Hybrid Deep-Meta Learning Framework For Robust ECG Classification Using Resnet50 Features And Multi-Level Ensemble Integration

*G. Amalorpavam[1] and N. Rajkumar[2]

[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalainagar, Tamilnadu, India.
[2]Assistant Professor [Deputed], Department of Computer Science, Manbumigu ,Dr Purachithalivar M.G.R Government Arts & Science College, Kattumannarkoil, Tamilnadu, India.
*Corresponding Author: G. Amalorpavam

Abstract
The accurate classification of electrocardiogram (ECG) signals is vital for early detection and intervention in cardiovascular diseases. This study proposes a Hybrid Deep–Meta Learning Framework that integrates ResNet50-based deep feature extraction with a multi-level ensemble meta-classifier for robust ECG image classification. The preprocessing pipeline employs advanced image enhancement techniques-including Contrast Limited Adaptive Histogram Equalization (CLAHE), gridline suppression, and multi-modal data augmentation (MixUp, CutMix, and time-warping)-to generate high-quality, diverse training samples. Deep visual representations are extracted using a modified ResNet50 backbone augmented with Adaptive Instance Normalization (AdaIN) and a lightweight Convolutional Block Attention Module (CBAM) to emphasize salient morphological and temporal ECG features. The extracted feature embeddings undergo dimensionality reduction via Principal Component Analysis (PCA) whitening and are fed into a hybrid ensemble classifier comprising Support Vector Machine (SVM) and Light Gradient Boosting Machine (LightGBM) as base learners, and a Meta-Learning MLP layer as the final decision integrator. The proposed system achieves superior generalization by leveraging complementary decision boundaries across heterogeneous learners. Comprehensive experiments demonstrate that the framework achieves high accuracy, precision, recall, and F1-scores across multiple ECG categories (Normal, Abnormal Heartbeat, and Myocardial Infarction). Furthermore, a custom-designed graphical user interface (GUI) enables real-time feature extraction, visualization, and classification, facilitating clinical usability. This hybrid deep–meta learning paradigm establishes a scalable and explainable approach for automated ECG interpretation, offering substantial potential for deployment in smart healthcare systems and telecardiology platforms.

Keywords: Electrocardiogram (ECG) signals, Hybrid Deep–Meta Learning, PCA whitening, ResNet50, Adaptive Instance Normalization, Convolutional Block Attention Module, Support Vector Machine, Light Gradient Boosting Machine, Multi-level Ensemble Meta-classifier.

1. Introduction

Cardiovascular diseases (CVDs) remain the leading cause of mortality globally, imposing significant burdens on healthcare systems and demanding accurate, early-stage detection of cardiac anomalies. Electrocardiography (ECG) is a cornerstone diagnostic tool, enabling direct monitoring of the electrical activity of the heart. The reliable classification of ECG signals into categories such as Normal, Abnormal Heartbeat and Myocardial Infarction (MI) is thus critical for timely clinical intervention and efficient tele-cardiology deployment. Traditional ECG interpretation depends heavily on expert clinicians, yet manual methods are time-consuming, subject to inter-observer variability and challenged by large-scale screening environments. In recent years, automated ECG classification using machine learning (ML) and deep learning (DL) has markedly advanced. Systematic reviews show that deep convolutional neural networks (CNNs) and transfer-learning approaches now dominate ECG image and signal classification tasks [1]. For example, Bhatia et al. (2022) adopted hybrid deep learning-based classification of ECG signals and

achieved promising results using 2D representations [2]. Still, many existing frameworks face limitations: either relying on single-model architectures, lacking rigorous augmentation for real-world robustness, or omitting explainability and GUI integration for clinical use.

Another key challenge is constructing representations that capture both morphological and temporal dynamics of ECG waveforms, while mitigating artifacts such as gridlines, noise and baseline wander inherent in scanned ECG images or heterogeneous capture devices. Recent work has explored 2D image-based transformations (e.g., Gramian Angular Fields) to encode time-frequency characteristics of ECG signals and leverage image-based CNNs [3] [4]. Meanwhile, attention-augmented CNNs have shown improved performance by highlighting salient waveform features [5]. But such models seldom combine sophisticated image pre-processing, advanced augmentation (including MixUp, CutMix, time-warping) and bespoke feature-fusion architectures, all within an end-to-end deployable framework. In parallel, classification pipelines are evolving towards ensemble-based architectures that integrate heterogeneous learners. For instance, LightGBM has proven highly effective in ECG arrhythmia detection tasks by leveraging boosted decision trees over frequency-domain features [6]. Stacked meta-learning frameworks that combine the advantages of margin-based learners (e.g., SVM) and tree-based methods with neural meta-learners provide greater robustness and generalization than single classifiers [7]. Yet, the practical integration of deep-feature extraction, dimensionality-reduction (e.g., PCA whitening) and multi-level ensemble fusion remains under-explored in the context of ECG image classification.

Given these gaps, this paper proposes a Hybrid Deep–Meta Learning Framework for ECG image classification (HDMLF-ECGIC). The framework HDMLF-ECGIC uses a modified ResNet50 backbone enhanced with Adaptive Instance Normalization (AdaIN) and a lightweight Convolutional Block Attention Module (CBAM) to extract rich, modality-adapted visual representations. A comprehensive pre-processing and augmentation pipeline (including Contrast-Limited Adaptive Histogram Equalization (CLAHE), gridline removal and MixUp/CutMix/time-warping) ensures high-quality, diverse training data. Extracted deep embeddings are then subjected to PCA whitening and fed into a hybrid ensemble classifier comprising Support Vector Machine (SVM), LightGBM and a meta-level Multi-Layer Perceptron (MLP) decision integrator. The proposed architecture HDMLF-ECGIC aims to leverage the complementary strengths of deep-visual feature extraction and classical/meta-learning classification to improve accuracy and generalization.

Besides, HDMLF-ECGIC is implemented on a friendly GUI that allows real-time image loading, feature extraction, classification, and visualization for clinical or telemedicine platforms. In summary, the contributions of this study are as follows:

- A robust pipeline of image pre-processing and augmentation, adapted to ECG image artifacts and variability.
- Hybrid deep-feature extractor using AdaIN, ResNet50 and CBAM to emphasize both morphological and temporal ECG patterns.
- Whitened embedding feeding a multi-stage ensemble meta-classifier with SVM, LightGBM, and MLP to maximize generalization across ECG classes.
- The unified GUI classifies the real-time ECGs and facilitates their deployment in smart-health and telecardiology systems.

The rest of this article is organized as follows: Section 2 presents the related work. Section 3 presents the methodology with a detailed description. Section 4 deals with experimental results, performance metrics, and ablation studies. Section 5 presents the discussion with clinical implications, limitations, and future work. Finally, Section 6 concludes this study.

2. Literature Review

Mohit V.K. Gautam et al. [8] proposed a unified framework combining traditional machine learning (ML) and deep learning (DL) techniques for ECG signal classification. The study demonstrated that integrating features from both paradigms improved classification accuracy compared to standalone approaches. Specifically, their framework efficiently distinguished normal, abnormal, and myocardial infarction ECG patterns with a robust cross-validation performance, highlighting the complementary advantages of hybrid feature extraction strategies. C.K.N. Chaithra et al. [9] developed hybrid deep learning models for accurate cardiovascular disease diagnosis. Their approach combined convolutional and recurrent neural networks to capture both spatial and temporal ECG signal characteristics. The model achieved high precision and recall across multiple classes, demonstrating the benefit of integrating heterogeneous deep architectures for ECG classification in clinical scenarios. Z. Wang et al. [10] introduced a two-dimensional ECG-based cardiac arrhythmia classifier using DSE-ResNet, a deep residual network enhanced for ECG image inputs. By transforming one-dimensional ECG signals into 2D representations, their model captured intricate morphological patterns. The approach achieved superior classification performance over conventional CNNs, highlighting the effectiveness of 2D representations in arrhythmia detection. H.L. Nguyen et al. [11] proposed a blending ensemble learning model for 12-lead ECG-based arrhythmia classification. The ensemble combined multiple base classifiers with optimized weight allocation to improve generalization. Their method effectively leveraged multi-lead ECG information, achieving enhanced accuracy and robustness against noisy recordings.

