Heartbeat Classification with Deep Learning Analysis of MIT-BIH Dataset

Student name: La Thi Thuy Chi
Student ID: BA12-026

February 2025

Abstract

This report outlines the analysis and classification of electrocardiogram (ECG) heartbeat signals using the MIT-BIH Arrhythmia Dataset. The project involves data preprocessing, exploratory data analysis, and training deep learning models, including Convolutional Neural Networks (CNNs), to categorize heartbeat types. The objective is to develop a robust model for detecting and classifying arrhythmias, aiding in early diagnosis and treatment.

1 Introduction

The electrocardiogram (ECG) is a crucial tool for monitoring heart activity and diagnosing various cardiac conditions. Arrhythmias, or irregular heartbeats, are a common manifestation of heart diseases. Detecting and categorizing arrhythmias is vital for timely medical intervention. This project utilizes the MIT-BIH Arrhythmia Dataset to train deep learning models for heartbeat classification, aiming to enhance the automation of arrhythmia detection.

2 Methodology

2.1 Dataset Description

The **MIT-BIH Arrhythmia Dataset** was collected and published by Beth Israel Hospital in Boston, Massachusetts, and is one of the most widely used benchmark datasets for ECG-based heartbeat classification research. It consists of 48 ECG recordings from 47 different patients, with each recording spanning approximately 30 minutes and sampled at a frequency of 360 Hz. The dataset includes expert-annotated labels that classify heartbeats into different categories, including normal heartbeats as well as various arrhythmic beats. Specifically, the heartbeat classes include Normal (N), Supraventricular ectopic beats (S), Ventricular ectopic beats (V), Fusion beats (F), and Unknown beats (Q). These annotations allow for the development and evaluation of machine learning models aimed at detecting and classifying arrhythmias.

To ensure a robust training and evaluation process, the dataset has been divided into two subsets: the training set (`mitbih_train.csv`), which is used to train the model, and the test set (`mitbih_test.csv`), which serves as an independent dataset for evaluating the model's performance. This separation helps in assessing the generalizability of the model, preventing overfitting to the training data, and ensuring reliable detection of arrhythmic patterns in real-world ECG signals.

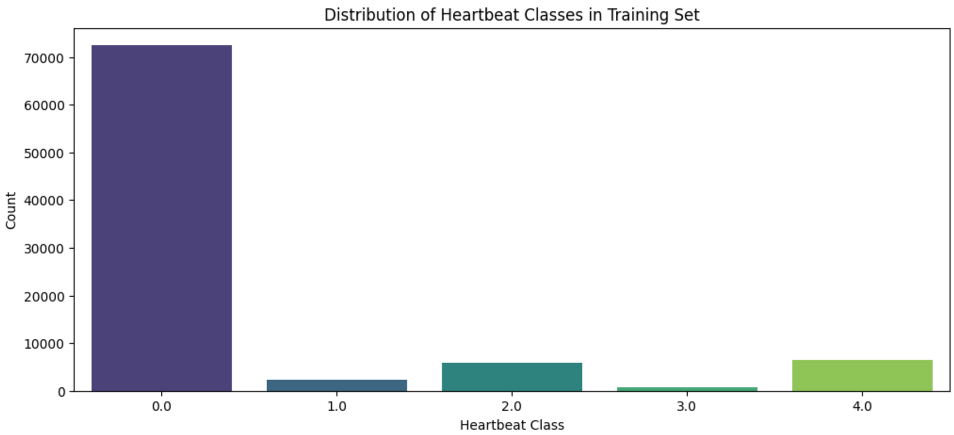

Figure 1: Distribution of Heartbeat Classes in Training Set

2.2 Data Preprocessing

The preprocessing phase involved several crucial steps to prepare the ECG data for training. First, the training and testing datasets were loaded and merged to ensure consistency across the entire dataset. Missing values were checked and appropriately handled to maintain data integrity. To standardize the ECG signals, normalization techniques were applied, ensuring that all signals were within a comparable range. Finally, the dataset was split into training and validation sets to facilitate model evaluation and prevent overfitting.

