10.48047/jocaaa.2024.33.08.207

ECG Signal Classification Using Machine Learning and Deep Learning: A Unified Framework

Viksit Kumar Gautam[a], Janak Kapoor[b], Mohd Mustafa Khan[c*],

[a]Department of Electronics and Communication Engineering, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, 243006, Uttar Pradesh, India; eviksit@gmail.com

[b]Department of Electronics and Communication Engineering, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, 243006, Uttar Pradesh, India; janak.kapoor@mjpru.ac.in

[c]Department of Computer Science and Information Technology, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, 243006, Uttar Pradesh, India; mustafamohd178@gmail.com

Abstract

The accurate classification of electrocardiogram (ECG) signals is vital for the early detection and diagnosis of cardiovascular diseases. This study evaluates and compares the performance of various Machine Learning (ML) and Deep Learning (DL) models in classifying ECG signals into normal and abnormal categories. A comprehensive set of models, including logistic regression, support vector machines (SVM), random forests, k-nearest neighbors (KNN), 1D/2D/3D convolutional neural networks (CNNs), denseNet, efficientNet, and long short-term memory (LSTM) networks, was implemented and tested on a benchmark ECG dataset. The models were assessed using standard metrics. Among all models, 1D-CNN and efficientNet achieved the highest classification accuracy (99.0%), closely followed by Random Forests and 3D-CNN (98.9%). The findings highlight the effectiveness of DL frameworks in capturing complex temporal patterns in ECG signals, making them suitable for real-time, automated cardiac diagnostics.

Keywords

Electrocardiogram (ECG), Machine Learning (ML), Deep Learning (DL), Signal Classification, Cardiovascular Disease, Automated Diagnosis

10.48047/jocaaa.2024.33.08.207

1. Introduction

The classification of ECG signals is essential for the prompt identification and treatment of cardiovascular diseases (CVDs), which continue to be the number one cause of death globally, accounting for nearly 18 million deaths annually [1]. With its widespread use, ECG facilitates the diagnosis of arrhythmia, myocardial infarction, and various other cardiac disorders, aiding in more than 70% of cardiac diagnoses worldwide [2]. Despite its clinical significance, conventional interpretation of ECGs still relies heavily on expert cardiologists, resulting in inconsistencies and interobserver variability, particularly in under-resourced clinical settings [3].

Advancements in artificial intelligence (AI), machine learning (ML), and deep learning (DL) have significantly enhanced the precision and reliability of ECG evaluation. These technologies enable efficient population screening and real-time health monitoring through wearable devices, leveraging the objective classification and strong signalprocessing capabilities of ML and DL models [4],[5]. As noted by [6], research in AI-based ECG classification has more than doubled in the past five years, underscoring the urgency for automation and rapid cardiac diagnosis.

2. Literature Survey

Recent advancements in technology have enhanced the classification of automated ECG signals. Arrhythmia detection systems were traditionally built with algorithmic techniques like decision trees, SVM, and k-NN. Unfortunately, these systems often relied on manually crafted features, and their signal dynamics' generalizability was poor [7], [8].

In response to these concerns, deep learning (DL) models have begun to develop rapidly. Models based on CNNs, RNNs, and LSTM networks have consistently proven to be more effective than traditional ML algorithms and can now directly extract relevant spatiotemporal features from raw ECG data [9], [10], [11].

The classification accuracy has increased with CNN-LSTM and residual networks because of their capability to spatially hierarchically and temporally decode dependencies [12], [13]. Noise reduction is achieved through preprocessing techniques like wavelet transform, empirical mode decomposition, and baseline wander correction, which are integrated into model pipelines [12], [11]. Heterogeneous ECG datasets have shown greater adaptability and robustness with ensemble and transfer learning [14], [15].

Other XAI techniques have also been developed that focus on the model’s visualization for borders and representations like Grad-CAM and SHAP that allow model-learned bounds and representational visualization [16], [17]. Addressing data imbalance, which is common in clinical ECG datasets, has been tackled with methods like the Synthetic Minority

10.48047/jocaaa.2024.33.08.207

Oversampling Technique (SMOTE), data augmentation, and class weighting. These approaches help improve how well models generalize [18], [19]. The existence of large, annotated, open-access datasets like PTBXL and MIT-BIH has also made it easier to reproduce results, conduct comparative evaluations, and speed up progress in ECG signal modeling [20], [21]. Additionally, there is increasing interest in using lightweight deep learning models in wearable and energy-efficient platforms for real-time cardiac monitoring [22], [23]. However, significant challenges remain, especially regarding cross-population generalization, multi-lead integration, and the clinical validation of predictive models [24].

3. Proposed Method

In this study, we propose a robust and comprehensive ECG signal classification framework that integrates both traditional ML algorithms and advanced DL architectures to achieve high classification performance. The ECG dataset used in this work comprises labeled signal segments corresponding to various cardiac conditions. Each ECG signal sample was pre-processed and formatted appropriately to suit both 1D signal input for CNN/LSTM models and 2D formats (e.g., spectrograms) for convolutional analysis. The complete flow of the proposed methodology is illustrated in Figure 1, which includes ECG signal acquisition, preprocessing, model training, and evaluation.

