

Early Detection of Cardiac Arrhythmia Using Bi-LSTM Networks on ECG Signals

^{1*}Dikshya Aryal, ²Hemant Joshi

^{1,2}Department of Computer Engineering, Universal Engineering and Science College

Email: ²hemantajoshi@uesc.edu.np

*Corresponding Author: ¹dikshyauesc@gmail.com

Abstract

Cardiovascular diseases (CVDs) are a top global cause of death. Early detection of cardiac abnormalities using non-invasive techniques, such as electrocardiogram (ECG) signals, plays a significant role in improving patient outcomes and reducing mortality. Traditional methods rely on handcrafted feature extraction, which limits scalability and robustness. Therefore, this paper presents the implementation of a Bidirectional Long Short-Term Memory (Bi-LSTM) neural network for binary arrhythmia classification using the MIT-BIH Arrhythmia Database. A structured preprocessing pipeline, which consists of bandpass filtering, Z-score normalization, and R-peak-centered segmentation, is applied to extract 112,572 labeled heartbeat segments. The implemented Bi-LSTM model is systematically evaluated against three classical machine learning baselines: Support Vector Machine (SVM), Logistic Regression, and Random Forest. Accuracy, precision, recall, F1-score, specificity, False Positive Rate (FPR), and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) are used to evaluate performance. The Bi-LSTM gets 96.41% overall accuracy, while under the same conditions, conventional models like Random Forest obtain 98.76% accuracy, showing that the benefits of deep learning depend on the quantity and balance of the data. The effect of class imbalance (90,603 normal vs. 21,969 abnormal beats) on model sensitivity is also shown in the paper. The results highlight the significance of domain-specific evaluation and balanced training procedures for ECG classification systems.

Keywords: *Arrhythmia Detection, Deep Learning, ECG, Arrhythmia Classification, and Bi-LSTM*

I. INTRODUCTION

CVDs are currently the leading cause of mortality globally, causing an estimated 17.9 million deaths each year [1]. Early diagnosis plays an important part in the timely detection of heart-related abnormalities, which can significantly improve patient outcomes. Electrocardiogram is one of the most effective and widely used diagnostic tools, which is a non-invasive, low-cost technique that records the electrical activity of the heart [2]. Although it has diagnostic values, manual interpretation of ECG signals is time-consuming, it requires specialized expertise, and it is prone to human error while dealing with long-term continuous monitoring [3]. In order to overcome these limitations, machine learning (ML) based automated methods have been developed to analyze ECG data quickly and accurately. Among these ML models, deep learning

models like LSTM networks have shown good results due to their ability to capture temporal dependencies in sequential data like ECG signals [4].

In this research, we investigate the effectiveness of a Bi-LSTM network by implementing existing architectural principles for the early detection of heart disease using ECG signals from the MIT-BIH Arrhythmia dataset [5]. Under identical experiment conditions, the Bi-LSTM model is benchmarked against three classical ML classifiers: SVM, Logistic Regression, and Random Forest.

One way to help with early detection and ease some workload is automated ECG analysis. This can be applied to areas that may have limited clinical resources. The goal of this project is to create a responsible pipeline and analyze deep learning versus traditional algorithms with realistic constraints from real-world data.

Specifically, this paper will:

1. Use a systematic ECG preprocessing pipeline that includes bandpass filtering, normalization, and heartbeat segmentation.
2. Design and implement a custom Bi-LSTM architecture for binary arrhythmia classification.
3. Use evaluation measures to compare the Bi-LSTM to SVM, Logistic Regression, and Random Forest under identical conditions.
4. Calculate how class imbalance affects model sensitivity.

Background: ECG fundamentals

An ECG is a diagnostic test that records the electrical signals generated by the heart during its cycle of contraction and relaxation. These signals are produced by depolarization and repolarization of the heart muscle cells and travel through the body to the skin, where they can be detected using surface electrodes. A standard 12-lead ECG uses electrodes placed on the chest and limbs to provide a comprehensive view of the heart's electrical activity [6].

