

ECG Heartbeat Categorization Report

Trinh Van Quyet

Data Science

quyetTV.22BI13387@usth.edu.vn

Abstract—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.

I. INTRODUCTION

In this study, we focus on ECG heartbeat classification using the MIT-BIH Arrhythmia Database, a well-known benchmark dataset in the field. We develop a deep learning-based model that integrates Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) to effectively capture both spatial and temporal patterns in ECG signals.

My approach is evaluated by comparing its performance with the results presented in the existed study [1]. Through rigorous experimentation and performance assessment, we aim to demonstrate the effectiveness of my deep learning model in improving ECG heartbeat classification accuracy.

II. DATASETS

A. Overview

The dataset contains preprocessed ECG heartbeat signals extracted from the MIT-BIH Arrhythmia Database. There are 109446 samples and each heartbeat instance is labeled into one of five categories. Each record contains 187 measurements of the heartbeat signal.

Label	Category	Description
0	N (Normal)	Normal heartbeat
1	S (Supraventricular)	Abnormal supraventricular beat
2	V (Ventricular)	Ventricular ectopic beat
3	F (Fusion)	Fusion of normal and abnormal beat
4	Q (Unknown)	Unclassified heartbeat

TABLE I: Five Distinct Type of Heartbeat.

B. Exploratory Data Analysis

In this section, we explore the MIT-BIH dataset to understand its structure and distribution.

1) Class Distribution

To gain insights into the dataset's number of samples of each class, we visualized the class distribution for both the training and test sets. Figure 1 & 2 present the class distribution in the MIT-BIH training and test sets.

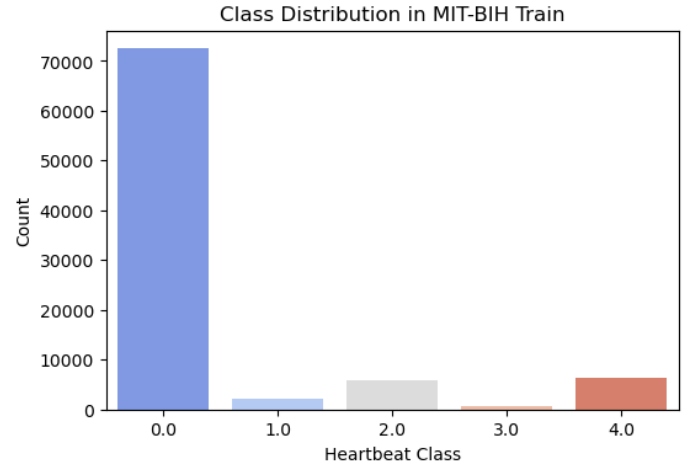


Fig. 1: Class Distribution in MIT-BIH Train

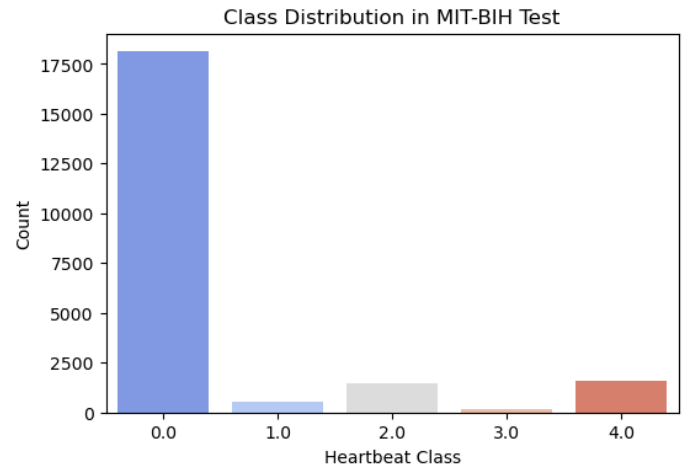


Fig. 2: Class Distribution in MIT-BIH Test

The training set exhibits a significant class imbalance, with one class dominating the dataset. This imbalance may impact model performance, requiring techniques such as resampling to address it.

