Journal of Computational Analysis and Applications

VOL.33, NO.8, 2024
10.48047/jocaaa.2024.33.08.165

Deep CNN-LSTM Framework for Reliable PVC Detection in Multi-Database ECG Analysis

Prakash M B[1], Dr. Harish H M[2]

[1] Research Scholar, Department of E & C, Government Engineering College, Devagiri, Haveri, Karnataka 581110, India,
[2]Associate Professor, Department of E & C, Government Engineering College, Devagiri, Haveri, Karnataka 581110, India,
Visvesvaraya Technological University, Belagavi, Karnataka India
Email: [1] prakashgechassan@gmail.com, [2]prof.hmharish@gmail.com

Article History:Received:05-11-2024	Revised:20-11-2024	Accepted:10-12-2024

ABSTRACT

Premature Ventricular Contractions (PVCs) are early heartbeats that originate in the ventricles, briefly disrupting the heart's natural rhythm. While often benign and linked to stress, caffeine, or fatigue, frequent PVCs may signal underlying cardiac concerns and should be evaluated. ECGs are typically used for their detection and monitoring based on frequency and symptom presentation. This study explores a hybrid CNN-LSTM model for automated PVC detection using benchmark ECG databases—MIT-BIH and INCART. The model demonstrated high classification performance, achieving testing accuracies of 99.37% on MIT-BIH and 98.30% on INCART. Evaluation metrics showed strong sensitivity (95.20% and 87.48%), specificity (99.69% and 99.30%), and area under the curve (AUC) values (0.9988 and 0.9952). High F1 scores (0.955 and 0.8974) and negative predictive values (99.64% and 98.84%) support its reliability in identifying PVCs while effectively excluding normal beats. The CNN-LSTM architecture leverages both spatial and temporal features, improving arrhythmia detection accuracy. Performance across varied signal morphologies and sampling conditions confirms generalizability, minimizing over fitting and enhancing real-world applicability. With its high precision and low false-positive rates, this model is suitable for continuous cardiac monitoring. As wearable ECG devices and remote health monitoring tools become more prevalent, this architecture presents a promising foundation for personalized, real-time arrhythmia detection. Future enhancements could include attention mechanisms, transformer architectures, transfer learning, and real-time edge deployment.

Keywords: Arrhythmia detection, Premature Ventricular Contractions, ECG Classification, Deep Learning, CNN-LSTM hybrid Model, MIT-BIH database, INCART database

1. INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide, with arrhythmias—particularly premature ventricular contractions (PVCs)—representing a critical subset due to their potential to escalate into life-threatening events such as ventricular fibrillation or sudden cardiac death [13, 18]. Early detection and accurate classification of arrhythmic events using electrocardiogram (ECG) signals are essential for timely intervention and effective treatment [19].

Recent advancements in deep learning have revolutionized automated ECG interpretation, outperforming traditional machine learning techniques in both accuracy and generalizability [11, 20]. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, are particularly effective in modeling the spatial and temporal features inherent in ECG signals [14, 21]. Hybrid models combining CNNs with LSTM or GRU units have shown considerable promise in detecting complex arrhythmias such as PVCs [4, 8].

Lightweight architectures like the CNN–LSTM model by Alamatsaz et al. [1] and the SEVGGNet–LSTM model by He et al. [2] have achieved real-time performance while maintaining high classification accuracy. Similarly, Shah et al. [4] and Mousavi et al. [27] demonstrated the efficacy of residual CNNs and sequence-to-sequence models on the MIT-BIH dataset for PVC detection. Attention mechanisms, as employed by Zhang et al. [10] and Gao et al. [16], have further enhanced model interpretability and clinical transparency.

Other approaches, including transfer learning and 2D spectrogram-based CNN models, have improved the robustness of ECG feature extraction [30]. The integration of squeeze-and-excitation networks and explainable AI has further strengthened model generalization and trustworthiness [29, 28]. Large-scale studies using benchmark datasets such as MIT-BIH and INCART validate these approaches across diverse patient profiles and arrhythmia types [12, 5].