The ECG waveform consists of several components corresponding to phases of cardiac electrical activity. The P wave represents atrial depolarization, the QRS complex corresponds to rapid ventricular depolarization, and the T wave reflects ventricular repolarization. By analyzing the shape, duration, and intervals of these waves, clinicians diagnose various cardiac conditions, such as arrhythmias, myocardial infarction, and conduction abnormalities [7]. Modern ECG recordings can range from short 10-second readings, which are often sufficient for routine diagnosis, to extended modern techniques such as continuous 24-hour Holter monitoring, which can capture rare and intermittent arrhythmias. These long recordings produce vast amounts of data that are impractical to analyze manually. Therefore, there is a significant need for automated methods that can process, interpret, and classify ECG signals in real time, which addresses this challenge by applying computational algorithms to classify signals as normal or abnormal. Preprocessing steps such as noise filtering, normalization, and segmentation are employed to enhance signal quality and improve classification accuracy [8]. The use of traditional models and

their dependence on handcrafted features limit their ability to generalize across diverse patient populations.

To address these limitations, deep learning models, particularly Bi-LSTM networks, have been widely adopted due to their ability to model temporal dependencies in sequential biomedical signals. Their application, ECG analysis, supports automated and scalable cardiac abnormality detection.

II. LITERATURE REVIEW

Automated detection of heart disease from ECG signals has been a widely explored research area in recent years, primarily driven by the growing availability of large-scale annotated datasets and the advancement of machine learning algorithms. This section reviews the progression of ECG classification methods from traditional machine learning to deep learning models, with a focus on LSTM networks. Traditional machine learning techniques have historically relied on handcrafted features extracted from ECG waveforms, such as heart rate variability, QRS complex duration, and frequency-domain parameters. For instance, [9] compared several classifiers and highlighted the effectiveness of SVM and k-Nearest Neighbors (KNN) in arrhythmia detection. While these models provided satisfactory results on small-scale datasets, they often lacked robustness and generalizability across diverse patient populations.

The introduction of deep learning significantly shifted the landscape of ECG analysis by eliminating the dependence on manual feature engineering. Convolutional Neural Networks (CNNs) have been particularly effective in automatically learning spatial features from ECG signals. [10] Proposed a real-time patient-specific ECG classification framework using 1D CNNs, demonstrating high accuracy and scalability for clinical deployment. Similarly, [8] presented a comprehensive review of deep learning applications in healthcare, emphasizing the ability of CNNs to outperform traditional methods in ECG-based diagnosis.

Recurrent Neural Networks (RNNs) and their variants, especially LSTMs, further advanced ECG analysis by modeling temporal dependencies in sequential data. [11] Developed a bidirectional LSTM model for ECG classification and showed superior performance compared to CNNs and conventional algorithms. In another study, [12] combined deep feature extraction with LSTM networks, achieving state-of-the-art results for arrhythmia detection. These findings confirm the suitability of LSTMs for capturing long-term dependencies and complex temporal patterns in ECG signals.

Most recently, hybrid architectures have been explored to enhance model performance. Like [12], which proposed a deep-coded feature approach integrating CNN and LSTM networks, achieving robust classification across multiple arrhythmia classes. Similarly, [13] applied LSTMs for anomaly detection in ECG signals, showing their potential in unsupervised learning scenarios. These hybrid and

advanced architectures indicate the continuous evolution of deep learning approaches in ECG-based diagnostics.

Despite these advancements, challenges remain in terms of dataset imbalance, computational cost, and real-time implementation. Many studies focus on arrhythmia detection using benchmark datasets like MIT-BIH, but broader applications in multi-disease classification and deployment in wearable healthcare devices are still limited. This research aims to address part of this gap by systematically benchmarking LSTM models against traditional classifiers, thereby providing insights into their comparative strengths and weaknesses in early detection of cardiac abnormalities.

III. METHODOLOGY

The methodology followed in this research contains dataset selection, a structured preprocessing pipeline with mathematical formulations, model architecture design, training configuration, and an evaluation framework.

A. Dataset

The MIT-BIH Arrhythmia Database from PhysioNet [7] was used in this study. It contains 48 annotated two-channel ambulatory ECG recordings sampled at 360 Hz. Each record includes waveform signal files (.dat, .hea) and expert beat annotations (.atr). All 48 records were processed to extract individual heartbeats for classification. The final dataset contains 112,572 heartbeat segments with 90,603 regular beats (class 0) and 21,969 abnormal beats (class 1), indicating a class imbalance ratio of 4.12:1. This dataset contains multiple arrhythmia types and is widely used as a benchmark for automated ECG analysis.

