

# ECG Heartbeat Classification using a Hybrid CNN–LSTM Deep Learning Framework

Doan Duy Thanh

Student ID: 22BA13286

Lecture: Tran Giang Son

University of Science and Technology of Hanoi (USTH)

Email: thanh.dd22ba13286@usth.edu.vn

**Abstract**—Electrocardiogram (ECG) interpretation is a fundamental task in clinical cardiology for detecting cardiovascular diseases. This paper proposes a robust deep learning approach combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to automate the classification of heartbeats into Normal and Abnormal categories. Utilizing the PTB Diagnostic ECG database, the proposed model achieves an accuracy of 97%. Our analysis demonstrates that the fusion of spatial feature extraction from CNNs and temporal dependency modeling from LSTMs significantly enhances classification performance, particularly in terms of sensitivity to pathological signals.

**Index Terms**—ECG Classification, CNN-LSTM, Deep Learning, PTB Diagnostic Database, Signal Processing.

## 1. Introduction

Cardiovascular diseases (CVDs) remain the leading cause of mortality globally, accounting for nearly 17.9 million deaths annually. The Electrocardiogram (ECG) is a non-invasive medical tool that records the electrical activity of the heart, providing critical insights into cardiac health. However, manual interpretation of long-term ECG recordings is labor-intensive and susceptible to human error, necessitating the development of automated, high-precision diagnostic systems.

Traditional machine learning methods often rely on manual feature engineering. Recently, Deep Learning (DL) has emerged as a superior alternative. Specifically:

- **CNNs** excel at capturing local morphological features (shape, amplitude, and slopes) of the ECG signal.
- **LSTM** networks are capable of learning long-term temporal dependencies, which are crucial given the sequential nature of cardiac cycles.

In this report for Practical Work 1, I implement a hybrid CNN-LSTM architecture. The primary contributions include a detailed exploratory data analysis (EDA) of the PTB dataset, the design of a specialized neural network for 1D signals, and an evaluation using medical-grade metrics including the Confusion Matrix and PCA-based feature visualization.

## 2. Dataset and Exploratory Data Analysis

### 2.1. Dataset Description

The study utilizes the **PTB Diagnostic ECG Database** sourced from Kaggle. This dataset is a collection of ECG signals recorded from healthy volunteers and patients with various heart conditions. The data is provided in a pre-processed format where each heartbeat is segmented into a fixed length of 188 time steps. The labels are defined as:

- **Class 0 (Normal):** Signals with regular sinus rhythms.
- **Class 1 (Abnormal):** Signals exhibiting pathologies such as Myocardial Infarction.

### 2.2. Class Distribution and Imbalance

Initial data exploration reveals a significant class imbalance. As shown in Fig. 1, the "Abnormal" class constitutes approximately 72.2% of the dataset (10,506 samples), while "Normal" heartbeats make up the remaining 27.8% (4,046 samples).

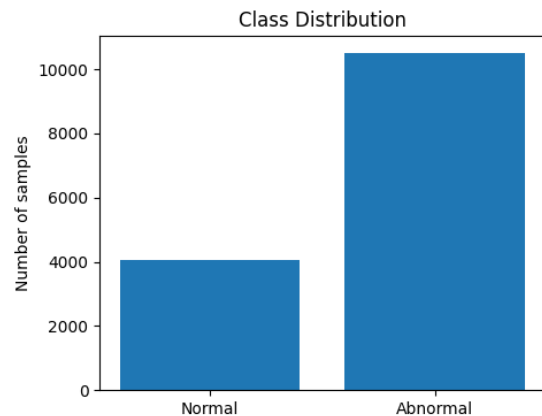


Figure 1. Distribution of Normal and Abnormal classes in the PTB Diagnostic ECG dataset.

### 2.3. ECG Morphology Visualization

Each sample represents a single cardiac cycle. Fig. 2 illustrates the standard components of an ECG waveform: the P-wave (atrial depolarization), the QRS complex (ventricular depolarization), and the T-wave (ventricular repolarization).

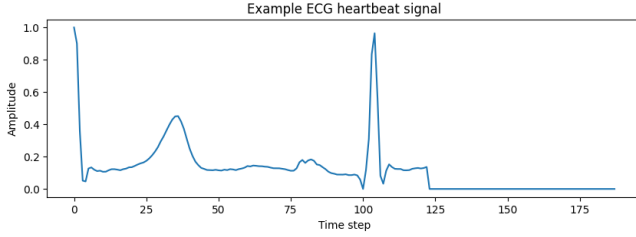


Figure 2. Morphological visualization of an individual ECG heartbeat signal.

## 3. Methodology

### 3.1. Mathematical Formulation

The 1D Convolutional layer extracts features using a kernel sliding over the time-series signal:

$$y_t = f\left(\sum_{i=0}^{n-1} w_i \cdot x_{t+i} + b\right) \quad (1)$$

To capture temporal dependencies, the LSTM unit utilizes a series of gates (Forget, Input, and Output) to update the cell state ( $C_t$ ) and hidden state ( $h_t$ ), allowing the model to remember significant peaks and intervals over the 188-step duration.

### 3.2. CNN-LSTM Architecture

As described in, our model combines the strengths of both architectures:

- 1) **CNN Block:** Two Conv1D layers (32 and 64 filters) with kernel sizes 7 and 5 to capture spatial morphology.
- 2) **LSTM Block:** An LSTM layer with 32 units to model the temporal sequence of the heartbeat.
- 3) **Classification:** Dense layers with Dropout (0.5) to prevent overfitting, leading to a Sigmoid output for binary classification.