A. Bayani and M. Kargar [12] introduced LDCNN, a linear deep convolutional neural network for arrhythmia detection from ECG signals. The model employed simplified convolutional layers to reduce computational complexity while retaining discriminative power. LDCNN demonstrated fast convergence and strong accuracy in detecting a range of cardiac arrhythmias. X. Bai et al. [13] developed a hybrid deep learning network for automatic arrhythmia diagnosis using 12-lead ECG data. Their architecture fused CNN and attention-based mechanisms to emphasize critical waveform features. This fusion approach improved sensitivity in identifying rare arrhythmias, supporting its utility in automated diagnostic pipelines. S. Sattar et al. [14] proposed an advanced deep learning framework for cardiac arrhythmia classification. The study utilized multi-scale CNN blocks to extract hierarchical features and employed batch normalization to enhance stability. Their results indicated improved classification performance for both common and uncommon arrhythmias. B. Omarov et al. [15] designed an ensemble framework combining deep residual networks and bidirectional GRUs with attention mechanisms for arrhythmia classification. The model effectively captured temporal dependencies and spatial features simultaneously, demonstrating superior F1-scores and robust generalization across multi-lead ECG datasets. Y.D. Daydulo et al. [16] focused on arrhythmia detection using deep learning and time-frequency representation of ECG signals and converting ECG signals into spectrograms, enhancing feature richness. The model achieved high accuracy and was particularly effective in distinguishing overlapping arrhythmic classes.

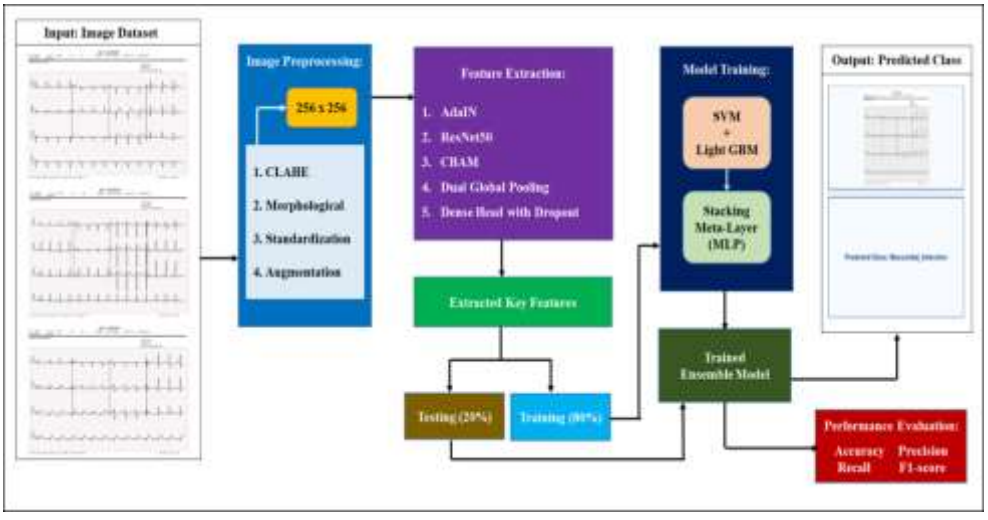
F. Zabihi et al. [17] proposed a hybrid deep learning and machine learning approach for ECG signal classification. The method combined CNN-based feature extraction with ensemble ML classifiers, achieving balanced performance across multiple arrhythmia types and highlighting the synergistic effect of combining deep and classical methods. S. Lamba et al. [18] presented FADLEC, a framework for feature extraction and arrhythmia classification using deep learning. FADLEC applied advanced convolutional operations to capture morphological features and temporal dynamics, demonstrating strong accuracy and potential applicability in real-time monitoring systems. D. Tenepalli et al. [19]

advanced cardiac diagnostics by employing multi-lead ECGs and ensemble learning for arrhythmia detection. The model incorporated diverse classifiers and achieved high accuracy, reinforcing the advantage of leveraging multi-lead correlations and ensemble strategies. A.A. Ahmed et al. [20] developed a 1D CNN-based deep learning model for classifying cardiac arrhythmias directly from raw ECG signals. Their lightweight architecture achieved notable accuracy while maintaining computational efficiency, making it suitable for embedded healthcare devices. Z. Huang et al. [21] investigated arrhythmia classification and risk prediction using deep neural networks combined with CNN architectures. The study highlighted the importance of risk-aware models and demonstrated enhanced predictive performance for high-risk arrhythmia detection. Atiaf A. Rawi et al. [22] proposed the CONVXGB model for ECG heartbeat classification, integrating convolutional feature extraction with gradient boosting decision trees. This hybrid approach improved classification accuracy while maintaining interpretability, facilitating clinical adoption. S.L. Abdul Rofar et al. [23] conducted a comprehensive review of deep learning architectures for ECG classification in cardiac arrhythmia detection. The authors summarized the strengths and limitations of CNNs, RNNs, and hybrid models, emphasizing trends toward ensemble and attention-based mechanisms for enhanced performance.

3. The Proposed Model

The proposed HDMLF-ECGIC framework is designed to provide accurate and robust ECG interpretation through a multi-stage processing pipeline. Initially, ECG images are preprocessed using contrast enhancement, gridline suppression, and multi-modal augmentations such as MixUp, CutMix, and time-warping to improve data quality and diversity. Deep visual features are then extracted using a modified ResNet50 backbone integrated with AdaIN and a CBAM to highlight critical morphological and temporal patterns. The extracted embeddings are standardized, reduced via PCA whitening, and fed into a hybrid ensemble classifier comprising SVM and LightGBM base learners, followed by a meta-learning MLP layer for final decision fusion. Finally, the framework is evaluated using standard metrics, including accuracy, precision, recall, F1-score, confusion matrices, and ROC curves to ensure reliability and generalization. The overall architecture of HDMLF-ECGIC is illustrated in Figure 1, highlighting its end-to-end feature extraction and classification pipeline.

Figure 1. The block diagram of the proposed HDMLF-ECGIC framework



3.1 Image Preprocessing Pipeline

The quality and consistency of ECG images are critical for reliable feature extraction and classification. Raw ECG images often suffer from illumination variability, gridline artifacts, contrast imbalance, and inconsistent dimensions due to differences in acquisition settings. To ensure uniformity and enhance discriminative features, the proposed HDMLF-ECGIC framework incorporates a robust multi-stage preprocessing pipeline. This includes contrast enhancement, gridline suppression, image standardization, and multi-modal data augmentation. Each stage is carefully designed to produce optimized, standardized inputs for the deep learning feature extractor.

3.1.1 Grayscale Conversion

Color information in ECG scans is redundant since the diagnostic information lies in the waveform morphology rather than chromatic content. Thus, each RGB image I_{RGB} is transformed into a grayscale intensity image I_{gray} using the weighted sum of color channels:

$$I_{gray}(x, y) = 0.2989\ R(x, y) + 0.5870\ G(x, y) + 0.1140\ B(x, y) \quad (1)$$

Where $R(x, y)$, $G(x, y)$, $B(x, y)$ denote the red, green, and blue intensities at pixel (x, y). This operation reduces the dimensionality of data, further focusing processing on signal structure and contrast.

