking Classifier	**94.5**	**98.1**	**98.3**

4.2 Performance Comparison of Proposed Ensemble Model

The existing models utilized only 12 leads for the ECG classification. For more in-depth examination and to mimic clinical lead-based interpretation, ECG images were divided into 13 leads rather than the traditional 12-lead setup. We have considered popular machine-learning models for the performance analysis. In the proposed model, we utilized a Gaussian filter with a 3×3 kernel filter size in pre-processing and PCA with 400 features for feature reduction. The proposed ensemble model also consists of an ensemble of SVM with $\mathrm{C} = 1$, SVM with Gamma $= 0.01$, KNN with 5 neighbors, and Random Forest with 300 trees. Table 5 lists out class-wise performance of proposed stacking classifier.

Table 5. Class-wise performance of proposed stacking classifier.

Class	**Precision**	**Recall**	**F1-Score**
0	100.00	100.00	100.00
1	100.00	100.00	100.00
2	97.95	95.79	96.80
3	93.52	96.54	94.87
Accuracy	98.06		
Weighted Avg.	98.17	98.06	98.06

Table 6. Model performance comparison with ML models.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
KNN	77.36	77.79	77.37	75.94
Logistic Regression	88.69	89.47	88.69	88.19
SVM	93.10	93.27	93.10	93.02
XGBoost	93.32	93.56	93.32	93.24
Voting Classifier	95.47	95.85	95.47	95.36
Stacking Classifier	98.06	98.17	98.06	98.06

Table 6 lists out performance comparisons of stacking-based ensemble models with basic machine-learning models. Among the individual classifiers, SVM and XGBoost were the best performers, with over 93% accuracy, followed by logistic regression in the third place with 88.69%. KNN was the worst performer with lower metrics owing to its vulnerability to noisy and high-dimensional data. The SVM model gave consistent classification accuracy for each of the four classes of ECG with a total accuracy of 93.10%. Weighted precision, recall, and F1-score were 93.27%, 93.10%, and 93.02%, respectively. The XGBoost classifier achieved an accuracy of 93.32%, weighted precision of 93.56%, recall of 93.32%, and F1-score of 93.24%. Its gradient-boosting mechanism allowed it to handle non-linear decision boundaries well, since it could capture the fine changes of ECG waveform features. However, the proposed ensemble model performs better on all measures, with an accuracy rate of above 95%. The voting classifier achieved 95.47% accuracy that uses a soft vote algorithm on predictions from SVM, KNN and Random Forest. The stacking classifier, which uses a meta-model to integrate the predictions of different base learners, like SVM, KNN, and Random Forest, also achieved the greatest overall performance, with 98.06% accuracy and a balanced F1-score of 98.06%. Its hierarchical-learning framework enables the comprehension of intricate interactions across models and enhances predictive accuracy, rendering it the optimal selection for multi-class classification challenges. The data suggests that utilizing ensemble methods, especially stacking, improves the accuracy, reliability, and overall effectiveness of diagnosing cardiovascular problems through ECG.

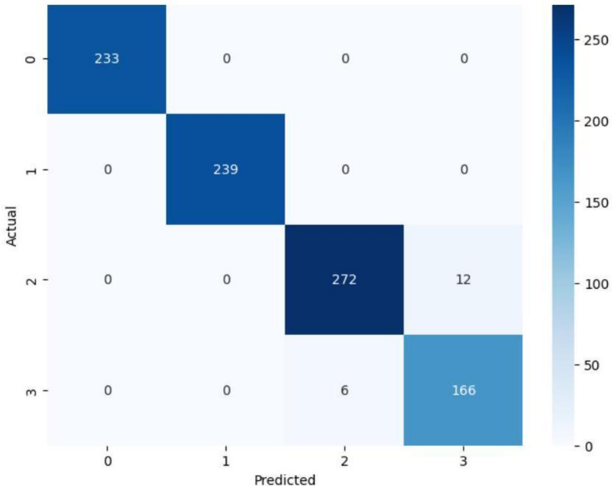


Figure 8. Confusion matrix of the proposed ensemble model.

The confusion matrix indicated that all four classes had more stability in their categorization. The ensemble mechanism provided the model with optimal confidence in recognizing borderline cases. The Abnormal and HMI classes exhibited a reduction in both false positives and false negatives. The results indicate that integrating various base learners enhances the system's resilience and reliability. Figure 8 illustrates the confusion matrix for the proposed stacking-based ensemble model. The confusion matrix indicates that the stacking classifier effectively distinguished between the classes, exhibiting minimal errors. The stacking classifier effectively distinguished between the MI and HMI classes, demonstrating exceptional precision and recall, indicative of robust discriminatory capability. The stacking design exhibited superior performance due to its ability to use inter-model interactions.

Table 7. Performance comparison with state-of-the-art models.