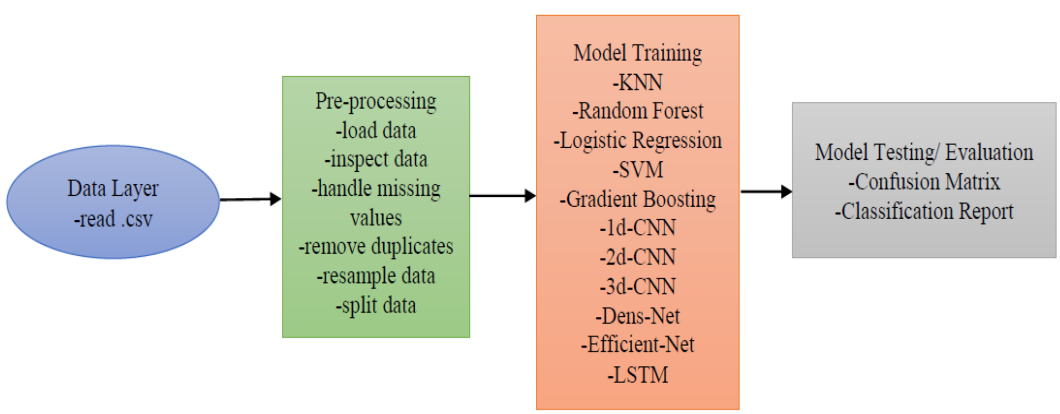

Figure 1: Process Flow.

10.48047/jocaaa.2024.33.08.207

3.1. Dataset Description

The ECG dataset consists of the ECG readings of patients. The dataset consists of 5 rows and 141 columns. Each row corresponds to a single complete ECG of a patient. Every single ECG is composed of 140 data points(readings). Columns 0-139 contain the ECG data point for a particular patient. These are floating point numbers and the label shows whether the ECG is normal or abnormal. It is a categorical variable with value either 0 or 1. Each

ECG signal is a time-series recording of electrical activity of the heart, sampled uniformly and segmented for classification. An example of the dataset structure and corresponding ECG waveform samples for each class

is provided in Figure 2, and Figure 3 respectively. For more details, refer to the dataset available at Kaggle.

	-0.11252183	-2.8272038	-3.7738969	-4.3497511	-4.376041	-3.4749863	-2.1814082	-1.8182865	-1.2505219	-0.47749208	...	0.79216787	0.93354122	0.7969577
0	-1.100878	-3.996840	-4.285843	-4.506579	-4.022377	-3.234368	-1.566126	-0.992258	-0.754680	0.042321	...	0.538356	0.656881	0.78749
1	-0.567088	-2.593450	-3.874230	-4.584095	-4.187449	-3.151462	-1.742940	-1.490659	-1.183580	-0.394229	...	0.886073	0.531452	0.31137
2	0.490473	-1.914407	-3.616364	-4.318823	-4.268016	-3.881110	-2.993280	-1.671131	-1.333884	-0.965629	...	0.350816	0.499111	0.60034
3	0.800232	-0.874252	-2.384761	-3.973292	-4.338224	-3.802422	-2.534510	-1.783423	-1.594450	-0.753199	...	1.148884	0.958434	1.05902
4	-1.507674	-3.574550	-4.478011	-4.408275	-3.321242	-2.105171	-1.481048	-1.301362	-0.498240	-0.286928	...	1.089068	0.983369	1.01412

5 rows × 141 columns

Figure 2: Represents the first 5 rows and the columns of the dataset.

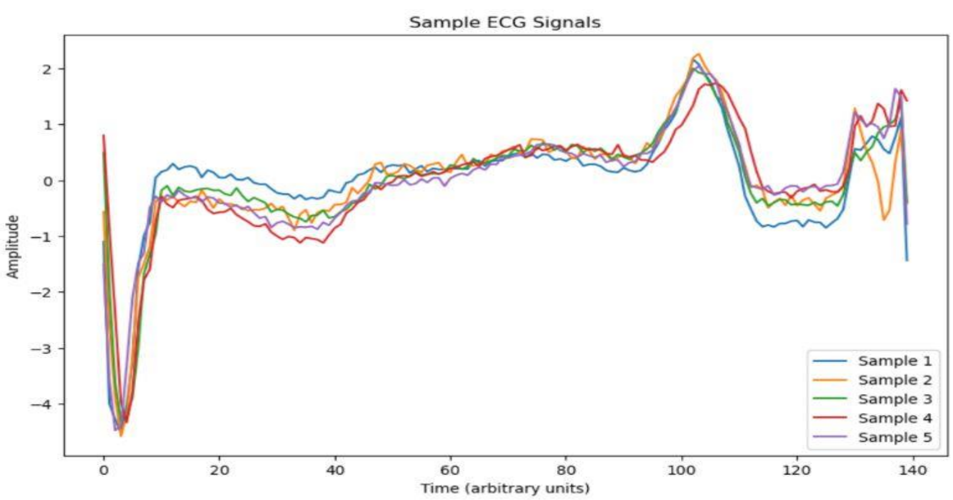

Figure 3: Sample ECG signals.

10.48047/jocaaa.2024.33.08.207

3.2. Traditional ML Algorithms

The proposed methodology focuses on the classification of ECG signals using a combination of traditional ML models and advanced DL architectures. The ECG dataset is first pre-processed to ensure consistency and remove missing or corrupted values. For all models, the raw ECG signals are either directly used or transformed into fixed-length vectors suitable for input into respective classifiers.