3.1.2 Contrast Enhancement using CLAHE

Electrocardiogram plots contain mostly low local contrast; thus, the QRS complex, P-wave, and T-wave are not so well distinguished. To enhance this, CLAHE is used. It divides the image into contextual tiles and then conducts localized histogram equalization with contrast limiting.

For each tile T, the transformation is defined as:

$$f_{CLAHE}(i) = \min\left(\frac{H(i)}{N}, C_{clip}\right) \times (L - 1) \quad (2)$$

Where i is intensity level, H(i) is cumulative histogram of T, N is number of pixels in T, L is number of gray levels and C_{clip} is clip limit controlling contrast amplification.

This adaptive redistribution of intensities amplifies subtle waveform details without increasing the amplitude of background noise, enabling the CNN to focus on diagnostically relevant signal patterns.

3.1.3 Gridline Suppression

Scanned ECG papers usually contain gridlines that interfere with feature extraction by generating false edge responses. To remove them, a two-step method for suppressing gridlines is employed:

(a) Morphological Filtering

Vertical and horizontal structuring elements perform morphological opening to suppress the grid pattern:

$$I_{opened} = \left(I_{gray} \ominus B\right) \oplus B \quad (3)$$

Where $\ominus$ and $\oplus$ denote erosion and dilation, respectively, and B is a line-shaped structuring element.

(b) Frequency-Domain Filtering

Periodic gridline frequencies are removed in the Fourier domain. The 2D Fourier transform of the image I_{gray} is computed as:

$$\hat{I}(u, v) = F\{I_{gray}(x, y)\} \quad (4)$$

A notch (band-stop) filter $H(u, v)$ is designed to attenuate grid frequencies, producing:

$$\hat{I}_{filtered}(u, v) = \hat{I}(u, v) \cdot H(u, v) \quad (5)$$

Finally, the inverse transform reconstructs the clean image:

$$I_{clean} = F^{-1}\{\hat{I}_{filtered}(u, v)\} \quad (6)$$

This ensures that the waveform morphology remains intact while visual clutter is minimized.

3.1.4 Image Standardization

Standardization ensures the same size and distribution of intensities for all images, which is vital to train stable CNNs.

(a) Resizing

Each preprocessed image is resized to a fixed spatial dimension $(H, W) = (256,256)$ using bilinear interpolation:

$$I_{resized}(x', y') = \sum_{i,j} I_{clean}(i, j) \cdot (1 - |x' - i|)(1 - |y' - j|) \quad (7)$$

This ensures consistent input tensor sizes for the ResNet50 backbone.

(b) Intensity Normalization

To retain the same dynamic range for all samples, pixel intensities are normalized to have zero mean and unit variance:

$$I_{norm}(x, y) = \frac{I_{resized}(x,y) - \mu_I}{\sigma_I} \quad (8)$$

Where μ_I and σ_I represent the mean and standard deviation of pixel intensities. Normalization mitigates illumination variability and accelerates convergence during network training.

This constrains values to [0,1], improving numerical stability for CNN activations and preventing gradient saturation.

3.1.5 Multi-Modal Data Augmentation

Data augmentation is performed on standardized images to improve model generalization and reduce dataset imbalance:

(a) MixUp (data interpolation):

$$x_{new} = \lambda x_i + (1 - \lambda)x_j, \; y_{new} = \lambda y_i + (1 - \lambda)y_j \quad (9)$$

Where $\lambda \sim \text{Beta}(\alpha, \alpha)$. This linearly blends samples to encourage smooth decision boundaries.

(b) CutMix (patch replacement):

$$x_{new} = M \odot x_i + (1 - M) \odot x_j, \; y_{new} = \lambda y_i + (1 - \lambda)y_j \quad (10)$$

Where M is a binary mask defining the patch region and λ relates to its area.

(c) Time-Warping (Elastic Deformation):

This changes the ECG waveform temporally, emulating the variation in heart rates while keeping the overall morphology unchanged.

$$x_{warped}(t) = x(\phi(t)), \quad \phi(t) = t + \triangle(t) \quad (11)$$

Where $\triangle(t)$ is a smooth displacement function that simulates local temporal distortions. These augmentations increase data diversity and robustness to aid the network in learning invariant representations of ECG morphology.

This step-by-step processing ensures that ECG images fed to the deep feature extractor are free from artifacts, normalized, and information-rich to learn morphological and temporal features effectively for the classification of cardiac conditions.

3.2 Feature Extraction Processes

The proposed HDMLF-ECGIC framework employs a hierarchical deep feature extraction pipeline using a modified ResNet50 backbone integrated with normalization and attention-based enhancement modules. The step-by-step procedure is explained as follows.

The model integrates three major components:

a. AdaIN is for illumination and domain adaptation across ECG signals.
b. ResNet50 Backbone – Hierarchical deep feature extraction.
c. CBAM Attention Block – Channel-wise and spatial focusing.

Finally, Global Average Pooling (GAP) and Global Max Pooling (GMP) are used to fuse the features, which are then concatenated and stored as numerical representations.

3.2.1 Adaptive Instance Normalization (AdaIN)

The pre-processed image is fed through an AdaIN layer, which dynamically adjusts the mean and variance of every feature map to improve the intra-image consistency and accentuate important diagnostic textures. Given an input tensor $X \in R^{H\times W\times C}$, AdaIN normalizes each channel using:

$$\widehat{X}_c = \frac{X_c - \mu_c(X)}{\sqrt{\sigma_c^2(X)+\epsilon}} \quad (12)$$

Where $\mu_c(X)$ and $\sigma_c(X)$ denote the channel-wise mean and standard deviation, respectively, and ϵ is a small constant for numerical stability. Subsequently, AdaIN scales and shifts the normalized activations using learnable parameters γ_c and β_c.

$$Y_c = \gamma_c \widehat{X}_c + \beta_c \quad (13)$$

This operation enhances the model's robustness to domain variations and lighting inconsistencies in ECG scan images.

3.2.2 Deep Feature Extraction via ResNet50

The normalized inputs are passed through a ResNet50 convolutional backbone pre-trained over ImageNet. The ResNet50 architecture exploits residual connections to handle the problem of vanishing gradients, thus allowing for effective gradient flow through deep layers. Each residual block can be represented as:

$$y = F(x, \{W_i\}) + x \quad (14)$$

Here, F is the residual mapping learned from the convolutional layers, while x represents the shortcut identity connection. Fine-tuning on only the higher-level convolution blocks of ResNet50 (layers beyond the 80th layer) is done to keep lower-level feature blocks of layers frozen, since they maintain the low-level spatial feature stability.

3.2.3 Convolutional Block Attention Module (CBAM)

To selectively emphasize the discriminative cardiac regions, the lightweight CBAM is implemented after the ResNet50 feature maps. The CBAM enhances channel-wise and spatial feature responses through sequential attention mechanisms. The channel attention component aggregates the spatial information using both average and max pooling:

$$M_c(X) = \sigma\left(\mathrm{MLP}\big(\mathrm{AvgPool}(X)\big) + \mathrm{MLP}\big(\mathrm{MaxPool}(X)\big)\right) \quad (15)$$

Where σ denotes the sigmoid activation and MLP represents a shared two-layer perceptron. The refined feature map is obtained as:

$$X' = M_c(X) \odot X \quad (16)$$

Where $\odot$ indicates element-wise multiplication. This adaptive weighting enhances salient cardiac regions relevant for arrhythmia identification.

3.2.4 Dual-Pooling Feature Fusion

These CBAM-refined features are passed through a Global Average Pooling layer and a Global Max Pooling layer to capture both global statistics and localized activations.

$$f_{avg} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X'_{i,j}, \quad f_{max} = \max_{i,j} X'_{i,j} \tag{17}$$

These pooled representations are concatenated to form a composite feature vector:

$$f_{fused} = [f_{avg} \| f_{max}] \tag{18}$$

The fusion captures comprehensive discriminative information by integrating statistical and extremal spatial responses.