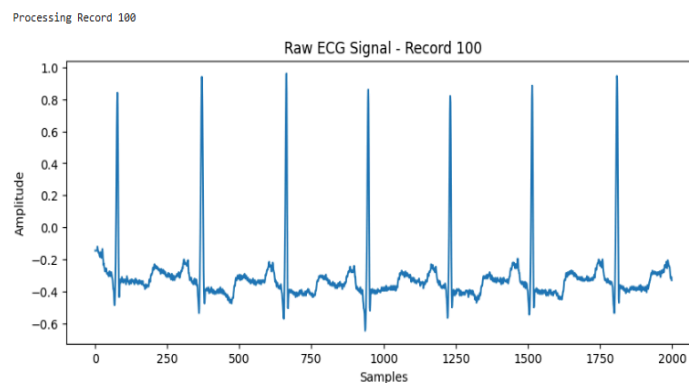


Fig. 1: Raw ECG signal from MIT-BIH Database

B. Preprocessing Pipeline

A consistent preprocessing pipeline was applied to improve signal quality and to prepare heartbeat segments for model training. The following preprocessing steps were applied sequentially to each ECG record. These steps make use of standard signal processing methods described in the arrhythmia classification literature [6], [10].

- Step 1: Signal Denoising (Bandpass Filtering)

Raw ECG signals contain baseline wander (low-frequency noise, $< 0.5 \text{ Hz}$) and high-frequency interference ($> 40 \text{ Hz}$). A Butterworth bandpass filter was applied with passband $0.5\text{-}40 \text{ Hz}$ was applied to remove noise components. The filter transfer function is:

$$H(f) = \frac{1}{\sqrt{\left(1 + \left(\frac{f}{f_c}\right)^{2n}\right)}}$$

where f_c is the cutoff frequency, and n is the filter order ($n = 4$). The filtered signal $x_{f(t)}$ was obtained by applying the filter in both forward and reverse directions to eliminate phase distortion.

- Step 2: Z-Score Normalization

Each ECG signal $x_{f(t)}$ was normalized using Z-score normalization to obtain a zero mean and unit variance.

$$x_{norm(t)} = \frac{(x_{f(t)} - \mu)}{\sigma}$$

where μ is the mean, and σ is the standard deviation of the signal $x_{f(t)}$ computed across the full record. This ensures that amplitude differences between records do not bias the classifier. A small constant ($\epsilon = 1 \times 10^{-8}$) was added to σ to prevent division by zero.

- Step 3: Heartbeat Segmentation

To extract individual heartbeat segments, annotated R-peak locations r_k were used as reference points. A fixed-length window $W = 200$ samples, approximately 556 ms at 360 Hz , was centered on each R-peak.

$$s_k = x_{norm} \left[r_k - \frac{W}{2} : r_k + \frac{W}{2} \right]$$

A fixed-length windowing preserves the temporal morphology of the P wave, QRS complex, and T wave within each beat. Segments at the record boundaries where the window exceeded the signal length were discarded.

- Step 4: Label Mapping

Beat annotation symbols from the MIT-BIH annotation standard were mapped to binary labels as follows:

$$y_k = 0 \text{ if } symbol_k \in \{N, L, R, e, j\} \text{ [Normal beat]}$$

$$y_k = 1 \text{ otherwise [Abnormal beat]}$$

The binary mapping follows established practice in automated arrhythmia screening [5], where all non-sinus beat types are treated as clinically significant abnormalities.

