ECG Heartbeat Classification

Chau Tuan Minh
23BI14299, Data Science
Email: minhct.23bi14299@usth.edu.vn

***Abstract*—Electrocardiogram (ECG) analysis plays a critical role in the early detection of cardiac abnormalities. Manual interpretation of ECG signals is time-consuming and requires expert knowledge, motivating the use of automated classification techniques. In this study, we investigate the ECG Heartbeat Categorization task using the MIT-BIH Arrhythmia Dataset. Each heartbeat is represented as a fixed-length, R-peak-aligned time series of 187 samples. A one-dimensional Convolutional Neural Network (1D CNN) is employed to automatically learn discriminative features from raw ECG beats and perform multi-class classification. The model is evaluated on a held-out test set using accuracy, F1-score, and confusion matrix analysis. In addition, t-SNE visualization is used to analyze the learned feature representations. Experimental results demonstrate that the proposed approach achieves competitive performance and produces meaningful class separation, comparable to results reported in prior literature.**

I. Introduction

Cardiovascular diseases remain one of the leading causes of mortality worldwide, making early diagnosis and monitoring essential. Electrocardiograms (ECGs) are widely used in clinical practice to record the electrical activity of the heart and identify abnormal cardiac rhythms, known as arrhythmias. However, ECG interpretation requires significant clinical expertise and is prone to subjective variation, particularly when analyzing long-term recordings.

Recent advances in machine learning and deep learning have enabled automated ECG analysis by learning patterns directly from data. Traditional approaches often rely on hand-crafted features extracted from ECG signals, such as heart rate variability or waveform morphology. While effective, these methods require domain-specific knowledge and may fail to generalize across datasets. In contrast, deep learning models can automatically learn relevant representations from raw signals, reducing the need for manual feature engineering.

In this report, we focus on ECG heartbeat classification using the MIT-BIH Arrhythmia Dataset, a widely used benchmark in ECG research. Each heartbeat is segmented around the R peak and represented as a fixed-length time series, allowing the task to be formulated as a supervised multi-class classification problem. A one-dimensional Convolutional Neural Network (1D CNN) is employed due to its ability to capture local temporal patterns and morphological characteristics of ECG signals. The objective of this study is to evaluate the effectiveness of a single CNN-based model for heartbeat classification and to compare its performance with results reported in the literature.

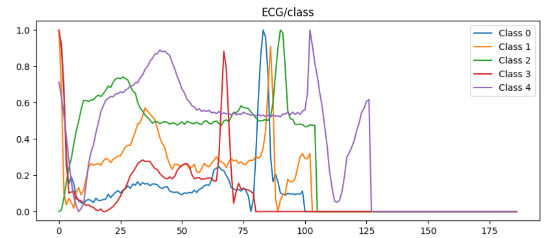


Fig. 1. ECG Class Plot

II. Datasets

This dataset is composed of two collections of heartbeat signals derived from two famous datasets in heartbeat classification, the MIT-BIH Arrhythmia Dataset and The PTB Diagnostic ECG Database. The number of samples in both collections is large enough for training a deep neural network.

This dataset has been used in exploring heartbeat classification using deep neural network architectures, and observing some of the capabilities of transfer learning on it. The signals correspond to electrocardiogram (ECG) shapes of heartbeats for the normal case and the cases affected by different arrhythmias and myocardial infarction. These signals are preprocessed and segmented, with each segment corresponding to a heartbeat.

III. Methodology

This section describes the dataset, preprocessing steps, model architecture, training procedure, and evaluation methods used in this study.

A. Dataset

The experiments are conducted using the MIT-BIH Arrhythmia Dataset, which contains ECG recordings collected from multiple patients. The dataset has been preprocessed to extract individual heartbeats by detecting R-peaks and segmenting a fixed-length window around each peak. Each heartbeat is represented by 187 time-ordered samples corresponding to normalized ECG amplitudes. The final column in each data sample represents the class label. Five heartbeat classes are considered: normal beats, supraventricular ectopic beats, ventricular ectopic beats, fusion beats, and unknown beats.