To classify ECG signals accurately, we implemented a variety of machine learning (ML) algorithms. Each method brings distinct strengths in pattern recognition and signal modeling. Below, we describe each algorithm along with its mathematical foundations.

3.2.1. (a) K-Nearest Neighbors (KNN)

KNN classifies a sample based on the majority label among its k nearest neighbours using Euclidean distance [25]. KNN is a non-parametric, instance-based learning algorithm that classifies an input sample based on the majority label among its k nearest neighbors. The distance metric is typically the Euclidean distance:

$$d(x, x_i) = \sqrt{\sum_{j=1}^{n} (x_j - x_{ij})^2} \quad (1)$$

where, $d(x, x_i)$ is the distance between point x and point x_i, x_j is the j-th feature (coordinate) of point x, x_{ij} is the j-th feature (coordinate) of point x_i, and n is the number of features (dimensions) .

The predicted label $y\hat{}$ is computed as:

$$\hat{y} = \arg\max_{c} \sum_{i \in N_k(x)} \mathbf{1}(y_i = c) \quad (2)$$

This formula finds the class c that is most common among the k nearest neighbors of input x. For each possible class c, it counts how many of the k nearest neighbors belong to class c. Then, it assigns to $y\hat{}$ the class c with the highest count.

10.48047/jocaaa.2024.33.08.207

3.2.2. (b) Random Forest (RF)

RF is an ensemble approach that constructs several decision trees and combines their outputs. Multiple decision tree outputs are aggregated to determine an RF classifier's prediction for a given input instance x [26]. In an ensemble classification model consisting of T base learners (e.g., decision trees), the final predicted label ŷ is obtained by majority voting over the individual predictions:

$$\hat{y} = \mathrm{mode}\left(\{h_t(x)\}_{t=1}^{T}\right) \tag{3}$$

where, $h_t(x)$ represents the prediction of the t^{th} base classifier for input x, T is the total number of classifiers in the ensemble, *mode*(·) returns the class label that occurs most frequently among the predictions.

This majority voting strategy helps improve classification accuracy and robustness by aggregating decisions from multiple models, thereby reducing variance and mitigating overfitting.

3.2.3. (c) Logistic Regression (LR)

LR uses the sigmoid function to estimate the probability of a binary outcome [27]. The sigmoid (logistic) function is used in LR to model the likelihood that a given input x belongs to the positive class y = 1, and it is defined as follows: In logistic regression, the probability that the binary target $y \in \{0, 1\}$ equals 1 given an input vector $x \in R^n$ is modelled as:

$$\hat{y} = P(y = 1 \mid \mathbf{x}) = \frac{1}{1 + \exp(-(\mathbf{w}^\top \mathbf{x} + b))} \tag{4}$$

where, $w \in R^n$ is the weight vector, $b \in R$ is the bias term, $y\hat{} \in (0, 1)$ is the predicted probability after applying the sigmoid activation function. To train the model, the binary cross-entropy (log loss) is commonly employed as the objective function, which penalizes the deviation between the predicted probability $y\hat{}$ and the true label y:

$$\mathcal{L} = -\left[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\right] \tag{5}$$

10.48047/jocaaa.2024.33.08.207

This loss function encourages the predicted probability to approach the true binary class label, with greater penalties for confident but incorrect predictions.

3.2.4. (d) Support Vector Machine (SVM)

The goal of the supervised learning algorithm SVM is to determine the best hyperplane for maximally separating data points from various classes. In its basic form for linearly separable data, SVM seeks to solve the following optimization problem [28].

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to} \quad y_i(w^T x_i + b) \geq 1 \tag{6}$$

In this formulation, x_i is the input feature vector of the i-th data sample, w and b are the weight vector and bias term, respectively. The term $\frac{1}{2}\|w\|^2$ is minimized to maximize the margin, which is the distance between the hyperplane and the nearest data points (support vectors). The constraint $y_i(w^T x_i+b) \geq 1$ ensures that each point lies outside or on the margin boundary. Once the model is trained, prediction for a new input x is performed using the decision function and is stated as:

$$f(x) = \text{sign}(w^T x + b) \tag{7}$$

Here, the sign of the decision function determines the class label, if the result is positive, the input is classified as +1 and if negative, it is classified as −1.

3.2.5. (e) Gradient Boosting (GB)

GB is an ensemble learning technique that builds predictive models in a sequential manner by combining the outputs of weak learners typically decision trees to minimize the overall prediction error. At each iteration m, the model updates its prediction function by adding a new base learner h_m(x), scaled by a learning rate γ_m. This process is mathematically represented as Friedman (2001) [29].

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{8}$$

10.48047/jocaaa.2024.33.08.207

In the above equation, $F_{m-1}(x)$ is the ensemble model from the previous iteration. The key idea behind gradient boosting is to fit $h_m(x)$ to the residuals or negative gradients of the loss function with respect to the prediction of the current model. By doing so, the algorithm gradually corrects the errors made by previous models, improving performance step-by-step.