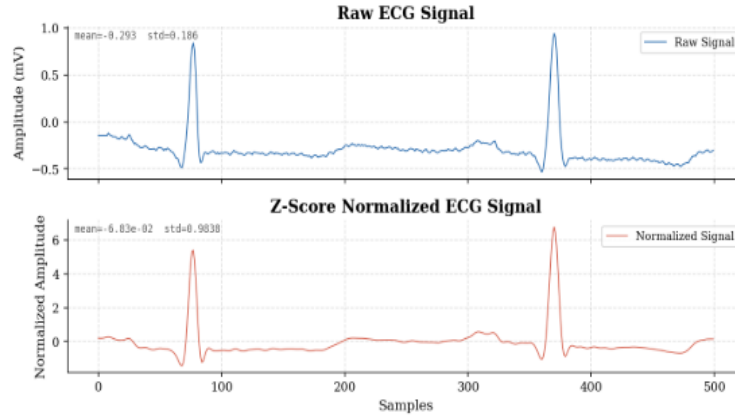


Fig. 2: Raw vs. Z-score Normalized ECG signal

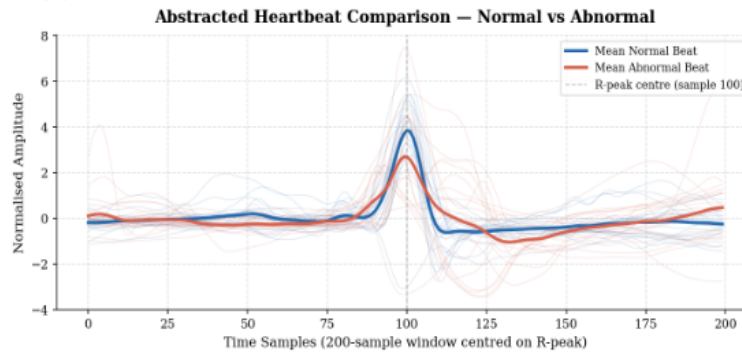


Fig. 3: Extracted heartbeat segments

C. Detailed Algorithm Flow

Algorithm 1: ECG Arrhythmia Classification Pipeline

Input: Raw ECG records from MIT-BIH (48 records, sampled at 360 Hz)

Output: Binary label per heartbeat (0 = Normal, 1 = Abnormal)

1. Load each ECG record and apply Butterworth bandpass filter (0.5-40 Hz) to obtain the filtered signal $x_{f(t)}$
2. Normalize the filtered signal using Z-score normalization: $x_{norm(t)} = \frac{(x_{f(t)} - \mu)}{\sigma}$ **for** each annotated R-peak r_k **do**
3. Extract a heartbeat segments: $s_k = x_{norm}[r_k - 100 : r_k + 100]$
4. Convert map annotation symbol to binary label $y_k \in \{0, 1\}$

5. Append (s_k, y_k) to dataset D
6. end
7. Split dataset D into training and testing sets (80% / 20%) using stratified sampling
8. Train the Bi-LSTM model using the training dataset
9. Train SVM, Logistic Regression, Random Forest classifiers using the flattened training samples
10. for each trained model do
11. Predict labels for the testing dataset and compute the confusion matrix
12. Calculate Accuracy, Precision, Recall, F1-score, Specificity, FPR, AUC-ROC
13. end
14. return the classification reports and confusion matrices for all models

D. Model Development

Four models were implemented under identical data splits:

- Bidirectional LSTM: An existing Bi-LSTM architecture was implemented using TensorFlow/Keras. The architecture consists of a Bi-LSTM layer with a total of 128 and 64 units per direction, followed by a Dropout rate of 0.3 for regularization, and a Dense output layer with sigmoid activation for binary classification. This architecture is consistent with the standard Bi-LSTM configuration described in [11].

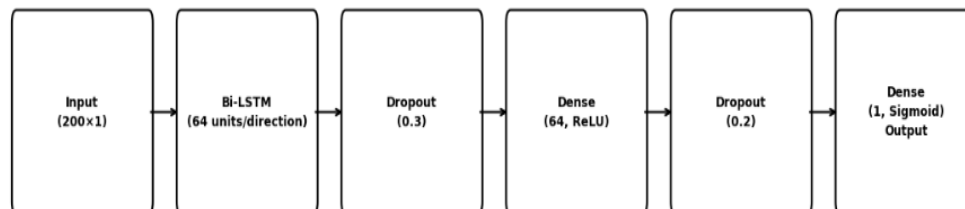


Fig. 4: Architecture of the Implemented Bi-LSTM Model

- Support Vector Machine: SVM was used as a strong baseline classifier for high-dimensional data, as it is effective for binary classification problems with clear class separation.
- Logistic Regression: It was implemented as an interpretable linear baseline model for binary classification.
- Random Forest: Random Forest was used as an ensemble learning method capable of modeling non-linear decision boundaries and handling noisy data.

E. Training Configuration

The dataset was split 80/20 into training and test sets using stratified sampling to maintain class proportions. The Bi-LSTM was trained up to 30 epochs with early stopping with a batch size of 64, binary cross-entropy loss, and the Adam optimizer

(learning rate = 0.001). Classical models received the same 80% training split, with heartbeat windows flattened to 1D feature vectors of size 200.