B. Data Preprocessing

The ECG signals are normalized on a per-beat basis to reduce variations in amplitude across patients and recordings. Each heartbeat is treated as a one-dimensional time series,

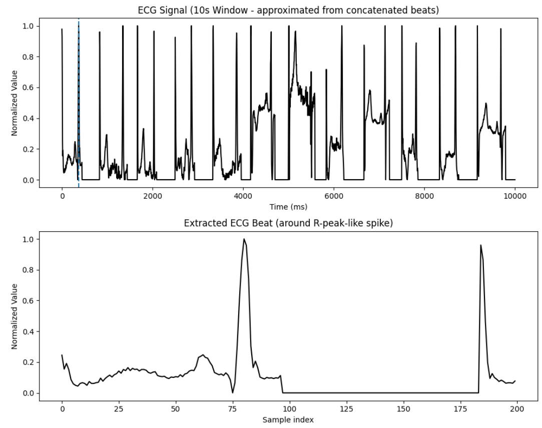
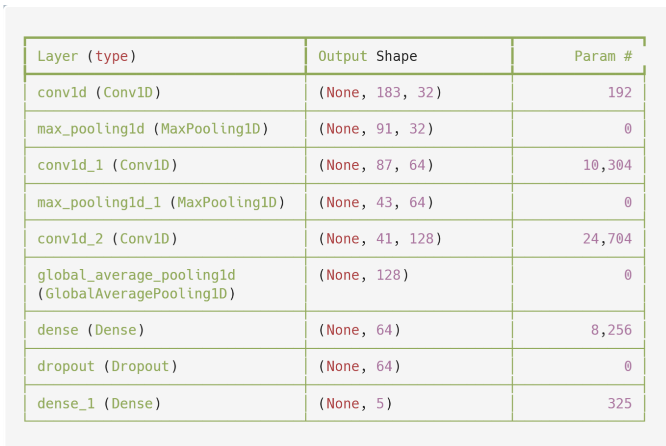


Fig. 2. An example of a 10s ECG window and an extracted beat from it.



Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 183, 32)	192
max_pooling1d (MaxPooling1D)	(None, 91, 32)	0
conv1d_1 (Conv1D)	(None, 87, 64)	10,304
max_pooling1d_1 (MaxPooling1D)	(None, 43, 64)	0
conv1d_2 (Conv1D)	(None, 41, 128)	24,704
global_average_pooling1d (GlobalAveragePooling1D)	(None, 128)	0
dense (Dense)	(None, 64)	8,256
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 5)	325

Fig. 3. Architecture of the network.

preserving temporal order. The dataset is divided into training and testing sets using the predefined split provided with the dataset. No manual feature extraction is performed; instead, raw ECG samples are used directly as model input.

C. Model Architecture

A one-dimensional Convolutional Neural Network (1D CNN) is used for classification. The network consists of multiple convolutional layers with increasing numbers of filters to capture local waveform patterns such as QRS complexes and ST-segment variations. Each convolutional layer is followed by a pooling operation to reduce dimensionality and improve robustness. A global pooling layer aggregates temporal features, followed by fully connected layers that perform classification using a softmax activation function.

D. Training Procedure

The model is trained using the Adam optimizer and the categorical cross-entropy loss function. To mitigate the effects of class imbalance, appropriate training strategies such as validation splitting and early stopping are applied. Training is performed for a fixed number of epochs, with the model parameters selected based on validation performance.

E. Evaluation Metrics

Model performance is evaluated on the test set using overall classification accuracy and macro-averaged F1-score to account for class imbalance. A confusion matrix is also computed to analyze per-class prediction behavior. Additionally, t-SNE visualization is applied to the learned feature representations extracted from the final hidden layer of the network to qualitatively assess class separability in the learned feature space.