3.3. Deep Learning (DL) Algorithms

In automatic feature extraction and classification tasks, DL models have demonstrated impressive performance, especially for biomedical signals such as ECG. DL architectures do not require handcrafted features, in contrast to traditional machine learning techniques, as they learn hierarchical representations straight from raw input data. Because of this, they are especially well-suited for ECG classification, where intricate temporal patterns and minute variations in waveform shape are crucial distinguishing characteristics. Below is a description of the DL models used in this framework.

3.3.1. (a) 1D Convolutional Neural Network (1D-CNN)

Sequential data, like time-series signals, audio waveforms, or text, is the main application for 1D CNNs [30]. Convolutional filters are used in this architecture to extract features from local sequences by sliding across one dimension, usually time or position.. Activation functions and pooling layers come after layers like Conv1D in the architecture, which ultimately leads to dense layers for regression or classification.

A 1D convolutional layer computes the output at position i in layer l using the following equation:

$$z_i^{[l]} = \phi\left(\sum_{j=1}^{K} \omega_j^{[l]} \cdot a_{i+j-1}^{[l-1]} + \beta^{[l]}\right) \tag{9}$$

where $_{i+j-1}$ is the input activation from the previous layer at index $i + j - 1$, $\omega^{[l]}{}_j$ is the j^{th} weight in the convolutional filter at layer l, $\beta^{[l]}$is the bias term associated with the filter, $\phi(\cdot)$ is the activation function (e.g., ReLU, sigmoid), K denotes the kernel size (number of weights).

The expression computes a weighted sum over a local segment of the input sequence (a sliding window), adds a bias, and applies a nonlinear activation to produce the output at each position.

10.48047/jocaaa.2024.33.08.207

3.3.2. (b) 2D Convolutional Neural Network (2D-CNN)

The most popular and extensively used kind of CNNs is 2D ones. They use 2D convolutional kernels to process data with two spatial dimensions: height and width. Every filter gains the ability to recognize distinct spatial patterns, including corners, edges, and textures. Conv2D layers, nonlinear activations (like ReLU), pooling layers to lower dimensionality, and dense output layers are all common components of a 2D CNN architecture. The data is first transformed into 2D matrices (such as spectrograms) and convolved over height and width while dealing with the ECG signals [31]. The output of a 2D convolutional layer at position (p, q) in the l^{th} layer can be defined as:

$$y_{p,q}^{[l]} = \phi\left(\sum_{u=0}^{K-1}\sum_{v=0}^{K-1}\theta_{u,v}^{[l]} \cdot a_{p+u,\,q+v}^{[l-1]} + \beta^{[l]}\right) \tag{10}$$

$_{p+u,\ q+v}$is the input activation from the previous layer $[l - 1]$ at spatial position $(p+u,\ q+v)$, $\theta^{[l]}{}_{u,v}$ is the convolution kernel weight at position (u, v) in layer l, $\beta^{[l]}$is the bias term associated with the filter, $\phi(\cdot)$ is the non linear activation function, $K \times K$ is the filter size (assuming a square kernel for simplicity), $y^{[l]}{}_{p,q}$ is the output at position (p, q) after applying convolution and activation.

This formulation describes the convolution operation as a weighted summation over a local region of the input (also called the receptive field), followed by the addition of a bias and the application of an activation function $\phi(\cdot)$. The sliding window moves across the input feature map to produce the full output feature map of the convolutional layer.

3.3.3. (c) 3D Convolutional Neural Network (3D-CNN)

By convolving over three dimensions, height, width, and depth, 3D CNNs expand on the idea of 2D CNNs. These dimensions are frequently used to represent time or multiple channels, as in MRI or CT scans. For tasks like 3D medical imaging, where the volumetric or temporal continuity between frames or slices is essential, they are particularly effective.

3D CNNs use cube-shaped kernels that move across the input volume to simultaneously capture temporal and spatial features [32]. 3D CNNs are applied to spatial-temporal ECG data.

10.48047/jocaaa.2024.33.08.207

$$z_{i,j,k}^{(l)} = f\left(\sum w_{m,n,o}^{(l)} x_{i+m,j+n,k+o}^{(l-1)} + b^{(l)}\right) \tag{11}$$

where the output at position (*i, j, k*) in layer *l* is computed by applying a 3D filter across the depth, height, and width of the input from the previous layer. Each filter weight $w^{(l)}$ value $x^{(l-1)}{}_{m,n,o}$ is multiplied by the corresponding input $_{i+m,j+n,k+o}$, and the results are summed. A bias term b^l is added, and the result is passed through an activation function *f.*

3.3.4. (d)Densely Connected Convolutional Networks (DenseNet)

DenseNet is a convolutional neural network architecture that introduces direct connections between any two layers with the same feature-map size. In a traditional CNN, each layer feeds into the next one sequentially. In DenseNet, however, each layer receives inputs from all previous layers, and its output is passed on to all subsequent layers [33]. DenseNet connects each layer to all previous ones to enhance feature propagation:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{12}$$

where H_l is a composition of batch normalization, ReLU, and convolution operations. Each block in DenseNet contains several layers, and these blocks are separated by transition layers that down sample the feature maps. DenseNet has been widely used in tasks ranging from image recognition to medical imaging because of its strong performance and parameter efficiency.