F. Evaluation Metrics

The model was evaluated using the following metrics (TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative):

- Accuracy: $(TP + TN) / (TP + TN + FP + FN)$ - The proportion of correctly predicted segments.
- Precision: $TP / (TP + FP)$ - The ability to identify only relevant instances.
- Recall (Sensitivity): $TP / (TP + FN)$ - The capacity to detect all relevant instances.
- F1-Score: $2 \times (Precision \times Recall) / (Precision + Recall)$ - It is the harmonic mean of precision and recall.
- Specificity: $TN / (TN + FP)$ - It is the model's ability to correctly identify normal beats. Clinically critical; low specificity leads to false alarms, which causes an alarm fatigue in monitoring systems and eroding clinician trust.
- False Positive Rate (FPR): $FP / (FP + TN)$ - It is known as a proportion of normal beats incorrectly classified as normal. In a system processing 100,000 beats per day, a 1% FPR generates 1,000 unnecessary alerts daily.
- AUC-ROC - It is a threshold-independent measure of discriminative ability essential for imbalanced datasets where a fixed 0.5 threshold is sub-optimal.

IV. RESULTS AND DISCUSSION

A. Dataset Characteristics

Following the preprocessing pipeline, a total of 112,572 heartbeat segments were extracted from all 48 MIT-BIH records. The class distribution was: Normal (class 0) 90,603 segments and Abnormal (class 1) 21,969 segments, containing a class imbalance ratio of 4.12:1. The key analytical objective of this study is to preserve this imbalance without artificial correction to reflect realistic clinical conditions and to directly observe the effect of imbalance on model behavior.

The extracted heartbeat segments have a uniform length of 200 samples $\approx 556ms$, which captures the complete P-QRS-T complex for each beat.

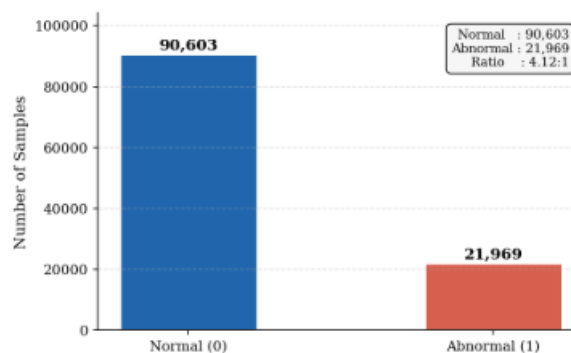


Fig. 5: Class Distribution of extracted heartbeats

B. Training Behavior of Bi-LSTM

The Bi-LSTM was trained over 16 epochs with early stopping. The training loss showed a general decreasing trend from approximately 0.31 to 0.25. However, the validation loss exhibited notable instability, with a significant spike at epoch 12-13 reaching approximately 0.52, before recovering toward the end of training. Despite this fluctuation, the model restored the best weights from the epoch with the lowest validation loss, as implemented through the EarlyStopping callback. The validation accuracy remained generally above 0.90 throughout training, indicating reasonable generalization on the held-out validation set. The observed instability in validation loss is likely attributable to the class imbalance in the dataset, which caused inconsistent gradient updates during certain epochs.