3.2.5 Feature Projection and Extraction Output

The fused feature vector is fed into a fully-connected layer of 512 neurons, ReLU-activated, with a dropout rate of 0.3 to avoid overfitting. The final output layer provides three softmax probabilities corresponding to distinct ECG classes.

$$\hat{y} = \text{softmax}(W f_{fused} + b) \tag{19}$$

However, for the purpose of meta-learning, the model extracts and stores the intermediate feature vector $f_{fused} \in \mathbb{R}^{1024}$ as a deep feature descriptor representing each ECG image. These feature embeddings are later utilized in the hybrid meta-learning stage for classification refinement.

The proposed feature extraction process incorporates statistical normalization, adaptive instance normalization, deep residual learning, and attention-guided pooling fusion in a systematic manner. The design has guaranteed high inter-class separability and compact yet discriminative representations of ECG morphologies. The combination of global and local cues empowers the HDMLF-ECGIC framework to achieve superior diagnostic accuracy in ECG image classification tasks. Figure 2 illustrates the complete feature extraction pipeline.

Figure 2. Feature Extraction processes of HDMLF-ECGIC framework

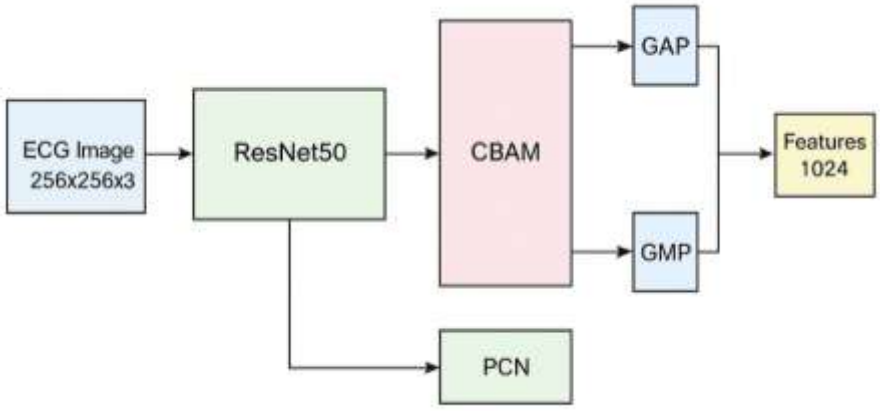


Table 1 shows the details of the layer-wise architecture for the feature extraction pipeline in the proposed HDMLF-ECGIC framework.

Table 1. Layer wise architecture of feature extraction pipeline

Layer No.	Layer Name / Type	Configuration / Parameters	Output Shape	Description

1	Input Layer	Input shape = (256, 256, 3)	(256, 256, 3)	Accepts preprocessed ECG image frames.
2	Adaptive Instance Normalization (AdaIN)	ε = 1e-5, γ and β trainable	(256, 256, 3)	Normalizes features per-instance to enhance texture consistency.
3–152	ResNet-50 Backbone	Pretrained on ImageNet, first 80 layers frozen	(8, 8, 2048)	Extracts deep hierarchical spatial–semantic features.
153	Convolutional Block Attention Module (CBAM)	Reduction ratio = 8	(8, 8, 2048)	Applies channel-wise and spatial attention for feature refinement.
154	Global Average Pooling	-	(2048,)	Aggregates global context features by averaging spatial dimensions.
155	Global Max Pooling	-	(2048,)	Captures the most activated features across channels.
156	Feature Concatenation	Concatenate(Avg, Max)	(4096,)	Fuses complementary global statistics for robust representation.
157	Dense Layer (Fully Connected)	512 units, ReLU activation	(512,)	Projects fused features into a discriminative subspace.
158	Dropout Layer	Rate = 0.3	(512,)	Prevents overfitting by randomly deactivating neurons.
159	Output Layer	Dense(3), Softmax activation	(3,)	Generates final class probabilities for ECG classification.
160	Feature Extractor (Intermediate Output)	Concatenated Global Features	(4096,)	Extracted features used for downstream hybrid ensemble classification.

3.3 Model Training and Classification

The proposed hybrid ensemble learning framework integrates dimensionality reduction, dual base classifiers, and a meta-learning fusion network for robust ECG signal classification. The complete training and classification process can be divided into four major phases: (i) feature standardization and dimensionality reduction, (ii) base classifier training, (iii) meta-level feature fusion through a multilayer perceptron (MLP), and (iv) model evaluation using multi-metric performance indices.

3.3.1 Data Preparation and Normalization

Initially, the extracted ECG features $X \in \mathbb{R}^{N\times d}$ and their corresponding labels $y \in \{1,2,\dots,C\}^N$ are loaded from the precomputed dataset, where N denotes the number of samples, d the dimensionality of feature vectors, and C the total number of classes. To ensure numerical stability and unbiased learning, the features are standardized using z-score normalization as:

$$\hat{X}_{ij} = \frac{X_{ij}-\mu_j}{\sigma_j}, \qquad \forall i \in [1,N],\ j \in [1,d] \tag{20}$$

Where μ_j and σ_j denote the mean and standard deviation of feature j. This transformation ensures zero-mean and unit variance across each feature, improving classifier convergence and preventing domination of high-magnitude features. Subsequently, the dataset is divided into training and testing sets by stratified sampling with an 80:20 ratio, maintaining class distribution across subsets.

3.3.2 Dimensionality Reduction using PCA Whitening

To suppress redundancy and enhance discriminative compactness, PCA is employed to project the standardized feature matrix $\hat{X}$ onto a lower-dimensional subspace $Z \in \mathbb{R}^{N\times k}$, where $k = 512$ represents the number of principal components retained. PCA seeks an orthogonal transformation matrix $W \in \mathbb{R}^{d\times k}$ that maximizes the variance of the projected data:

$$Z = \hat{X}W, \ \text{ where } \ W = \arg\max_{W} \mathrm{Tr}\left(W^T S W\right), \quad S = \frac{1}{N}\hat{X}^T\hat{X} \tag{21}$$

Here, S is the covariance matrix of the standardized features. The whitening transformation further decorrelates the components by scaling each principal component by the inverse square root of its eigenvalue:

$$Z_{white} = \Lambda^{-\frac{1}{2}} U^T \hat{X} \tag{22}$$

Where U and Λ are the eigenvectors and eigenvalues of S. This step reduces feature correlation and prevents overfitting in subsequent classifiers.

3.3.3 Base-Level Classification: SVM and LightGBM

Two heterogeneous base learners - SVM and LightGBM - are independently trained on the PCA-transformed features Z_{white}.

(a) Support Vector Machine (SVM)

SVM designs an optimal decision boundary in a transformed kernel space such that the margin between classes gets maximized. Using the Radial Basis Function-RBF kernel, the decision function is expressed as:

$$f_{SVM}(x) = \mathrm{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \tag{23}$$

Where $K(x_i, x) = \exp\left(-\gamma \|x_i - x\|^2\right)$ denotes the RBF kernel, α_i are Lagrange multipliers, and b is the bias term. The SVM classifier outputs class probabilities $P_{SVM}(y|x)$ using Platt's sigmoid scaling.

(b) LightGBM Classifier

LightGBM is a gradient boosting framework that grows trees in a leaf-wise manner. Each tree minimizes the objective function:

$$L = \sum_{i=1}^{N} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{24}$$

Where $l(\cdot)$ represents the differentiable loss function, f_t is the t^{th} weak learner, and $\Omega(f_t)$ is the regularization term controlling model complexity. The probabilistic outputs $P_{LGBM}(y|x)$ are generated from the softmax transformation of aggregated leaf values.