3.3.5. (e) EfficientNet

EfficientNet is a family of CNN architectures that aim to optimize both accuracy and efficiency. EfficientNet uses a compound scaling method to scale all three dimensions uniformly and systematically. Starting from a small baseline model (EfficientNet-B0), the architecture scales up to larger versions (EfficientNet-B1 to B7) by using a compound coefficient that balances the scaling of depth, width, and resolution. EfficientNet is built on a backbone called MBConv (Mobile Inverted Bottleneck Convolution), which is highly optimized for both speed and performance [34]. EfficientNet scales network dimensions using compound scaling:

$$\text{depth} \propto \alpha^{\phi}, \quad \text{width} \propto \beta^{\phi}, \quad \text{resolution} \propto \gamma^{\phi} \tag{13}$$

10.48047/jocaaa.2024.33.08.207

with the constraint: $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$

3.3.6. (f) Long Short-Term Memory (LSTM)

LSTM networks, a variant of recurrent neural networks (RNNs), are specifically designed to learn long-term temporal dependencies in sequential data such as ECG signals. Unlike standard RNNs, which often suffer from vanishing or exploding gradients during training, LSTMs employ memory cells and gating mechanisms—input, forget, and output gates—to retain relevant information across long sequences [35].

In ECG classification, LSTM networks are capable of capturing temporal patterns that span multiple cardiac cycles, which is critical for identifying arrhythmias or irregularities such as atrial fibrillation and ventricular tachycardia [36]. The model inputs are typically sequences of ECG signal windows or heartbeat segments, which are processed through one or more LSTM layers to produce class probabilities.

Several studies have demonstrated the superior performance of LSTM-based architectures in classifying ECG signals compared to traditional ML techniques, especially in noisy and multi-class environments [37]. Moreover, hybrid models combining LSTM with CNNs have been used to extract both spatial and temporal features from ECG signals, significantly improving classification accuracy in multi-lead datasets [38].

Despite their effectiveness, LSTMs are computationally more demanding than feedforward networks and may require careful hyperparameter tuning and regularization techniques to avoid overfitting, especially in smaller datasets [39].

An efficient formulation of the Long Short-Term Memory (LSTM) cell uses vectorized operations, where all gates are computed simultaneously through a single matrix multiplication:

$$\begin{bmatrix} f_t \\ i_t \\ o_t \\ \tilde{C}_t \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} \left(W \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b \right) \tag{14}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{15}$$

$$h_t = o_t \odot \tanh(C_t) \tag{16}$$

10.48047/jocaaa.2024.33.08.207

Table 1: Performance Comparison of ML and DL Models on ECG Signal Classification

Model	Accuracy (%)	Precision	Recall	F1-Score
Logistic Regression	98.200	0.980	0.985	0.980
SVM	98.900	0.985	0.990	0.990
Random Forests	98.900	0.990	0.990	0.990
KNN	98.500	0.985	0.980	0.985
1D-CNN	99.000	0.990	0.990	0.990
2D-CNN	98.300	0.985	0.980	0.985
3D-CNN	98.900	0.990	0.990	0.990
DenseNet	96.000	0.955	0.965	0.965
EfficientNet	99.000	0.985	0.985	0.985
LSTM	97.800	0.980	0.975	0.975

Here, $x_t \in \mathrm{R}^m$ is the input vector, $h_{t-1} \in \mathrm{R}^n$ is the previous hidden state, and $W \in \mathrm{R}^{4n\times(n+m)}$ is the concatenated weight matrix for all gates. $b \in \mathrm{R}^{4n}$ is the corresponding bias vector. The functions σ and tanh represent the element-wise sigmoid and hyperbolic tangent functions, respectively. The operator $\odot$ denotes the Hadamard (element-wise) product.

This vectorized formulation improves computational efficiency and aligns with how LSTM cells are typically implemented in deep learning libraries.

3.4. Evaluation Metrics

All models are evaluated using classification reports and confusion matrices. The metrics include Accuracy, Precision, Recall, and F1-score. This help assess the performance across all ECG classes and ensure both ML and DL models are compared fairly under the same experimental conditions.

10.48047/jocaaa.2024.33.08.207

4. Results and Discussion

The classification performance of various ML and DL models for ECG signal analysis is summarized in Table 1, and their respective confusion matrix plots are shown in Figure 5. Among traditional ML models, SVM and RF demonstrated superior accuracy (98.9%), precision (0.990), and recall (0.990), reflecting their robustness in handling complex and high-dimensional data. L also performed well, achieving 98.2% accuracy with a precision of 0.980.

For distance-based classification, KNN yielded 98.5% accuracy, slightly trailing SVM, but still maintaining high precision (0.985). These results confirm the effectiveness of classical classifiers in capturing discriminative signal patterns.

In the deep learning category, 1D-CNN, 3D-CNN, and EfficientNet achieved the highest accuracy (99.0%), demonstrating the potential of automatic feature extraction. Both 1D-CNN and 3D-CNN obtained the best F1-score (0.990), indicating excellent balance between precision and recall. The success of these models highlights the importance of temporal and spatial feature learning in biomedical signal classification.