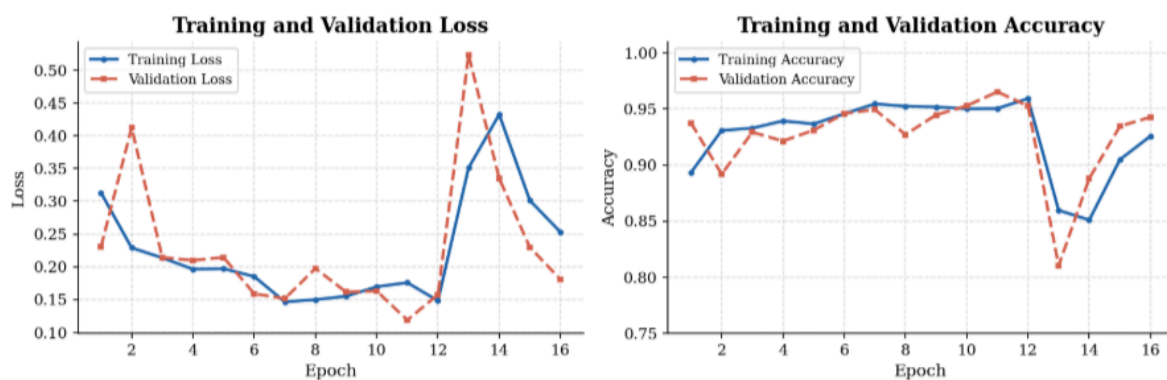


Fig. 6: Training and validation loss and accuracy curves for the Bi-LSTM model

C. Quantitative Performance Comparison

Table 1 presents the complete evaluation of all four models across all seven metrics. The three domain-specific metrics: Specificity, FPR, and AUC-ROC are included alongside standard evaluation measures.

Table 1: Quantitative Performance Comparison

Model	Accuracy	Precision	Recall	F1-score	Specificity	FPR	AUC-ROC
Bi-LSTM	0.9641	0.8841	0.9392	0.9108	0.9701	0.0299	0.9867
SVM	0.9813	0.9473	0.9577	0.9525	0.9871	0.0129	0.9933
Logistic Regression	0.8649	0.6171	0.8116	0.7011	0.8779	0.1221	0.9029
Random Forest	0.9876	0.9881	0.9479	0.9676	0.9972	0.0028	0.9984

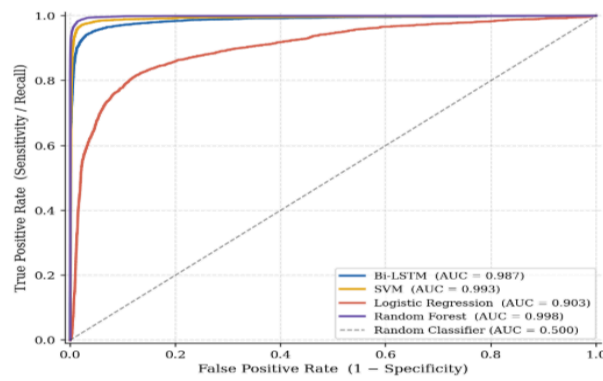


Fig. 7: ROC curves for all models

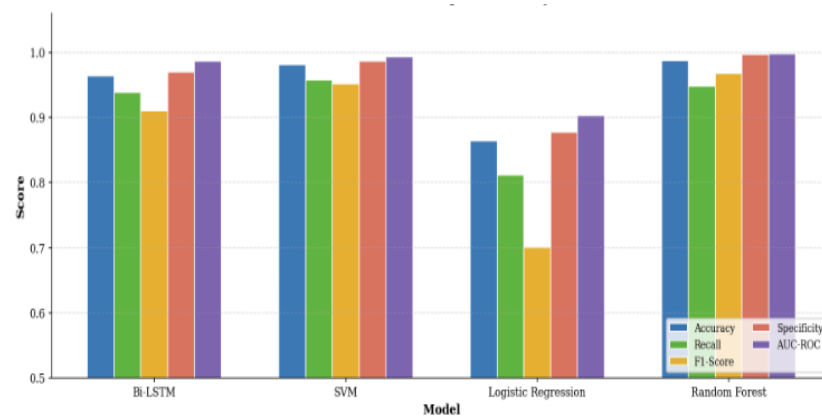


Fig. 8: Side-by-side comparison of key performance metrics across all four classifiers

D. Confusion Matrix Analysis of Models

Confusion matrices for all four models were computed on the held-out test set. The following discussion interprets the LSTM and Random Forest confusion matrices in terms of specific error counts and their clinical implications.