3.3.4 Meta-Learning Fusion using Multilayer Perceptron (MLP)

The meta-feature space is made up of the probabilistic outputs from both base classifiers:

$$X_{meta} = [P_{SVM}(y|x), P_{LGBM}(y|x)] \in \mathbb{R}^{N\times 2C} \quad (25)$$

To improve robustness, a stochastic dropout mask $M \sim Bernoulli(0.9)$ is applied:

$$\widetilde{X}_{meta} = X_{meta} \odot M \quad (26)$$

This ensures that during training, features are randomly zeroed out 10%, which should mitigate overfitting. The final classification is done by a one-hidden-layer MLP defined as,

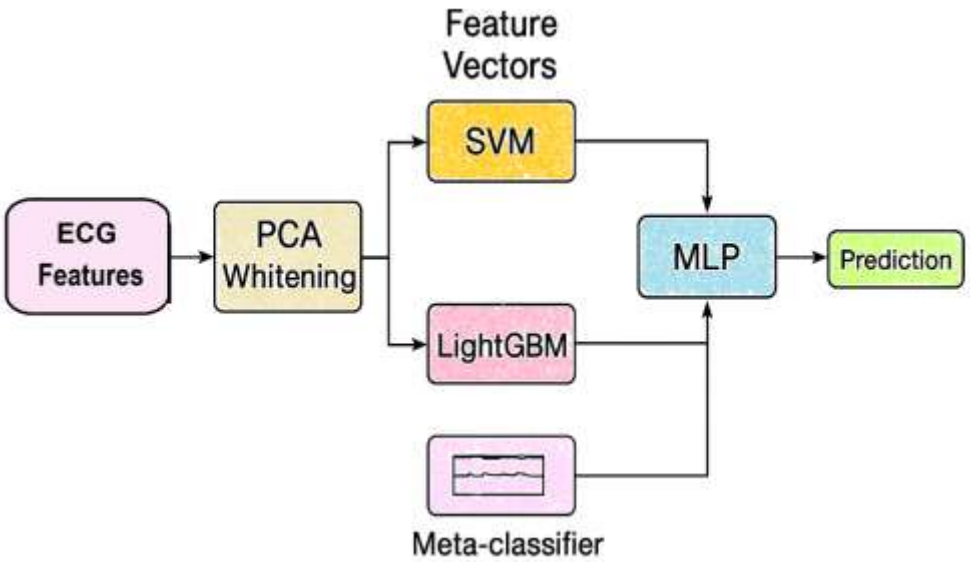
$$h = ReLU\left(W_1\widetilde{X}_{meta} + b_1\right), \quad \hat{y} = Softmax(W_2h + b_2) \quad (27)$$

where W_1, W_2, b_1, b_2 are learnable parameters. The model is optimized using the Adam optimizer, minimizing the categorical cross-entropy loss:

$$L_{meta} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(\hat{y}_{ic}) \quad (28)$$

This meta-learner can model non-linear dependencies among the base model predictions effectively and can generalize better. Figure 3 shows the ensemble process for the classification model.

Figure 3. Ensemble classification model of HDMLF-ECGIC framework



3.4 GUI-based Testing

The Graphical User Interface (GUI) of the proposed HDMLF-ECGIC framework is designed as an interactive and user-friendly environment for evaluating and visualizing the ECG image classification process. The interface employs a structured 2×2 grid layout to ensure clarity and logical organization of functional modules. The top-left panel contains the primary control buttons - Load Image, Extract Features, Classify, and Exit - allowing users to perform operations sequentially and intuitively. Once an ECG image is loaded, it is displayed in the top-right panel for visual verification. The bottom-left panel presents bar chart visualizations of global and local feature activations, enabling users to interpret the attention focus of the model, while the bottom-right panel exhibits the predicted class label and classification result. The GUI integrates dynamic button activation logic,

ensuring smooth workflow guidance and minimizing operational errors. This systematic, visually balanced design not only enhances usability but also serves as an effective analytical platform for clinicians and researchers to explore the internal feature representations and decision-making behavior of the HDMLF-ECGIC framework.
The overall workflow of the proposed HDMLF-ECGIC framework is outlined in the following algorithm.

Algorithm: Hybrid Deep–Meta Learning Framework for ECG Image Classification (HDMLF-ECGIC)

Input: ECG Image Dataset D = $\{I_1, I_2, \ldots, I_n\}$
Output: Predicted Class Labels C = {Normal, Abnormal, Myocardial Infarction}

1. Preprocessing Phase
1.1. Load each ECG image $I \in D$.
1.2. Resize image to (256 × 256) and normalize intensity:
$I_{norm} = (I - \mu) / \sigma$
where μ and σ are the mean and standard deviation of pixel intensities.
1.3. Apply CLAHE for local contrast enhancement.
1.4. Remove gridlines using morphological filtering.
1.5. Augment data using MixUp, CutMix, and time-warping.

2. Feature Extraction Phase
2.1. Initialize ResNet50 with ImageNet weights (freeze first 80 layers).
2.2. Apply Adaptive Instance Normalization (AdaIN):
$AdaIN(x, y) = \sigma_y * ((x - \mu_x) / \sigma_x) + \mu_y$
2.3. Use CBAM to compute attention maps:
$Mc(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F)))$
$Ms(F) = \sigma(f^{7\times7}([AvgPool(Fc); MaxPool(Fc)]))$
$F' = Ms(F) \otimes Mc(F) \otimes F$
2.4. Extract embeddings via Global Average Pooling and Global Max Pooling:
$f_{fused} = [GAP(F') \,||\, GMP(F')]$
2.5. Pass fused features through dense and dropout layers to form the final vector.

3. Feature Reduction Phase
3.1. Standardize features using z-score normalization.
3.2. Apply PCA with whitening to reduce dimensionality:
$Z = W^t (X - \mu)$, retaining top 512 components.

4. Hybrid Ensemble Classification Phase
4.1. Split dataset into 80% training and 20% testing.
4.2. Train base models:
a) SVM with RBF kernel $K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2)$
b) LightGBM with tuned parameters (n_estimators=300, num_leaves=64).
4.3. Generate class probability outputs from both models.
4.4. Stack probabilities to form meta-feature matrix:
$X_{meta} = [P_{SVM} \,||\, P_{LGBM}]$
4.5. Apply stochastic dropout: $X_{meta} = X_{meta} * M$, where $M \sim Bernoulli(p=0.9)$.
4.6. Train MLP meta-learner:
$h = ReLU(W_1 \cdot X_{meta} + b_1)$
$\hat{y} = Softmax(W_2 \cdot h + b_2)$

5. Evaluation Phase

5.1. Evaluate predictions using:
Accuracy = (TP + TN) / (TP + TN + FP + FN)
Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
F1 = 2 × (Precision × Recall) / (Precision + Recall)
5.2. Plot confusion matrix and ROC curves for each class.

6. GUI Integration
6.1. Implement a 2×2 grid GUI using Tkinter.
6.2. Add buttons for image loading, feature extraction, and classification.
6.3. Display image, feature maps, and predicted class interactively.