2D-CNN achieved slightly lower accuracy (98.3%) compared to 1D and 13 3D variants, suggesting that 2D convolution may not fully capture the temporal dynamics inherent in ECG signals. DenseNet reported relatively lower accuracy (96.0%), possibly due to overfitting or sub-optimal representation of time-series features. LSTM, although

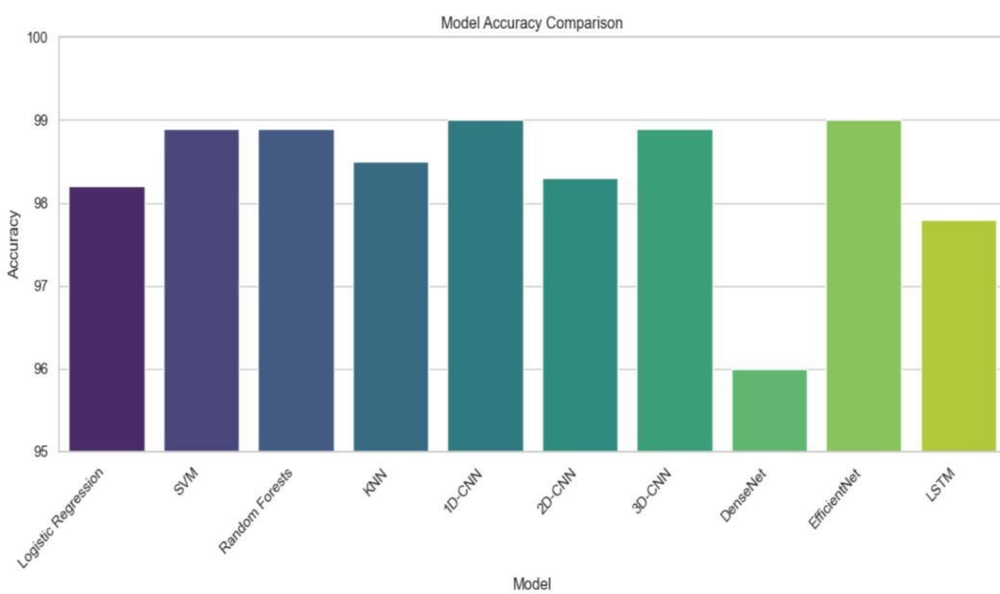

Figure 4: Model accuracy comparison.