Model	**Accuracy (%)**
MobileNetV2 Transfer Learning [1]	93.00
MobileNetV2 Fine Tuning [1]	95.00
VGG16 Transfer Learning [1]	91.00
VGG16 Fine Tuning [1]	95.00
Sakli et al. [7]	96.70
Proposed Voting Classifier	95.47
Proposed Stacking Classifier	**98.06**

The efficacy of the proposed stack-based ensemble model is compared with that of existing models, as illustrated in Table 7. The findings indicate that the proposed stack-based ensemble model surpasses current state-of-the-art models, achieving an accuracy of 98.06%.

4.3 Discussion

This paper presents an extremely accurate model for predicting cardiovascular disease (CVD) based on ensemble machine-learning models. Of all the classifiers examined, the best performance was exhibited by the stacking classifier at 98.06% accuracy. Compared to [1], which used computationally costly models, like MobileNetV2 (94%) and VGG16 (92%), our stacking classifier outperformed them, even though it used less computationally costly models. This result proves the feasibility of ensemble learning even without using computationally costly models. Reference [7] employed a range of standard machine-learning models, including KNN, logistic regression, XGBoost, and SVM, and achieved a best accuracy of 96.7%. In [8], there were only regular ML techniques tried out, and the highest documented accuracy was less than 92.4%. Reference [9] used the models mentioned above, such as SVM, KNN, RF, and logistic regression, achieving an accuracy up to 92.4% with ensemble learning using GridSearch to optimize.

These comparisons also highlight the fact that, although there are excellent deep-learning architectures, such as MobileNetV2 and VGG16, well-hyperparameterized ensemble machine-learning algorithms can provide similar or even superior performance without the need for deep neural networks. This not only makes our method accurate, but also lightweight, interpretable, and computationally efficient with a significant advantage in real-world deployments to resource-constrained environments. The performance summary shows that ensemble-learning methods are better than regular classifiers when it comes to detecting heart diseases using ECGs.

5. CONCLUSION

This paper presents a scalable framework for multi-class classification of cardiovascular diseases from ECG images. ECG images can be processed for multi-class heart-disease classification through better pre-processing, contour-based feature extraction, and an ensemble-learning pipeline. Our results indicate that the ensemble-stacking classifier significantly outperforms individual models and all the earlier published works. The stacking classifier, with an accuracy of 98.06%, not only performed better than traditional machine-learning models, but also deep learning-based classifiers, such as MobileNetV2 and VGG16. Compared with deep architectures, the improved performance and reduced computational load of our architecture render it highly suitable for real-world use in resource-constrained environments. Additional class-wise accuracy improvement, particularly for valuable classes, like MI and HMI, also renders the system more practical. Briefly, the work illustrates the ability of carefully crafted traditional and ensemble-learning approaches to state-of-the-art performance on the cardiac-disease detection from ECG images. Some potential areas of future extensions of this paper include its application in real time on edge hardware, integration in clinical decision-support systems, or multi-modal health data to more general diagnostic applications.

REFERENCES

[1] L. Mhamdi, O. Dammak, F. Cottin and I. B. Dhaou, "Artificial Intelligence for Cardiac Diseases' Diagnosis and Prediction Using ECG Images on Embedded Systems," Biomedicines, vol. 10, no. 8, 2022.

[2] M. T. Tahmid, M. E. Kader, T. Mahmud and S. A. Fattah, "MDCardioNet: A Multi-dimensional Deep Neural Network for Cardiovascular Disease Diagnosis from Electrocardiogram," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 4, pp. 2005-2013, April 2024.

[3] M. B. Abubaker and B. Babayigit, "Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods," IEEE Transactions on Artificial Intelligence, vol. 4, no.2, pp. 373-382, April 2023.

[4] N. Mitra and B. I. Morshed, "Analyzing Clinical 12-lead ECG Images Using Deep Learning Algorithms for Objective Detection of Cardiac Diseases," Proc. of the 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 0517-0523, New York, USA, 2022.

[5] M. Balipa, M. Murugappan and A. Castalino, "Artificial Intelligence-based ECG Analysis to Assess Cardiac Health," Proc. of the 2024 3rd Int. Conf. on Sentiment Analysis and Deep Learning (ICSADL), pp. 587-593, Bhimdatta, Nepal, 2024.

[6] M. Gulhane and S. Kumar, "Deep Learning based Heart Diseases Detection Using ResNet50 Architecture," Proc. of the IEEE 2023 Intelligent Methods, Systems and Applications (IMSA), pp. 193-198, Giza, Egypt, 2023.

[7] M.Sakli, N. Sakli and H. Sakli, "ECG Images Automated Diagnosis based on Machine Learning Algorithms," Proc. of the 2023 20th Int. Multi-Conf. on Systems, Signals & Devices (SSD), pp. 934-939, Mahdia, Tunisia, 2023.

[8] M. D. Gresa and V. Sathya, "Exploring the Potential of Machine Learning and Deep Learning in ECG Image Analysis for Cardiovascular Disease Diagnosis," Proc. of the 2024 5th Int. Conf. on Electronics and Sustainable Communication Systems (ICESC), pp. 13441349, Coimbatore, India, 2024.

[9] D. Rautela, M. Bajeli, A. Kumar and H. Vaidya, "Identifying Cardiovascular Disorders through ECG Image Analysis," Proc. of the 2024 Int. Conf. on Cognitive Robotics and Intelligent Systems (ICC-ROBINS), pp.607-612, Coimbatore, India,2024.