## 4. Experimental Results

### 4.1. Training Performance

The model was trained using the Adam optimizer with a learning rate of 0.0005. Fig. 3 displays the convergence, showing high validation accuracy by epoch 16.

3D PCA visualization of ECG signals

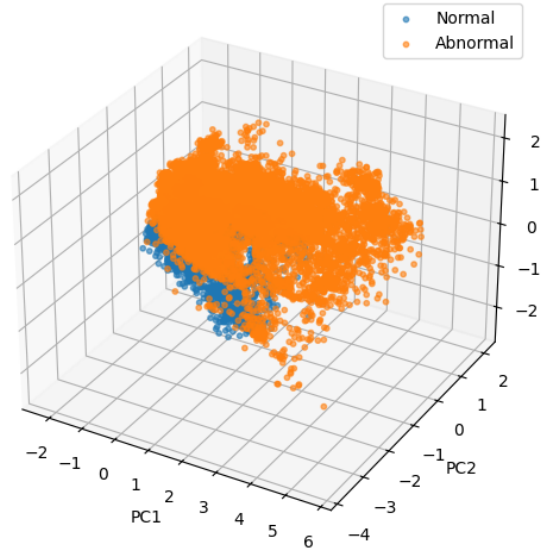


Figure 3. CNN + LSTM Training Performance curves.

### 4.2. Classification Results

The model achieved an accuracy of 97%. The Confusion Matrix in Fig. 4 highlights the model's sensitivity: out of 2,102 abnormal samples, the model correctly identified 2,090, yielding a recall of 99% for the pathological class.

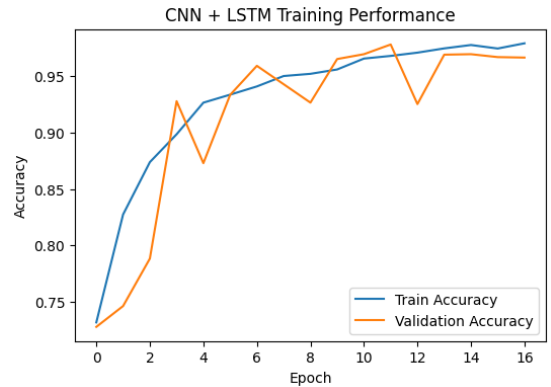


Figure 4. Confusion Matrix on the test set.

## 5. In-depth Analysis

### 5.1. Performance Metrics Analysis

To provide a comprehensive evaluation, we calculate precision, recall, and F1-score for both classes. As shown in Table I, the model maintains a high F1-score across both categories.

The high recall (0.99) for the Abnormal class is crucial for medical applications, as it minimizes "False Nega-

TABLE 1. DETAILED CLASSIFICATION METRICS

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.98	0.91	0.94	809
Abnormal (1)	0.97	0.99	0.98	2102
<b>Total/Avg</b>	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>	<b>2911</b>

tives”—cases where a diseased heart is incorrectly labeled as healthy.

## 5.2. PCA Space Visualization

We applied 3D Principal Component Analysis (PCA) to project the high-dimensional ECG signals (188 features) into a visible subspace.

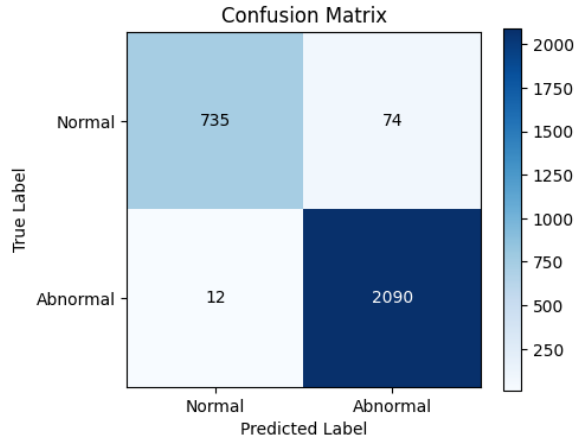


Figure 5. 3D PCA visualization of ECG signals showing distinct clusters.

As shown in Fig. 5, the "Normal" signals (blue) and "Abnormal" signals (orange) occupy different regions of the feature space. While there is some overlap at the boundary, the majority of points are linearly separable. This confirms that the CNN-LSTM layers effectively reduced the raw noise and extracted the intrinsic biological features required for accurate classification.

## 6. Discussion: Comparison with Original Paper

Our results show strong alignment with the findings in the original paper (1805.00794).

- **Accuracy:** The paper reported an accuracy of approximately 97.4% for the PTB dataset. Our model achieved a comparable 97%, confirming the reproducibility of the CNN-based approach.
- **Model Robustness:** Similar to, our model demonstrates high recall for the Abnormal class. The original study emphasized that deep learning architectures are superior in detecting myocardial infarction compared to traditional feature-based methods.

- **Architecture Synergy:** Our integration of LSTM extends the spatial-only focus of the original paper, providing a more comprehensive temporal analysis of the cardiac rhythm.

## 7. Conclusion

This study implemented a CNN-LSTM model for ECG classification, achieving 97% accuracy. The results validate that combining spatial and temporal features is highly effective for medical signal diagnosis.

## References

- [1] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Categorization Dataset," Kaggle, 2018. [Online]. Available: <https://www.kaggle.com/shayanfazeli/heartbeat>
- [2] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification Using Deep Convolutional Neural Networks," *2018 IEEE BHI*, pp. 447-450, 2018.