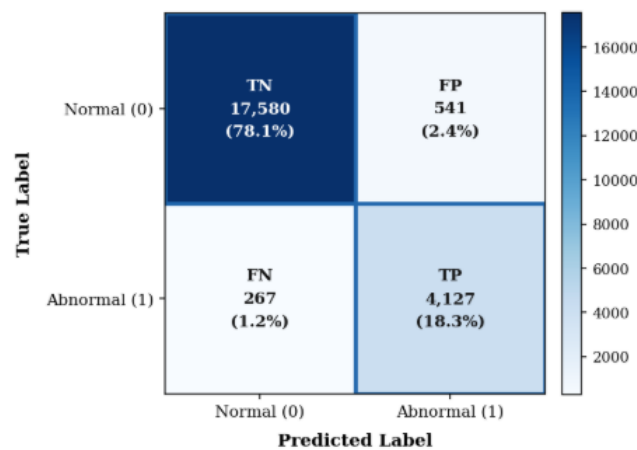


Fig. 9: Confusion Matrix of Bi-LSTM

For the Bi-LSTM, the model correctly classified 17,580 normal beats (TN) and 4,127 abnormal beats (TP). It misclassified 541 normal beats as abnormal (FP) and 267 abnormal beats as normal (FN). The 267 false negatives represent missed arrhythmias, which are a more clinically dangerous error type since a missed abnormal beat may result in an untreated cardiac event. The Bi-LSTM achieved a recall of 0.9392 for the abnormal class, confirming that the model failed to detect approximately 6% of true arrhythmias under these conditions.

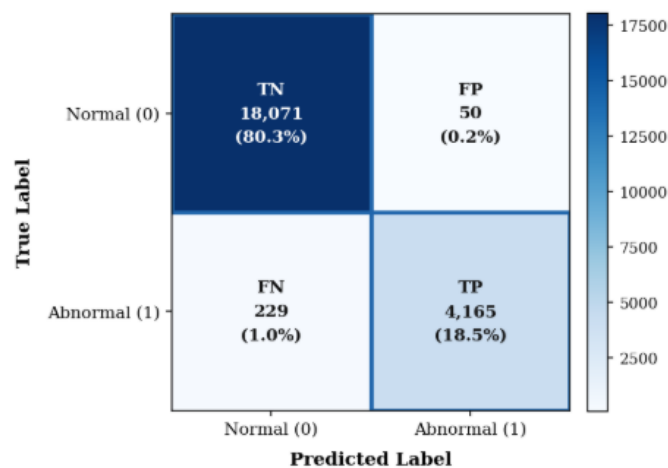


Fig. 10: Confusion Matrix of Random Forest

For Random Forest, the model correctly classified 18,071 normal beats (TN) and 4,165 abnormal beats (TP), with only 50 false positives and 229 false negatives. Random Forest produced fewer false positives which is 50 vs. 541 compared to the Bi-LSTM, and specificity of 0.9972 vs. 0.9701 and FPR of 0.0028 vs. 0.0299. This difference corresponds to approximately 2,710 additional false alerts daily in a high volume clinical monitoring scenario processing 100,000 beats per day.

For Logistic Regression and SVM, among all four classifiers, logistic regression produced the highest false negative, which is reflected on its recall of 0.8116 on the abnormal class. This data confirms that a linear decision boundary is insufficient to separate normal and abnormal beat morphologies in the normalized ECG feature space and, that the classification task is inherently non-linear. Whereas, SVM performed competitively with Random Forest with an accuracy of 0.9813 and AUC-ROC of 0.9933.

E. Analysis of Model Performance Differences

Random forest and SVM marginally outperformed the Bi-LSTM with the accuracy of 0.9876 and 0.9813 vs. 0.9641. This is explained by three factors specific to our setup. First, fixed-length R-peak-centered segmentation produces structurally consistent heartbeat windows by reducing the temporal modeling advantage of Bi-LSTM compared with classical machine learning classifiers. Second, the 4.12:1 class imbalance may have affected Bi-LSTM training more than ensemble methods, which are often more robust to class imbalance due to their bagging-based learning strategy. Third, the dataset may not be large enough for the Bi-LSTM to fully exploit its representation learning capability although it contains approximately 90,000 samples compared with classical machine learning models.

F. Limitations

This study has the following limitations: first, results are based on a single database (MIT-BIH), and cross-dataset generalization is invalidated. Second, only binary classification was performed. Third, class imbalance was intentionally not addressed through resampling or class-weighting, and a single 80/20 split was used without k-fold cross-validation. Lastly, there is no clinical validation against cardiologist interpretations are conduct.