2) ECG Signal Samples

The figure 3 highlights the consistency between the training and test ECG signals, ensuring that the model is trained on representative data.

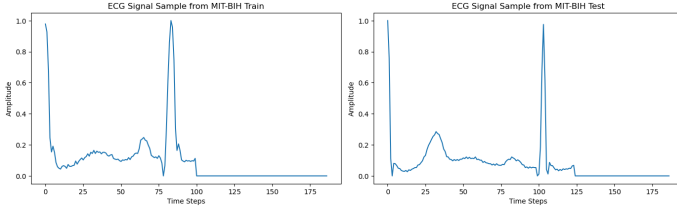


Fig. 3: ECG Signal Sample from MIT-BIH Train & Test

Both signals exhibit similar patterns, characterized by an initial high amplitude followed by a decline and subsequent fluctuations.

Despite their overall similarity, slight variations can be noted in the amplitude and waveform morphology between the two plots. These differences could be attributed to variations in individual heartbeats or noise present in the dataset. The gradual decrease in amplitude towards the end of both signals indicates the termination of a heartbeat cycle.

3) Peak Detection

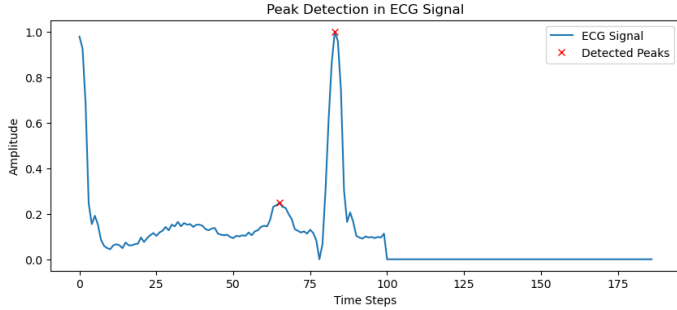


Fig. 4: Peak Detection in ECG Signal

The figure 4 illustrates peak detection in an ECG signal, highlighting key points of interest. The blue line represents the raw ECG signal, while the red markers indicate the detected peaks. These peaks correspond to significant events in the heartbeat cycle, with the highest peak likely representing the R-wave in the QRS complex.

4) Frequency of ECG Signal

The figure 5 presents a frequency analysis of ECG signals for both the MIT-BIH training and test datasets. The x-axis represents frequency in Hertz (Hz), while the y-axis shows the power spectral density. The training dataset exhibits higher power at lower frequencies, peaking between 0 and 10 Hz before gradually decreasing.

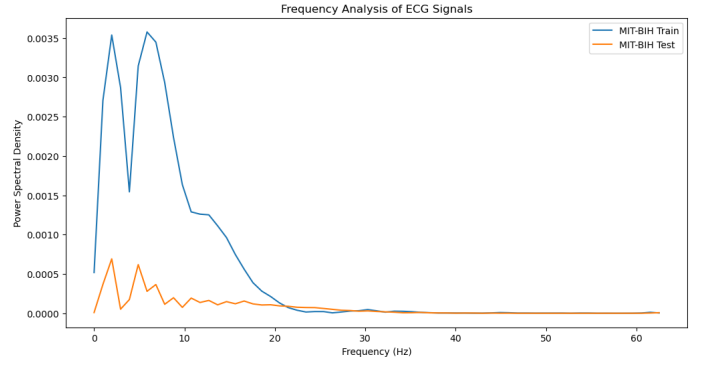


Fig. 5: Frequency of ECG Signal

The test dataset follows a similar trend but with lower power spectral density values, suggesting possible variations in signal amplitude or noise levels between the two datasets. The presence of frequency components beyond 20 Hz is minimal, indicating that most ECG signal information is contained within the lower frequency range.