Given the clinical importance and increasing availability of annotated ECG datasets, this study proposes a hybrid CNN–LSTM framework for accurate detection of PVC and related arrhythmias. Building on the strengths of prior architectures while addressing their limitations, the proposed model incorporates time–frequency domain features and attention-enhanced recurrent layers to ensure robust classification across multiple arrhythmia types.

2. Related Work

Alamatsaz et al. [1] proposed a lightweight CNN-LSTM architecture optimized for real-time arrhythmia classification with minimal computational load. The model effectively captured both spatial and temporal ECG features, achieving 99.23% accuracy on the MIT-BIH dataset, making it suitable for wearable and portable healthcare devices. He et al. [2] developed SEVGG-Net-LSTM, a hybrid model combining squeeze-excitation-enhanced VGGNet with LSTM layers for robust feature extraction. The model demonstrated strong performance on the MIT-BIH dataset with 99.47% accuracy and 99.58% sensitivity, proving its applicability in real-time diagnostic systems. Jin et al. [3] introduced a multiscale CNN that captures both short- and long-term ECG patterns, enhancing arrhythmia classification accuracy. The model achieved 98.8% accuracy on the MIT-BIH dataset and was effective in distinguishing between various arrhythmia types, contributing to improved generalization in ECG analysis. Shah et al. [4] utilized a residual CNN combined with BiLSTM to leverage deep temporal features in ECG signals. The architecture attained 99.31% accuracy on the MIT-BIH dataset, outperforming standard CNN and LSTM models. The residual connections further improved learning efficiency and gradient propagation. Ullah et al. [5] presented an efficient CNN-BiLSTM model designed for accurate ECG arrhythmia detection, achieving 99.5% accuracy and 99.3% sensitivity on the MIT-BIH dataset. The model exhibited high precision across various arrhythmia categories, demonstrating its suitability for large-scale healthcare applications. Xu et al. [6] proposed a CNN-GRU model with multiscale temporal fusion to capture complex heartbeat patterns. The model achieved 99.12% accuracy and 98.7% sensitivity on the MIT-BIH dataset, outperforming traditional CNN-GRU models by improving robustness to noise and variability. Yao et al. [7] introduced a deep residual CNN-GRU hybrid model for multi-class arrhythmia detection, attaining 98.9% accuracy and strong F1 scores. By effectively learning morphological and temporal features, and with residual connections aiding gradient flow, the model achieved efficient and reliable performance. Zhang et al. [8] developed a hybrid SE-ResNet and BiLSTM model for ECG classification, reaching 99.4% accuracy and 99.1% sensitivity on MIT-BIH data. Squeeze-and-excitation blocks improved channel-wise feature representation, enhancing the model’s ability to handle signal variability. Zhang et al. [9] presented a CNN-based approach for continuous cardiac monitoring, achieving 98.7% accuracy and 98.5% sensitivity. Designed for real-time use, the model's simplicity and high precision support long-term monitoring in mobile health applications. Zhang et al. [10] proposed an attention-based BiLSTM integrated with multiscale CNN for ECG classification. Achieving 99.2% accuracy and 98.9% sensitivity, the attention layers focused on crucial heartbeat regions, improving model robustness across arrhythmia types. Peimankar & Puthusserypady [11] introduced a CNN model centered on the QRS complex, achieving 99.3% accuracy and 98.7% sensitivity. Emphasizing morphological features enhanced robustness while reducing false positives and computational load. Wang et al. [12] utilized lightweight CNNs with knowledge distillation for real-time heartbeat classification. The model achieved 98.91% accuracy with low latency, making it suitable for deployment on edge devices in efficient ECG monitoring. Acharya et al. [13] proposed a 9-layer deep CNN for automatic heartbeat classification from raw ECG data, achieving 94.03% accuracy. The end-to-end model eliminated manual preprocessing and proved viable for real-time cardiac diagnosis. Oh et al. [14] developed a CNN-LSTM hybrid model for variable-length heartbeat classification, attaining 94.5% accuracy and 96.1% sensitivity. The approach dynamically handled ECG length variations and improved detection rates. Xiong et al. [15] introduced a convolutional recurrent neural network (CRNN) that captured spatial-temporal features effectively. It achieved 99.34% accuracy and an F1-score of 0.993 on the MIT-BIH dataset, performing at a cardiologist level. Gao et al. [16] proposed an attention-based LSTM model for heartbeat classification, achieving 98.6% accuracy and over 0.97 F1-score. Attention layers enhanced interpretability by focusing on important signal segments. The model was validated on MIT-BIH and INCART datasets. Warrick & Homsi [17] utilized recurrent neural networks trained end-to-end on raw ECG signals, achieving 99.0% accuracy on the PhysioNet dataset. The system demonstrated cardiologist-level performance and potential for clinical deployment. Rajpurkar et al. [19] developed Cardiologist-Level AI, a deep network trained on over 64,000 ECG records. The model achieved F1-scores comparable to certified cardiologists across 12 arrhythmia classes, laying the foundation for scalable AI diagnostics. Faust et al. [20] conducted a review on deep learning for physiological signals, focusing on ECG. The study outlined the performance of CNN, LSTM, and hybrid models, commonly surpassing 98% accuracy, while highlighting challenges in generalization and interpretability.