10.48047/jocaaa.2024.33.08.207

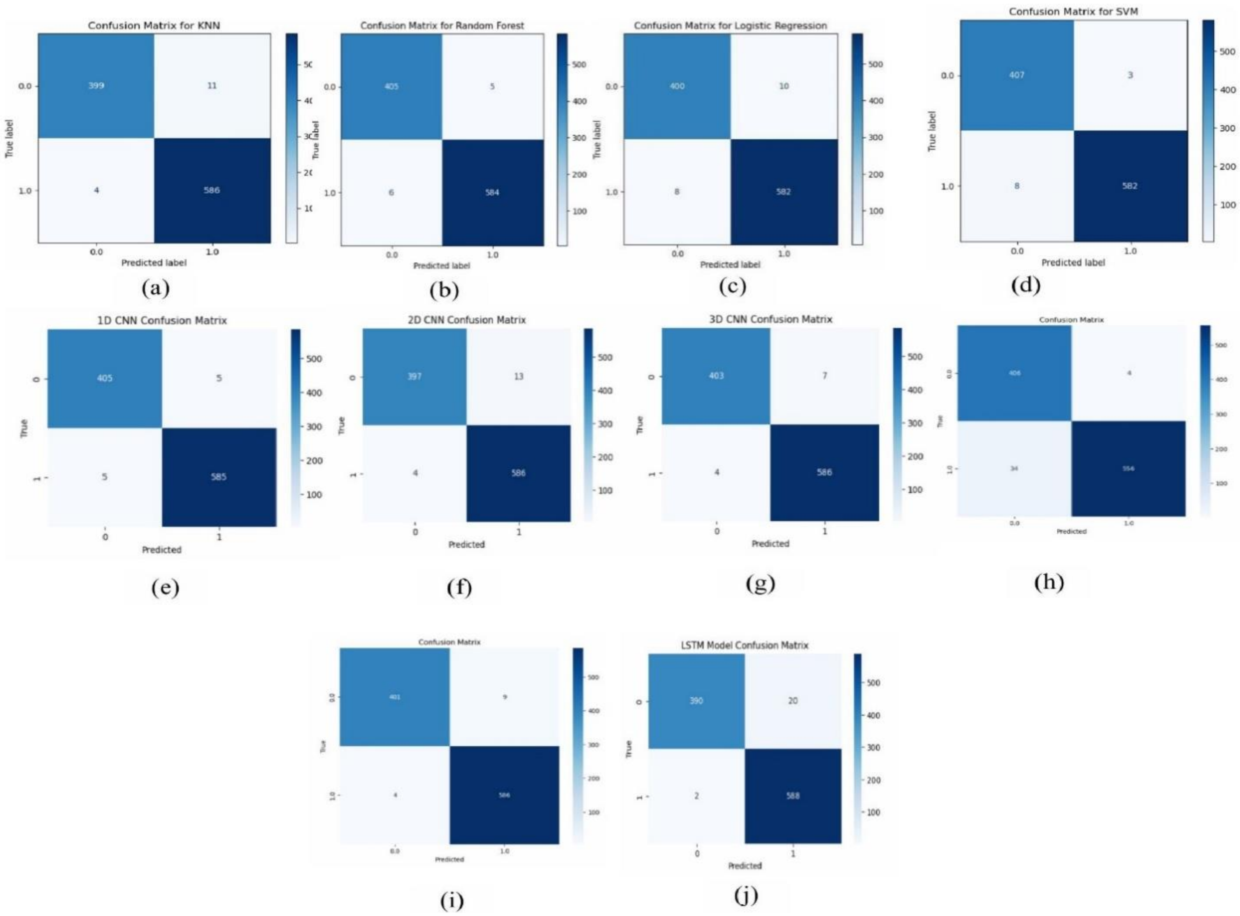

Figure 5: Confusion Matrix for (a)KNN, (b)Random Forest, (c)Logistic Regression, (d)SVM, (e)1d-CNN, (f)2d-CNN, (g)3D-CNN, (h)DenseNet, (i)EfficientNet, (j)LSTM.

designed for sequential data, achieved 97.8% accuracy, indicating that its temporal modeling capability is beneficial, though not outperforming CNN-based models in this setting.

Overall, deep learning architectures, particularly 1D-CNN, 3D-CNN, and EfficientNet, consistently outperformed traditional models, affirming their suitability for ECG signal classification tasks.

5. Conclusion

This research presented a detailed comparative analysis of multiple ML and DL models for ECG signal classification. Based on experimental results, DL models such as 1D-CNN, 3D-CNN, and EfficientNet outperformed traditional ML algorithms in terms of accuracy, precision, recall, and F1-score, with 1D-CNN and EfficientNet achieving the highest

10.48047/jocaaa.2024.33.08.207

accuracy of 99.0%. Random Forests and SVM also showed competitive performance among ML models, highlighting their continued relevance for ECG classification tasks. However, DL models demonstrated a stronger ability to learn from raw signal data without extensive manual feature engineering. The superior performance of convolutional models emphasizes their potential for real-time, scalable, and automated cardiac diagnostics. Future work will focus on optimizing model architectures, expanding to multi-class classification tasks, and deploying these models in edge devices for real-time monitoring applications.

References

[1] World Health Organization. (2023). *Cardiovascular diseases (CVDs)*. https://www.who.int/news-room/factsheets/detail/cardiovascular-diseases-(cvds)

[2] Singh, A. K., & Krishnan, S. (2023). ECG signal feature extraction trends in methods and applications. *Biomedical Engineering Online, 22*(1), 22. https://doi.org/10.1186/s12938-023-01064-y

[3] Strodthoff, N., Wagner, P., Schaeffter, T., & Samek, W. (2020). Deep learning for ECG analysis: Benchmarks and insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics, 25*(5), 1519–1528. https://doi.org/10.1109/JBHI.2020.3022989

[4] Xiao, Q., et al. (2023). Deep learning-based ECG arrhythmia classification: A systematic review. *Applied Sciences, 13*(9), 4964. https://doi.org/10.3390/app13094964

[5] Xiong, Z., et al. (2018). ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiological Measurement, 39*(9), 094006. https://doi.org/10.1088/1361-6579/aad852

[6] Ansari, Y., et al. (2023). Deep learning for ECG arrhythmia detection and classification: An overview of progress for period 2017–2023. *Frontiers in Physiology, 14*, 1246746. https://doi.org/10.3389/fphys.2023.1246746

[7] Ameen, M. S., et al. (2024). Lightweight hybrid CNN-RNN model for ECG classification using raw signals. *Journal of Biomedical Informatics, 147*, 104506. https://doi.org/10.1016/j.jbi.2024.104506

[8] Hammad, M., et al. (2021). Detection of abnormal heart conditions based on ECG signals using multiple features fusion and one-dimensional convolutional neural network. *Measurement, 178*, 109290. https://doi.org/10.1016/j.measurement.2021.109290

[9] Xiong, Z., et al. (2018). ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiological Measurement, 39*(9), 094006. https://doi.org/10.1088/1361-6579/aad852

10.48047/jocaaa.2024.33.08.207

[10] Jameel, M., et al. (2024). Multiclass ECG arrhythmia classification using an improved deep residual neural network. *Biomedical Signal Processing and Control, 87*, 105250. https://doi.org/10.1016/j.bspc.2023.105250

[11] Xiao, Q., et al. (2023). Deep learning-based ECG arrhythmia classification: A systematic review. *Applied Sciences, 13*(9), 4964. https://doi.org/10.3390/app13094964

[12] Bhatia, A., et al. (2022). A hybrid CNN-LSTM model for arrhythmia classification using ECG signals. *Computers in Biology and Medicine, 144*, 105403. https://doi.org/10.1016/j.compbiomed.2022.105403

[13] Zabihi, M., et al. (2024). ECG signal classification using deep residual networks with attention mechanisms. *Computer Methods and Programs in Biomedicine, 239*, 107581. https://doi.org/10.1016/j.cmpb.2024.107581

[14] Paul, J., et al. (2023). ECG classification with denoising autoencoder-based ensemble learning. *Journal of Medical Systems, 47*(2), 15. https://doi.org/10.1007/s10916-023-01963-w

[15] Cheng, X., et al. (2021). Transfer learning for ECG classification using CNN. *IEEE Access, 9*, 146192–146202. https://doi.org/10.1109/ACCESS.2021.3123032