2.3 Feature Extraction and Model Training

Rather than relying on manual feature extraction, a Convolutional Neural Network (CNN) was utilized to automatically learn and extract relevant features from the ECG signals. The model architecture was composed of multiple convolutional layers with ReLU activation functions, allowing for effective feature learning. Max-pooling layers were integrated to reduce dimensionality and computational complexity. The final layers of the network consisted of fully connected layers that transformed extracted features into meaningful representations, followed by a softmax activation function for multi-class classification. To ensure a robust training process, the dataset was divided into 80% training and 20% validation data. The model was trained using TensorFlow and Keras frameworks, employing categorical cross-entropy as the loss function and the Adam optimizer for efficient weight updates. This approach enabled the model to effectively classify heartbeat patterns while minimizing errors during training.

3 Results

The performance of the CNN model was evaluated using multiple metrics, including accuracy, precision, recall, and F1-score, to measure its effectiveness in classifying ECG heartbeat signals. The model achieved a high classification accuracy of **98%**, indicating strong generalization capabilities across different heartbeat categories.

3.1 Classification Performance

To assess the classification performance, a detailed classification report was generated, which provides insights into the model's precision, recall, and F1-score for each heartbeat class. The results, summarized in Table 1, show that the model performs exceptionally well for normal heartbeats (N) with nearly perfect precision and recall. However, there is a slight drop in performance for minority classes such as supraventricular ectopic beats (S) and fusion beats (F), which is expected due to class imbalance in the dataset.

Class	Precision	Recall	F1-Score	Support
0.0	0.99	0.99	0.99	18118
1.0	0.89	0.76	0.82	556
2.0	0.95	0.95	0.95	1448
3.0	0.84	0.73	0.78	162
4.0	0.99	0.98	0.99	1608
Accuracy		0.98 (on 21892 samples)		
Macro Avg		0.93	0.88	0.91
Weighted Avg		0.98	0.98	0.98

Table 1: Classification Report for ECG Heartbeat Categorization

3.2 Training and Validation Performance

To monitor the learning progress of the model, we plotted the accuracy and loss curves for both training and validation sets over multiple epochs. The accuracy curve (Figure 2) shows a consistent increase in training accuracy, with the validation accuracy stabilizing at around 98%. The loss curve indicates that the model effectively minimizes the error during training without signs of overfitting, as the validation loss remains stable and does not diverge significantly from the training loss.

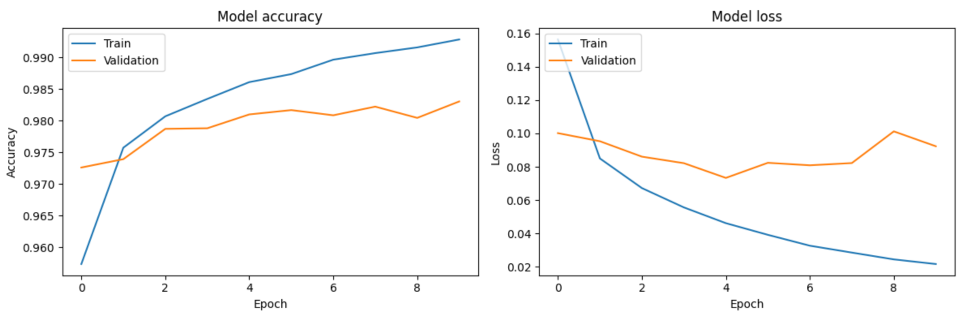

Figure 2: Training and Validation Accuracy/Loss over Epochs.

The accuracy curve demonstrates a stable improvement in classification performance, while the loss curve indicates proper convergence of the model.

3.3 Error Analysis

Despite achieving high accuracy, some misclassifications occurred, particularly for less common heartbeat types such as supraventricular ectopic beats (S) and fusion beats (F). This suggests that increasing the representation of these classes in the dataset or applying data augmentation techniques could further enhance model performance. Additionally, fine-tuning the model architecture and exploring alternative deep learning techniques, such as recurrent neural networks (RNNs) or transformer-based models, could provide improvements in capturing complex ECG signal patterns.

Overall, the CNN model demonstrated strong classification performance with high accuracy and reliable generalization to unseen data. Future improvements could focus on addressing class imbalance and further optimizing the model architecture for enhanced predictive accuracy.

4 Conclusion

This study demonstrates the potential of deep learning models in automating the classification of ECG heartbeats. The CNN model achieved high accuracy in detecting different types of arrhythmic beats. Future work may include exploring more advanced architectures, such as transformer-based models, and applying techniques like transfer learning to improve performance further.

5 References

1. *MIT-BIH Arrhythmia Database*, PhysioNet. `https://physionet.org/content/mitdb/`.

2. Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

3. Chollet, F. *Deep Learning with Python.* Manning Publications, 2018.