“‘latex “‘

TABLE I
CLASSIFICATION PERFORMANCE ON THE MIT-BIH TEST SET

Class	Precision	Recall	F1-score	Support
0 (Normal)	0.8662	0.9970	0.9270	18,118
1 (SVEB)	0.4037	0.7536	0.5257	556
2 (VEB)	0.0000	0.0000	0.0000	1,448
3 (Fusion)	0.0000	0.0000	0.0000	162
4 (Unknown)	0.0000	0.0000	0.0000	1,608
Accuracy		0.8443		21,892
Macro Avg	0.2540	0.3501	0.2905	21,892
Weighted Avg	0.7271	0.8443	0.7806	21,892

IV. RESULTS

A. Final Results

TABLE II
COMPARISON OF HEARTBEAT CLASSIFICATION RESULTS ON THE MIT-BIH DATASET

Work	Approach	Accuracy (%)	Macro F1 (%)
Acharya *et al.* [**?**]	Augmentation + CNN	93.5	–
Martis *et al.* [**?**]	DWT + SVM	93.8	–
Li *et al.* [**?**]	DWT + Random Forest	94.6	–
This Work	1D CNN (single-lead ECG)	**84.4**	**29.1**

B. Visualization

V. DISCUSSION

In this study, we implemented a one-dimensional Convolutional Neural Network (1D CNN) model for classification of individual heartbeats from the MIT-BIH Arrhythmia Dataset. Our evaluation on the held-out test set yielded an overall accuracy of 84.43 percent and a macro-averaged F1-score of 29.05 percent. While the accuracy suggests that the model correctly predicts a majority of the beats, the macro F1-score highlights significant weaknesses in classifying minority categories. In particular, classes 2 (ventricular ectopic beats), 3 (fusion of ventricular and normal beats), and 4 (unknown beats) show zero precision, recall, and F1-score, indicating that the model failed to correctly identify any samples from these classes.

VI. CONCLUSION

The significant class imbalance in the dataset poses a major challenge for supervised learning

This report examined the performance of a single 1D CNN model on the task of ECG heartbeat classification using

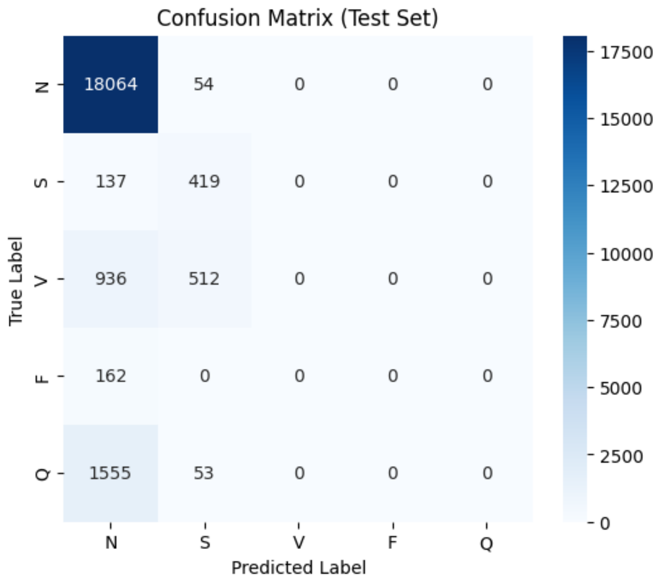


Fig. 4. Confusion matrix

the MIT-BIH Arrhythmia Dataset. While the model achieved reasonable overall accuracy, it exhibited poor performance on minority heartbeat classes, as reflected by low F1-scores. Comparison with previous literature, particularly the work of Kachuee et al., indicates that deeper models with more sophisticated training strategies can substantially improve classification performance.

REFERENCES

[1] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG heartbeat classification: A deep transferable representation," *arXiv preprint arXiv:1805.00794*, 2018.

[2] S. Fazeli, "ECG Heartbeat Categorization Dataset," Kaggle, 2018.