[16] Ansari, Y., et al. (2023). Deep learning for ECG arrhythmia detection and classification: An overview of progress for period 2017–2023. *Frontiers in Physiology, 14*, 1246746. https://doi.org/10.3389/fphys.2023.1246746

[17] Zhao, H. (2023). Explainable AI techniques in ECG diagnosis. *Journal of Biomedical Research, 37*(1), 35–44. https://doi.org/10.7555/JBR.37.2023005

[18] Strodthoff, N., et al. (2020). Deep learning for ECG analysis: Benchmarks and insights from PTB-XL. *IEEE Journal of Biomedical and Health Informatics, 25*(5), 1519–1528. https://doi.org/10.1109/JBHI.2020.3022989

[19] Islam, M. T., et al. (2022). Class imbalance handling in deep learning for ECG classification. *Biomedical Signal Processing and Control, 72*, 103336. https://doi.org/10.1016/j.bspc.2021.103336

[20] Singh, A. K., & Krishnan, S. (2023). ECG signal feature extraction trends in methods and applications. *Biomedical Engineering Online, 22*(1), 22. https://doi.org/10.1186/s12938-023-01064-y

[21] Zhang, X., et al. (2021). Comparative study of ECG classification methods using MIT-BIH and PTB-XL. *Computer Methods and Programs in Biomedicine, 208*, 106255. https://doi.org/10.1016/j.cmpb.2021.106255

10.48047/jocaaa.2024.33.08.207

[22] Lu, Y., et al. (2024). Efficient 1D-CNN model for ECG classification in wearable devices. *Sensors, 24*(1), 168. https://doi.org/10.3390/s24010168

[23] Kim, H., et al. (2020). Ultra-low-power DL model for real-time ECG monitoring. *IEEE Transactions on Biomedical Circuits and Systems, 14*(3), 504–514. https://doi.org/10.1109/TBCAS.2020.2972293

[24] Admass, M., & Bogale, M. (2024). Challenges in DL-based ECG classification: A systematic review. *Health Information Science and Systems, 12*(1), 9. https://doi.org/10.1007/s13755-024-00237-6

[25] Plawiak, P. (2018). Novel ECG recognition system based on extended modular neural networks with DWT for RR interval extraction. *Biomedical Signal Processing and Control, 45*, 68–83. https://doi.org/10.1016/j.bspc.2018.05.012

[26] Smíšek, R., Mautner, P., & Smital, L. (2018). ECG arrhythmia classification using a combination of convolutional and recurrent neural networks. *Measurement Science Review, 18*(6), 241–248. https://doi.org/10.1515/msr-2018 0031 [27] Yang, C., Zhang, H., & Zhu, Z. (2019). ECG heartbeat classification based on multiscale convolutional neural network. In *Proceedings of the IEEE International Conference on Artificial Intelligence Circuits and Systems(AICAS)* (pp. 287–290). https://doi.org/10.1109/AICAS.2019.8771490

[28] Yildirim, O. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Computers in Biology and Medicine, 96*, 189–202. https://doi.org/10.1016/j.compbiomed.2018.03.016

[29] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

[30] Tan, H. P., Zhang, H., & Tan, M. L. (2018). Arrhythmia detection using CNN-LSTM architecture with attention mechanism. In *International Conference on Artificial Intelligence and Big Data (ICAIBD)* (pp. 62–68). https://doi.org/10.1109/ICAIBD.2018.8396207

[31] Xiang, Y., Zhang, L., & Wu, H. (2018). ECG classification using 1D convolutional neural network. *Journal of Physics: Conference Series, 1087*(6), 062020. https://doi.org/10.1088/1742-6596/1087/6/062020

10.48047/jocaaa.2024.33.08.207

[32] Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3D convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 4489–4497). https://doi.org/10.1109/ICCV.2015.510

[33] Peimankar, A., & Puthusserypady, S. (2021). A novel deep learning model for automatic diagnosis of cardiac arrhythmia using convolutional and recurrent neural networks. *Computers in Biology and Medicine, 136*, 104645. https://doi.org/10.1016/j.compbiomed.2021.104645

[34] Midani, M., Faris, H., Al-Zoubi, A., Aljarah, I., & Alqudah, A. M. (2023). CNN BiLSTM model for ECG classification using heartbeat segmentation. *Biomedical Signal Processing and Control, 81*, 104440. https://doi.org/10.1016/j.bspc.2023.104440

[35] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[36] Yildirim, O. (2018). A novel wavelet sequence-based deep bidirectional LSTM network model for ECG signal classification. *Computers in Biology and Medicine, 96*, 189–202. https://doi.org/10.1016/j.compbiomed.2018.03.016 [37] Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine, 161*, 1–13. https://doi.org/10.1016/j.cmpb.2018.04.005

[38] Oh, S. L., Ng, E. Y. K., Tan, R. S., & Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heartbeats. *Computers in Biology and Medicine, 102*, 278–287. https://doi.org/10.1016/j.compbiomed.2018.06.002

[39] Ismail, M. F., & Sadiq, M. T. (2021). Comparative analysis of deep learning models for ECG classification: Challenges and solutions. *IEEE Access, 9*, 131060–131072. https://doi.org/10.1109/ACCESS.2021.3114133