III. MODEL ARCHITECTURE

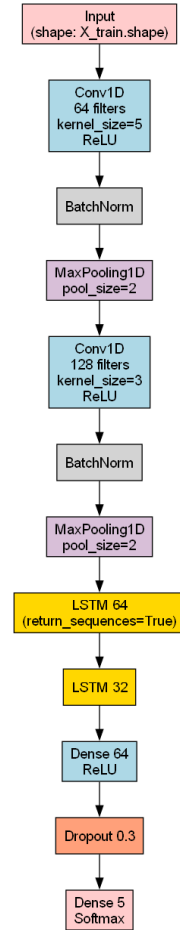


Fig. 6: Model Architecture

I used a simple one-dimensional convolutional neural network combining with LSTM layers 6 is designed classifying sequential data into five categories.

The first layer is a 1D convolutional layer with 64 filters and a kernel size of 5, using the ReLU activation function. This layer extracts local features from the input ECG signal. A batch normalization layer follows, which helps stabilize training by normalizing activations. Next, a max pooling layer with a pool size of 2 reduces the spatial dimensions, lowering computational complexity while retaining important features.

A second 1D convolutional layer with 128 filters and a kernel size of 3 is applied, further refining feature extraction. Another batch normalization layer follows for stability, and a second max pooling layer with a pool size of 2 is used to downsample the feature maps.

Following the convolutional layers, two LSTM layers are introduced to capture the sequential dependencies in the ECG signals. The first LSTM layer has 64 units and returns sequences, allowing the second LSTM layer with 32 units to process the extracted temporal features.

The model then includes a fully connected dense layer with 64 neurons using the ReLU activation function to learn high-level representations. A dropout layer with a 30% dropout rate is added to prevent overfitting by randomly deactivating neurons during training.

Finally, a dense output layer with 5 neurons and a softmax activation function is used for classification, corresponding to the five heartbeat classes.

The model is compiled using the Adam optimizer, categorical cross-entropy loss, and accuracy as the evaluation metric. It is trained for 20 epochs with a batch size of 64, using validation data to monitor performance.

IV. RESULT

A. Performance Evaluation

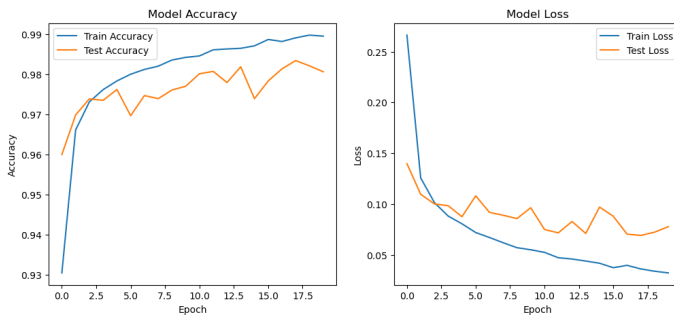


Fig. 7: Performance Metrics of Model

The left plot 7 represents the accuracy of both the training and test datasets over multiple epochs. The training accuracy follows a smooth upward trend, reaching above 99%, indicating that the model effectively learns from the training data. The test accuracy also remains relatively high, mostly above 97%, demonstrating good generalization. However, the fluctuations in test accuracy suggest that the model might

not be completely stable when making predictions. These fluctuations could be caused by potential overfitting, where the model performs exceptionally well on the training data but struggles slightly with variations in the test set.

The right plot 7 illustrates the loss curves for both training and testing datasets. The training loss steadily decreases, which is an expected behavior for a well-trained model. However, the test loss exhibits significant fluctuations, with sharp spikes at certain points. These variations suggest that the model struggles with some batches of the test set, leading to occasional high losses. This behavior may indicate high variance, meaning the model might be overfitting to the training data and not generalizing consistently.