[10] K. Bhangale et al., "Machine Learning-based Heart Disease Prediction Using ECG Image," Proc. of the 2024 5th Int. Conf. for Emerging Technology (INCET), pp. 1-9, Belgaum, India, 2024.

[11] I. Farady, V. Patel, C. -C. Kuo and C. -Y. Lin, "ECG Anomaly Detection with LSTM-Autoencoder for Heartbeat Analysis," Proc. of the 2024 IEEE Int. Conf. on Consumer Electronics (ICCE), pp. 1-5, Las Vegas, NV, USA, 2024.

[12] M. L. Meghana et al., "Cardiovascular Disease Detection in ECG Images Using CNN-BiLSTM Model," Proc. of the 2024 IEEE Int. Conf. on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), pp. 1-6, Bangalore, India, 2024.

[13] P. G. Aublin, J. Felblinger and J. Oster, "A Generalizable Heartbeat Classifier Leveraging Self-supervised Learning for ECG Analysis during Magnetic Resonance Imaging," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 9, pp. 5147-5155, Sept. 2024

[14] R. Krishna Priya, L. Alias and F. S. S. Al Salehiya, "Cardiac Health Assessment through Advanced Computational Models for ECG Image Analysis," Proc. of the 2024 10th Int. Conf. on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 2024.

[15] Z. M. Tun and M. A. Khine, "Cardiac Diagnosis based ECG Images Classification System Using Convolution Neural Network," Proc. of the 2023 IEEE Conf. on Computer Applications (ICCA), pp. 387-392, Yangon, Myanmar, 2023.

[16] S. Krishnakumar et al., "Detection of Arrhythmia and Congestive Heart Failure Through Classification of ECG Signals Using Deep Learning Neural Network," Proc. of the 2021 Int. Conf. on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021.

[17] M. Moutaib, M. Fattah, Y. Farhaoui and B. Aghoutane, "Machine Learning in Fetal Health: Improving ECG Analysis with Random Forest," Proc. of the 2024 Int. Conf. on Circuit, Systems and Communication (ICCSC), pp. 1-5, Fes, Morocco, 2024.

[18] H. R. Esther T. et al., "Identification of Cardiovascular Disease Using ECG Images Based on Deep Learning Procedure," Proc. of the 2023 Int. Conf. on Advances in Computing, Communication and Applied Informatics (ACCAI), pp. 1-6, Chennai, India, 2023.

[19] J. E. Madias, "The 13th Multiuse ECG Lead: Shouldn't We Use It More Often, and on the Same Hard Copy or Computer Screen, As the Other 12 Leads?," Journal of Electrocardiology, vol. 37, no. 4, pp. 285-287, 2004.

[20] A. H. Khan and M. Hussain, "ECG Images Dataset of Cardiac Patients," Mendeley Data, V2, DOI: 10.17632/gwbz3fsgp8.2, 2021.

[21] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th Edn., ISBN 978-0-13-335672-4, Pearson Education, 2018.

[22] C. Y. Saw and Y. C. Wong, "Neuromorphic Computing Based on Stochastic Spiking Reservoir for Heartbeat Classification," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 8, no. 2, pp. 182-193, 2022.

ملخص البحث:

تُعدُّ أمراض القلب في طليعة أسباب الوفيات على مستوى العالم، وتتطّلب تشخيصاً دقيقاً ومبكّراً. نقدّم في هذه الدّراسة نظاماً لتصنيف أمراض القلب للاستفادة منه في تشخيص تلك الأمراض بالاعتماد على صُور تخطيط القلب وتقنيات التّعلُّم الآلي. تتألف الشبكة المقترحة من ثلاث خطواتٍ، تشمل: المعالجة الأوّلية، واستخلاص السِّمات، والتّعلُّم الجماعي.

في البداية تخضع صورة تخطيط القلب إلى معالجةٍ أوّليةٍ شاملة. يلي ذلك استخلاص السِّمات من الصّورة ومن ثم تقليل عددها باستخدام تحليل المكوِّنات الرئيسية (PCA) للحفاظ على السِّمات المميِّزة والمعلومات المهمّة المتضمَّنة في الصّورة. وفي نهاية المطاف، يتمّ استخدام نظامٍ مجمَّع من عددٍ من نماذج التّعلُّم الآلي لتحسين أداء إطار العمل المقترح استناداً إلى طريقتي الانتخاب والتّرزيم.

وقد جرى تقييم النِّظام المقترح على مجموعة بياناتٍ عامّةٍ تشتملُ على صُور لتخطيط القلب، تمّ تصنيفها إلى أربع فئاتٍ: صُور طبيعية، وصُور غير طبيعية، وصُور تشير إلى نوبة قلبية، وصُور تشير إلى تاريخٍ من الإصابة بالنّوبات القلبية. وقد حقّق النّموذج المقترح أعلى دِقّة تصنيف تجاوزت 98% متفوِّقاً على النّماذج المماثلة الواردة في أدبيات الموضوع.