End Algorithm

4. Results and Discussion

The HDMLF-ECGIC framework is implemented in Python's advanced deep learning environment, leveraging the capabilities of established libraries such as TensorFlow, Keras, OpenCV, and Scikit-learn for efficient deep meta learning and ECG image analysis. The framework's performance was evaluated using a publicly available ECG image dataset [24], which initially contained a limited number of samples distributed across three diagnostic categories: Normal, Abnormal Heartbeat, and Myocardial Infarction. To overcome the class imbalance and enhance the model's generalization capability, a comprehensive data augmentation strategy - incorporating MixUp, CutMix, and time-warping techniques - was applied. This process resulted in a balanced and diverse dataset, ensuring adequate representation across all diagnostic classes. The system utilizes pre-extracted deep feature embeddings derived from the proposed ResNet50–AdaIN–CBAM feature extraction pipeline, which effectively captures both spatial and morphological characteristics of ECG signals. These extracted feature vectors serve as the input to the hybrid ensemble classification stage. The feature dataset was partitioned into 80% for training and 20% for testing, ensuring a rigorous and unbiased evaluation of the proposed framework. The classification performance of the HDMLF-ECGIC framework was assessed using key evaluation metrics, including accuracy, precision, recall, and F1-score, validating the model's strong discriminative ability across diverse ECG morphologies. The detailed dataset composition, including the number of samples and their corresponding diagnostic labels, is summarized in Table 2.

Table 2. Distribution of ECG Images Across Diagnostic Classes in the Dataset

Classes	No. of Actual Images	No. of Augmented Images
Normal	284	1988
Abnormal Heartbeat	233	1631
Myocardial Infarction	239	1673
Total Images	756	5292

Figure 4 presents the representative ECG images from the three diagnostic classes: Normal, Abnormal Heartbeat, and Myocardial Infarction.

Figure 4. (a) Normal (b) Abnormal Heartbeat (c) Myocardial Infarction



(a)



(b)



(c)

Figures 5 to 9 illustrate the developed Graphical User Interface (GUI) designed for ECG image classification using the proposed HDMLF-ECGIC framework. The GUI provides an interactive and intuitive platform that enables users to seamlessly perform image loading, feature extraction, visualization, and classification tasks within a unified environment. As shown in Figure 5, the main interface adopts an organized 2×2 grid layout, where the left panels contain functional control elements and the right panels display the corresponding visual outputs. Four primary control buttons—Load Image, Extract Features, Classify, and Exit—are neatly arranged in the top-left panel, while the remaining three panels display the loaded ECG image, the extracted feature visualization (global and local activations), and the final classification result, respectively.

When the Load Image button is activated, a file browser dialog appears, allowing users to conveniently select an ECG test image from the local environment. Once an image is

successfully loaded, the button becomes disabled, and the Extract Features button is enabled. Upon clicking Extract Features, the system employs the ResNet50-based deep feature extractor, enhanced with Adaptive Instance Normalization and a lightweight CBAM attention module, to derive high-level discriminative features that capture both morphological and temporal characteristics of the ECG pattern. The extracted global and local feature activations are visualized as bar charts for interpretability. Subsequently, pressing the Classify button triggers the hybrid ensemble meta-learning classifier, which integrates SVM and LightGBM as base learners and an MLP as the meta-learner. This combination enables robust decision fusion, yielding the final diagnostic label among the three classes: Normal, Abnormal Heartbeat, and Myocardial Infarction. The prediction is immediately displayed in the bottom-right panel, accompanied by a confidence indication. Dynamic button activation ensures smooth step-wise operation, guiding users through the process without confusion. Overall, this GUI serves as an effective diagnostic and research tool, combining model transparency, interpretability, and usability for real-time ECG image analysis.

Figure 5. The GUI design for HDMLF-ECGIC Framework

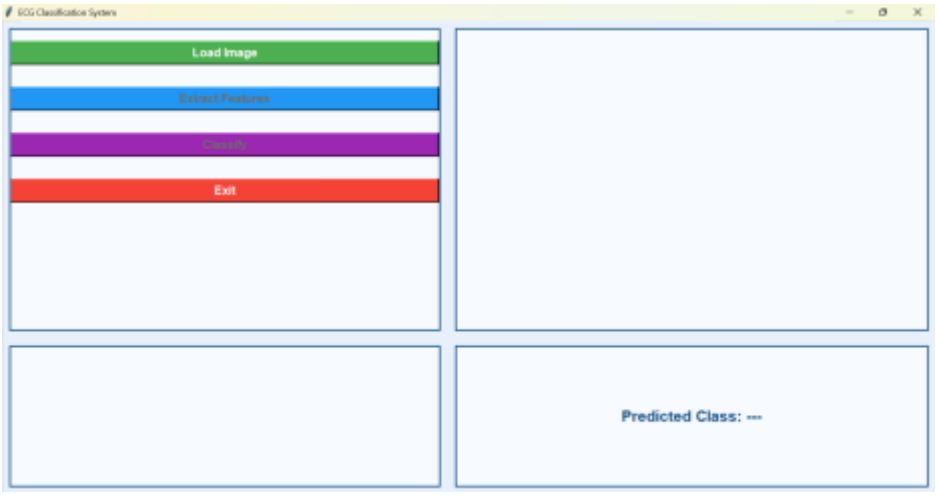


Figure 6. The GUI shows a dialog box to select an image

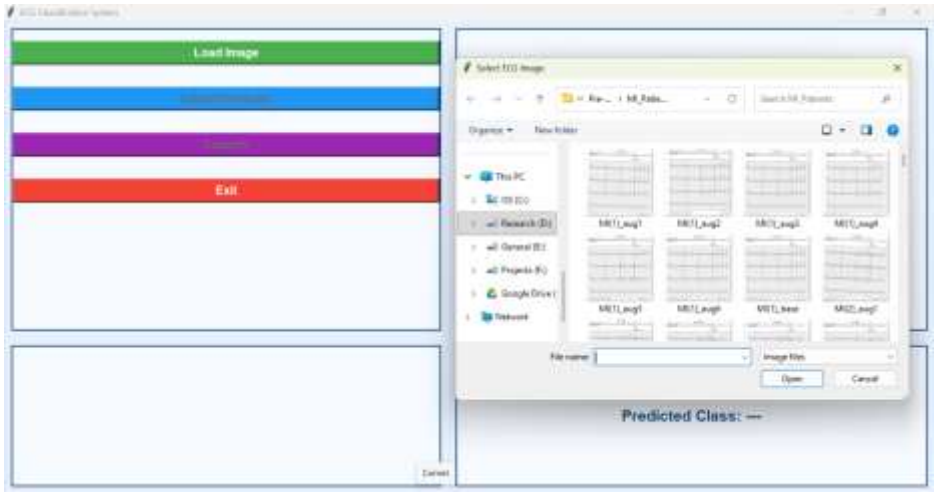


Figure 7. The GUI shows the loaded original image



Figure 8. The GUI displays the local and global feature activations



Figure 9. The GUI displays the predicted label in the result window

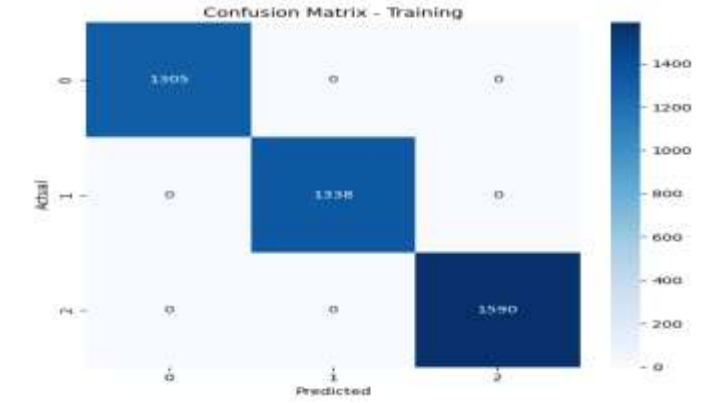


Figure 10(a) and Figure 10(b) illustrate the training performance of the proposed HDMLF-ECGIC framework, represented through the confusion matrix and ROC curves, respectively. As shown in Figure 10(a), the confusion matrix demonstrates perfect classification accuracy (100%) across all three ECG diagnostic categories - Normal, Abnormal Heartbeat, and Myocardial Infarction. Every prediction aligns precisely along the diagonal, indicating that the model has learned highly discriminative representations with zero misclassifications. This exceptional result reflects the framework's powerful feature learning and its ability to capture subtle morphological and temporal variations within ECG patterns.