From the figure 8, we observe that Class 0 has the highest number of correctly classified samples, with 18,012 instances predicted correctly, and very few misclassified samples. This suggests that the model performs exceptionally well for this majority class. Class 1 shows a noticeable number of misclassifications, with 119 samples being incorrectly predicted as Class 0. This indicates that the model struggles to distinguish Class 1 from Class 0, potentially due to similarities in their feature distributions.

Similarly, Class 2 achieves strong performance, with 1,340 correct predictions. However, there are some misclassified cases, particularly 68 samples being predicted as Class 0 and 29 as Class 3. This suggests some level of overlap between these classes, which could be mitigated through improved feature extraction or additional training data.

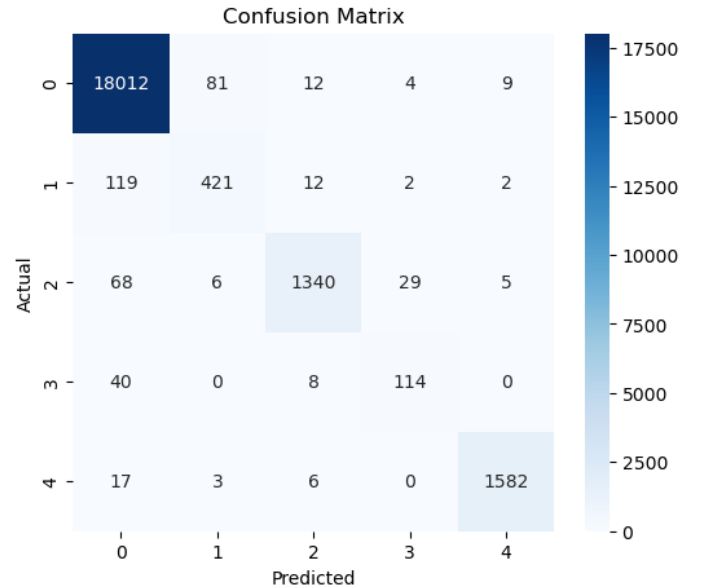


Fig. 8: Confusion Matrix

For Class 3, the model correctly classifies 114 samples but shows minor confusion with Classes 2 and 0. Lastly, Class 4 achieves 1,582 correct classifications, with only a few misclassified instances, demonstrating relatively high precision for this category.

B. Compare With Existed Paper

Work	Approach	Average Accuracy (%)
This report	CNN + LSTM	98.0
Acharya <i>et al.</i> [2]	Augmentation + CNN	93.5
Martis <i>et al.</i> [3]	DWT + SVM	93.8
Li <i>et al.</i> [4]	DWT + random forest	94.6

TABLE II: Comparison of heartbeat classification results.

In comparison to previous studies II, my CNN + LSTM model achieves a significantly higher accuracy of 98.0%, outperforming all referenced methods. Acharya *et al.* [2] implemented an augmentation-based CNN approach, achieving an accuracy of 93.5%, while Martis *et al.* [3] and Li *et al.* [4] employed Discrete Wavelet Transform (DWT) with Support Vector Machine (SVM) and Random Forest, obtaining 93.8% and 94.6%, respectively. The superior performance of my model can be attributed to the integration of CNN for feature extraction and LSTM for temporal sequence learning, allowing for a more comprehensive understanding of ECG signals.

V. CONCLUSION

In this report, I presented a CNN + LSTM-based approach for ECG heartbeat classification, achieving an impressive accuracy of 98.0%, surpassing previous methods. By leveraging CNN for feature extraction and LSTM for temporal sequence learning, my model effectively captures both spatial and temporal dependencies in ECG signals.

My results demonstrate that deep learning models, particularly those integrating CNN and LSTM, provide a more robust and accurate solution for heartbeat classification.

References

- [1] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," *arXiv preprint arXiv:1804.06812*, 2018.
- [2] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.
- [3] R. J. Martis, U. R. Acharya, C. M. Lim, K. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ECG signals," *International Journal of Neural Systems*, vol. 23, no. 4, p. 1350014, 2013.
- [4] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.