V. CONCLUSIONS, SUGGESTIONS, AND RECOMMENDATIONS

In this study, we implemented a custom Bi-LSTM network and compared it against SVM, Logistic Regression, and Random Forest for binary arrhythmia classification on the MIT-BIH Arrhythmia Database. 112,572 heartbeat segments were retrieved from 48 records using a preprocessing pipeline that included bandpass filtering, Z-score normalization, and R-peak centered segmentation. Under an identical conditions, a Random Forest with a highest accuracy of 0.9876 and lowest FPR of 0.0028 outperformed Bi-LSTM with an accuracy of 0.9641 and FPR of 0.0299. Despite the class imbalance, the Bi-LSTM showed good discriminative abilities with a strong AUC-ROC of 0.9867. Logistic Regression with an accuracy of 0.8649 and recall of 0.8116 performed poorest confirming the task requires non-linear decision boundaries. The observed class imbalance 4.12:1 likely contributed to reduced minority-class recall across all evaluated models.

In the future, this project can be focus on:

- Cross-dataset validation on external databases such as PTB-XL.
- Apply SMOTE or class-weighted loss to address imbalance.
- Extend to multi-class arrhythmia classification.
- Adding a self-attention mechanism to the Bi-LSTM.
- Clinical validation against cardiologist interpretations.

Despite using a benchmark dataset, this work provides one of the first comparative evaluations of Bi-LSTM and classical machine learning models for ECG arrhythmia classification while discussing their potential applicability in resource-constrained healthcare settings.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the PhysioNet project for providing access to the MIT-BIH Arrhythmia Database, which was used in this research. We would also like to acknowledge the researchers and contributors of the publicly available ECG datasets and algorithms used in this study.

We also acknowledge the financial and institutional support provided by ICAS under the UESC program for supporting this research work.

AUTHOR CONTRIBUTIONS

The authors were responsible for the complete research work, including literature review, dataset preparation, preprocessing, model development, experimentation, analysis, and manuscript writing.

DATA AVAILABILITY

The dataset used in this research is publicly available from the MIT-BIH Arrhythmia Database provided by PhysioNet.

CONFLICT OF INTEREST

We declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

- [1] World Health Organization, "Cardiovascular diseases (CVDs)." Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] P. Kligfield *et al.*, "Recommendations for the Standardization and Interpretation of the Electrocardiogram," *Circulation*, vol. 115, no. 10, pp. 1306–1324, Mar. 2007, doi: 10.1161/CIRCULATIONAHA.106.180200.
- [3] I. Dotsinsky, "Review of 'Advanced Methods and Tools for ECG Data Analysis', by Gari D. Clifford, Francisco Azuaje and Patrick E. McSharry (Editors)," *Biomed. Eng. Online*, vol. 6, no. 1, p. 18, 2007, doi: 10.1186/1475-925X-6-18.
- [4] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.
- [5] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001, doi: 10.1109/51.932724.
- [6] B. Surawicz, R. Childers, B. J. Deal, and L. S. Gettes, "AHA/ACCF/HRS Recommendations for the Standardization and Interpretation of the Electrocardiogram," *Circulation*, vol. 119, no. 10, Mar. 2009, doi:

- 10.1161/CIRCULATIONAHA.108.191095.
- [7] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.,” *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.CIR.101.23.E215/ASSET/9716B30D-2E4D-4D06-8F9E-A21157DFE97C/ASSETS/GRAPHIC/HC2304183003.JPEG.
 - [8] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, “Deep learning for healthcare applications based on physiological signals: A review,” *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018, doi: 10.1016/J.CMPB.2018.04.005.
 - [9] R. Ceylan and Y. Özbay, “Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network,” *Expert Syst. Appl.*, vol. 33, no. 2, pp. 286–295, Aug. 2007, doi: 10.1016/J.ESWA.2006.05.014.
 - [10] S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016, doi: 10.1109/TBME.2015.2468589.
 - [11] Ö. Yildirim, “A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification,” *Comput. Biol. Med.*, vol. 96, pp. 189–202, May 2018, doi: 10.1016/J.COMPBIOMED.2018.03.016.
 - [12] O. Yildirim, U. B. Baloglu, R. S. Tan, E. J. Ciaccio, and U. R. Acharya, “A new approach for arrhythmia classification using deep coded features and LSTM networks,” *Comput. Methods Programs Biomed.*, vol. 176, pp. 121–133, Jul. 2019, doi: 10.1016/J.CMPB.2019.05.004.
 - [13] S. Chauhan and L. Vig, “Anomaly detection in ECG time signals via deep long short-term memory networks,” *Proc. 2015 IEEE Int. Conf. Data Sci. Adv. Anal. DSAA 2015*, Dec. 2015, doi: 10.1109/DSAA.2015.7344872.