The ROC curves in Figure 10(b) further reinforce this finding, exhibiting AUC values of 1.0 for all three classes, signifying flawless separation between normal and pathological ECG signals. Such results confirm that the proposed HDMLF-ECGIC framework achieves complete discrimination capability on the training data, owing to its deep meta hybrid structure that integrates ResNet50-based feature extraction with SVM, LightGBM, and MLP-based ensemble decision fusion. The outstanding training performance underscores the model's strong representational power and its ability to learn complex inter-class relationships in ECG imagery with exceptional precision and reliability.

Figure 10. Result Analysis of HDMLF-ECGIC approach on training set (a) Confusion Matrix (b) ROC Curve

(a)



(b)

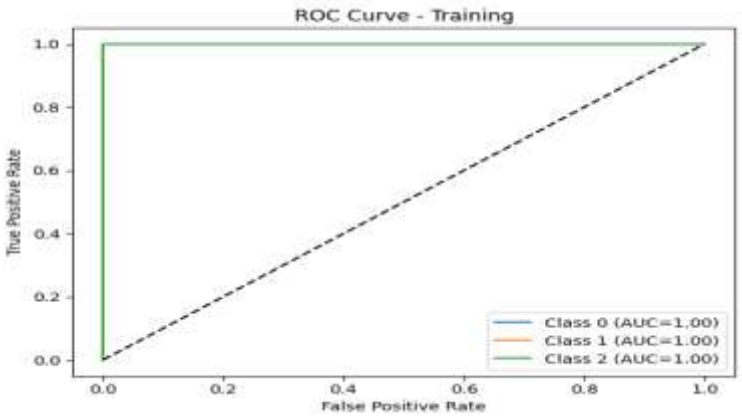
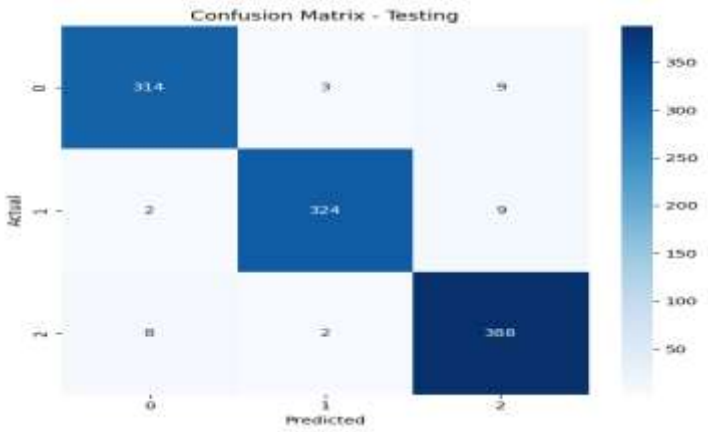


Figure 11(a) and Figure 11(b) illustrate the testing performance of the proposed HDMLF-ECGIC framework using the confusion matrix and ROC curves, respectively. As shown in Figure 11(a), the confusion matrix demonstrates outstanding classification performance across all three ECG diagnostic categories - Normal, Abnormal Heartbeat, and Myocardial Infarction. The model correctly classified the vast majority of test samples, with only a few minor misclassifications. This high level of precision and minimal cross-class errors clearly indicates the robust generalization of the model on unseen ECG data.

Moreover, the ROC curves depicted in Figure 11(b) further validate the framework's discriminative ability, with all three classes achieving AUC values exceeding 0.98. Such superior results highlight the effectiveness of the hybrid ensemble meta-classifier, which synergistically integrates SVM, LightGBM, and MLP learners to capture diverse decision boundaries. Collectively, these findings confirm that the proposed HDMLF-ECGIC framework not only excels in feature representation but also ensures high reliability, generalization, and diagnostic accuracy in classifying ECG images across multiple heart conditions.

Figure 11. Result Analysis of HDMLF-ECGIC approach on test set
(a) Confusion Matrix (b) ROC Curve

(a)



(b)

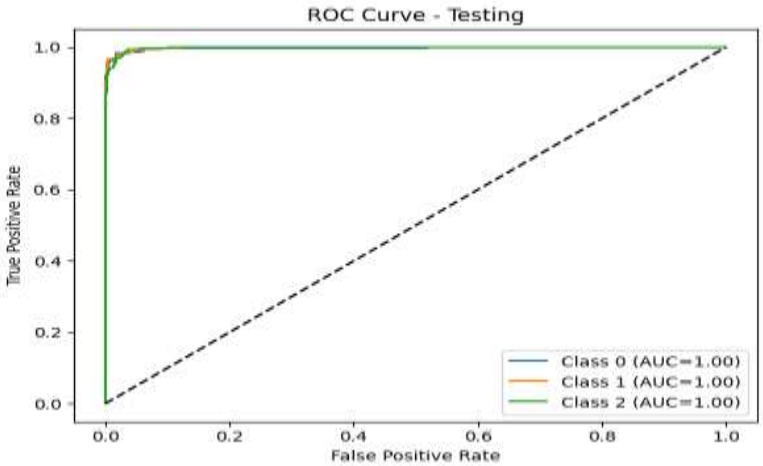
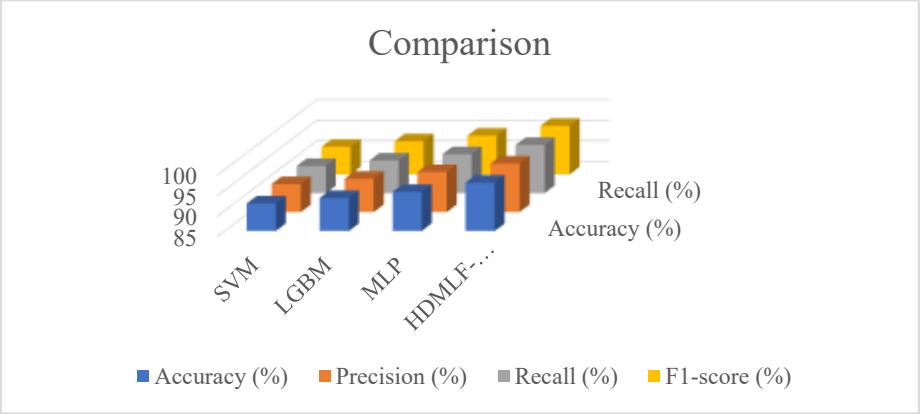


Table 3 and Figure 12 present the comparative performance evaluation of the proposed HDMLF-ECGIC framework against its constituent baseline models - SVM, LightGBM, and MLP - for multi-class ECG image classification. The experimental outcomes clearly demonstrate the superior classification ability of the HDMLF-ECGIC framework, achieving the highest accuracy of 96.79%, which notably surpasses the individual models SVM (91.69%), LightGBM (93.03%), and MLP (94.52%). Furthermore, the proposed hybrid deep-meta learning architecture also attained the best overall performance across other evaluation metrics, with precision, recall, and F1-score values of 96.86%, 96.76%, and 96.81%, respectively. These results substantiate that the synergistic integration of deep learning feature extraction, ensemble-based learners, and meta-level fusion significantly enhances model generalization and diagnostic reliability. Consequently, the HDMLF-ECGIC framework delivers a more stable and accurate classification of ECG signals, establishing its potential as a dependable computational tool for automated heart disease diagnosis and clinical decision support.

