ECG Arrhythmia Classification Using Machine Learning

Lab Report

Abstract—This study compares four machine learning approaches for automated ECG arrhythmia classification using the MIT-BIH dataset. We evaluated Random Forest, XGBoost, 1D CNN, and LSTM models on 109,446 training and 21,892 test samples across five arrhythmia classes. The 1D CNN achieved the best performance with 98.26% accuracy, demonstrating deep learning's effectiveness for temporal pattern recognition in ECG signals.

I. INTRODUCTION

Electrocardiogram (ECG) analysis is crucial for detecting cardiac arrhythmias. Manual interpretation is time-consuming and prone to human error. This work explores automated classification using both traditional machine learning and deep learning approaches on the MIT-BIH Arrhythmia Database containing five classes: Normal (N), Supraventricular ectopic beats (S), Ventricular ectopic beats (V), Fusion beats (F), and Unknown beats (Q).

II. EXPLORATORY DATA ANALYSIS

The dataset exhibits significant class imbalance (Fig. 1). Normal beats dominate with 72,471 samples (66.2%), while minority classes like Fusion (641 samples, 0.6%) pose classification challenges. Each ECG signal contains 187 time-step features sampled at 125Hz.
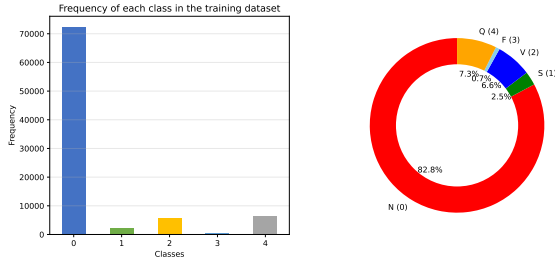

Fig. 1. Class distribution showing severe imbalance.

Sample ECG waveforms (Fig. 2) reveal distinct morphological patterns across classes, with ventricular beats showing wider QRS complexes compared to normal beats.

III. MODELS

A. Random Forest

Ensemble method with 100 decision trees (max depth=42). Handles class imbalance well and provides feature importance analysis showing critical time steps around the QRS complex.
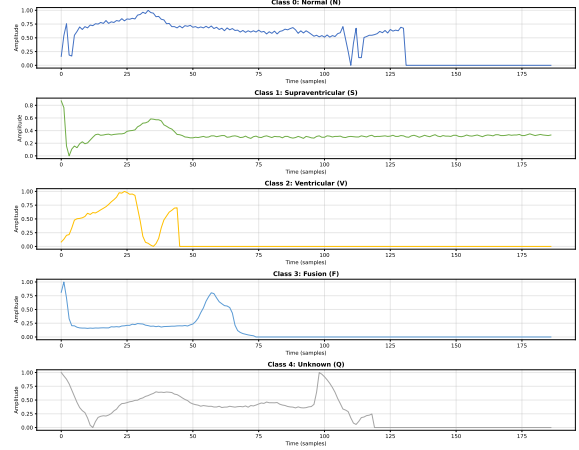

Fig. 2. Representative ECG signals for each class.

B. XGBoost

Gradient boosting with 100 estimators trained with class weights to address imbalance. Uses early stopping (10 rounds) to prevent overfitting.

C. 1D Convolutional Neural Network

Three-layer CNN architecture with 64 filters per layer, batch normalization, and max pooling. Automatically learns temporal features from raw ECG signals. Trained for 50 epochs with early stopping and learning rate scheduling.

D. LSTM

Bidirectional LSTM with 2 layers (128 hidden units) captures long-term temporal dependencies. Includes dropout (0.3) for regularization.

IV. EVALUATION AND RESULTS

All models were evaluated on the same test set using accuracy, weighted F1-score, and macro F1-score metrics. Table I summarizes the performance.

TABLE I
MODEL PERFORMANCE COMPARISON

Model	Accuracy (%)	F1 (Weighted)	F1 (Macro)
Random Forest	97.45	0.9728	0.8716
XGBoost	94.78	0.9528	0.8010
1D CNN	**98.26**	**0.9825**	**0.9048**
LSTM	96.62	0.9638	0.8144

The 1D CNN outperformed all models across all metrics, achieving 98.26% accuracy and 0.9048 macro F1-score. The confusion matrix (Fig. 3) shows excellent per-class performance with minimal misclassification.
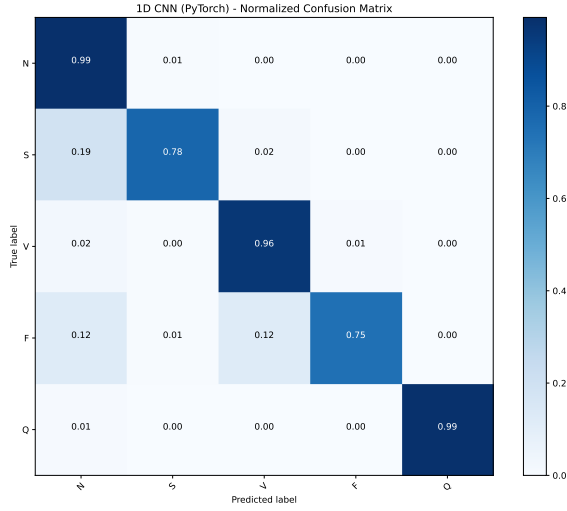

Fig. 3. 1D CNN confusion matrix (normalized).

Random Forest achieved competitive results (97.45%) with much faster training. XGBoost struggled with minority classes despite class weighting. LSTM performance (96.62%) was limited by sequential processing constraints.

The 1D CNN's training curves (Fig. 4) show stable convergence without overfitting, validating the model architecture and regularization strategy.
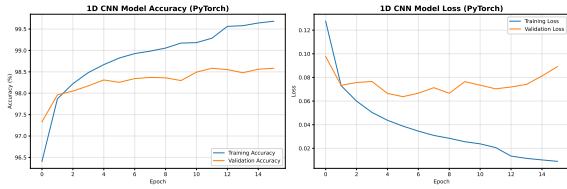

Fig. 4. 1D CNN training history showing convergence.

V. CONCLUSION

This study demonstrates that 1D CNN architectures are highly effective for ECG arrhythmia classification, achieving 98.26% accuracy by automatically learning discriminative temporal features. While Random Forest offers a good accuracy-speed trade-off, deep learning approaches excel at capturing complex patterns in time-series medical data. Future work should explore data augmentation and ensemble methods to further improve minority class detection.