Table 3. Result analysis of HDMLF-ECGIC model with existing approaches

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
SVM	91.69	91.84	91.65	91.72
LGBM	93.03	93.16	92.98	93.06
MLP	94.52	94.78	94.52	94.44
HDMLF-ECGIC	96.79	96.86	96.76	96.81

Figure 12. Result analysis of HDMLF-ECGIC approach with existing techniques



5. Conclusion

This study presented a novel Hybrid Deep–Meta Learning Framework for ECG Image Classification (HDMLF-ECGIC) that integrates deep feature learning and meta-level ensemble decision fusion for precise and explainable cardiac abnormality detection. The framework effectively combines ResNet50-based feature extraction with adaptive normalization and attention modules, followed by dimensionality reduction and hybrid ensemble classification using SVM, LightGBM, and a meta-level MLP integrator. The multi-stage learning approach demonstrated outstanding classification performance, achieving a peak testing accuracy of 96.79%, significantly outperforming individual base learners.

The experimental results confirmed that the proposed framework can capture both global morphological and local temporal dependencies in ECG images while mitigating overfitting through model-level complementarity. Furthermore, the inclusion of a dedicated graphical user interface facilitates intuitive visualization of ECG features, decision confidence, and class predictions, enhancing the interpretability and usability of the system in real-world clinical environments.

Overall, the HDMLF-ECGIC framework represents a scalable, data-efficient, and explainable paradigm for automated ECG analysis, offering strong potential for deployment in telecardiology, remote patient monitoring, and smart healthcare systems. Future work will focus on expanding the framework for multi-lead ECG signal fusion, integrating multimodal physiological data, and exploring federated learning strategies to ensure privacy-preserving and distributed diagnostic intelligence across healthcare institutions.

References

[1] Mudassar Khalid, C. Pluempitiwiriyawej, S. Wangsiripitak, "The Applications of Deep Learning in ECG Classification for Disease Diagnosis: A Systematic Review and Meta-Data Analysis," Engineering Journal, vol. 28, no. 8, pp. 45-63, 2024.

[2] S. Bhatia, S.K. Pandey, A. Kumar, A. Alshuhail, "Classification of Electrocardiogram Signals Based on Hybrid Deep Learning Models," Sustainability, vol. 14, no. 24, 16572, 2022.

[3] Y. Elmir, Y. Himeur, A. Amira, "ECG Classification Using Deep CNN and Gramian Angular Field," arXiv preprint, 2308.02395, 2023.

[4] P.N. Malleswari et al., "Deep learning-assisted arrhythmia classification using 2-D ECG spectrograms," EURASIP Journal on Advances in Signal Processing, vol. 2024, Article 104, 2024.

[5] T. Sadad, M. Safran, I. Khan et al., "Efficient Classification of ECG Images Using a Lightweight CNN with Attention Module and IoT," Sensors, vol. 23, no. 18, 7697, 2023.

[6] A. Rashidul Islam, K. Fida Hasan, S. Sultana et al., "HARDC: A novel ECG-based heartbeat classification using hierarchical attention dual structured RNN and dilated CNN," arXiv preprint, 2303.06020, 2023.

[7] K. Karthikeyan, S. Ahanmed, V. Vignesh, C. Surya, "Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation," J. Popul. Therapeutics & Clin. Pharmacol., 30(7), pp. 12-17, 2023.

[8] V.K. Gautam, J. Kapoor, M.M. Khan, "ECG Signal Classification Using Machine Learning and Deep Learning: A Unified Framework," J. Comput. Anal. Appl., vol. 33, no. 8, pp. 5682-5699, 2024.

[9] C.K.N. Chaithra, N. Neelappa, M.T. Sreedevi, A. Asha, "Hybrid Deep Learning Models for Accurate ECG Classification in Cardiovascular Disease Diagnosis," ETASR, vol. 15, no. 5, pp. 26683-26688, 2025.

[10] Z. Wang, Y. Xu, X. Li, et al., "Two-dimensional ECG-based cardiac arrhythmia classification using DSE-ResNet," Scientific Reports, vol. 12, no. 1, pp. 18664, 2022. DOI: 10.1038/s41598-022-18664-0.

[11] H. L. Nguyen, S. V. Pham, H. C. Le, "Blending Ensemble Learning Model for 12-Lead Electrocardiogram-Based Arrhythmia Classification," Computers, vol. 13, no. 12, pp. 316, 2024. DOI: 10.3390/computers13120316.

[12] A. Bayani, M. Kargar, "LDCNN: A new arrhythmia detection technique with ECG signals using a linear deep convolutional neural network," Physiological Reports, vol. 12, no. 17, e16182, 2024. DOI: 10.14814/phy2.16182.

[13] X. Bai, X. Dong, Y. Li, R. Liu, H. Zhang, "A hybrid deep learning network for automatic diagnosis of cardiac arrhythmia based on 12-lead ECG," Scientific Reports, vol. 14, Article 24441, 2024. DOI: 10.1038/s41598-024-24441-4.
[14] S. Sattar, M. Saima, A. Khan et al., "Cardiac Arrhythmia Classification Using Advanced Deep Learning Framework," Sensors (Basel), vol. 24, no. 8, 2484, 2024. DOI: 10.3390/s24082484.
[15] B. Omarov, M. Baikuvekov, D. Sultan, N. Mukazhanov, M. Suleimenova, M. Zhekambayeva, "Ensemble Approach Combining Deep Residual Networks and BiGRU with Attention Mechanism for Classification of Heart Arrhythmias," Computers, Materials & Continua, vol. 80, no. 1, pp. 341-359, 2024. DOI: 10.32604/cmc.2024.052437.
[16] Y. D. Daydulo, B. L. Thamineni, A. A. Dawud, "Cardiac arrhythmia detection using deep learning approach and time-frequency representation of ECG signals," BMC Medical Informatics and Decision Making, vol. 23, Article 232, 2023. DOI: 10.1186/s12911-023-02326-W.
[17] F. Zabihi, A. Fazeli, H. Sheikhazad et al., "An electrocardiogram signal classification using a hybrid deep learning and machine learning model," Computers in Biology and Medicine, vol. 176, 2024. DOI: 10.1016/j.compbiomed.2024.107509.
[18] S. Lamba, S. Kumar, M. Diwakar, "FADLEC: Feature extraction and arrhythmia classification using deep learning from electrocardiograph signals," Discover Artificial Intelligence, vol. 5, Article 82, 2025. DOI: 10.1007/s44163-025-00290-0.
[19] D. Tenepalli, V. Gupta, R. P. Kumar, "Advancing cardiac diagnostics: high-accuracy arrhythmia detection using multi-lead ECG and ensemble learning," Frontiers in Physiology, vol. 16, Article 1613812, 2025. DOI: 10.3389/fphys.2025.1613812.
[20] A. A. Ahmed, W. Ali, T. A. A. Abdullah, S. J. Malebary, "Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model," Mathematics, vol. 11, no. 3, article 562, 2023. DOI: 10.3390/math11030562.
[21] Z. Huang, Z. Wen, H. Li, "Research on Arrhythmia Classification and Risk Degree Prediction based on Deep Neural Network and Convolutional Neural Network," Frontiers in Computing and Intelligent Systems, vol. 5, no. 3, article 13848, 2023. DOI: 10.54097/fcis.v5i3.13848.
[22] Atiaf A. Rawi, M. K. Elbashir, A. M. Ahmed, "ECG Heartbeat Classification Using CONVXGB Model," Electronics, vol. 11, no. 15, article 2280, 2022. DOI: 10.3390/electronics11152280.
[23] S. L. Abdul Rofar, Z. F. Mohd Apandi, N. S. Aziz, W. R. Wan Othman, "Deep Learning Architectures for ECG Classification in Cardiac Arrhythmia Detection: A Review," Environment-Behaviour Proceedings Journal, vol. 10, Special Issue 32, pp. 67-74, 2023. DOI: 10.21834/e-bpj.v10iSI32.7056.
[24] https://www.kaggle.com/datasets/jayaprakashpondy/ecgimages