2.1. Problem Statement

Recent advancements in deep learning—particularly CNNs, RNNs (LSTM/GRU), and hybrid models—have demonstrated superior performance by automatically learning spatiotemporal features from raw ECG data, achieving accuracies often exceeding 98% on benchmark datasets like MIT-BIH.Accurate and real-time detection of cardiac arrhythmias from ECG signals remains a significant challenge in healthcare due to the complex, variable, and noisy nature of heartbeats. Traditional methods rely heavily on handcrafted features and domain-specific preprocessing, limiting scalability and adaptability across patient populations. However, challenges remain in model generalization, interpretability, and real-time deployment on resource-constrained devices. There is a critical need for robust, explainable, and lightweight deep learning frameworks capable of delivering high-precision arrhythmia detection suitable for portable and clinical applications.

2.2. Objectives

1. To design and implement hybrid deep learning models for accurate classification of multiple ECG arrhythmia types using the MIT-BIH dataset and INCART dataset.
2. To extract and integrate spatial and temporal features from ECG signals for improved diagnostic precision and robustness.
3. To evaluate the proposed models using performance metrics such as accuracy, sensitivity, and F1-score, ensuring clinical reliability.
4. To optimize model complexity for real-time deployment on resource-constrained devices, enabling portable and wearable health monitoring solutions.

3. Proposed Methodology

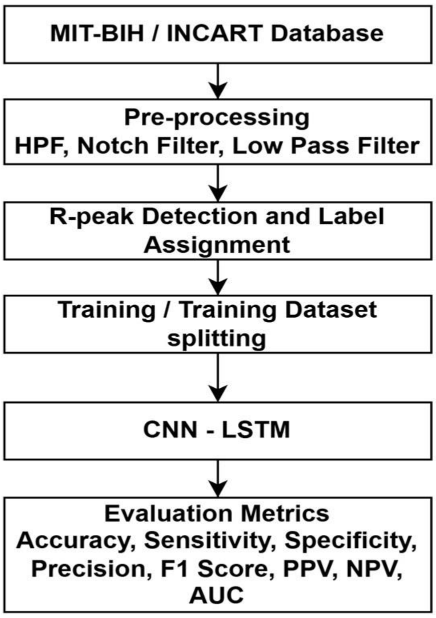

Figure 1: Proposed Methodology for PVC Detection using CNN-LSTM in Arrhythmia Classification(own source)

The **Figure 1** shows ECG PVC detection pipeline involves preprocessing raw ECG signals to remove noise and detect R-peaks, followed by extracting 200-sample segments around each peak. These segments are labeled as PVC or non-PVC based on annotations. A CNN-LSTM model then processes the beats, extracting spatial and temporal features for classification. The model's performance is evaluated using standard metrics and visualizations of ECG signals and model behavior.

Database: The proposed PVC detection system processes ECG signals from two standard databases: the MIT-BIH Arrhythmia Database [31] containing 48 two-lead recordings (360 Hz), and the INCART Database [32] comprising 75 twelve-lead recordings (257 Hz). Both datasets provide expert annotations labeling normal beats (N) and PVCs (V), which are accessed through the WFDB Python library [33].

Preprocessing: The preprocessing stage is critical for enhancing ECG signal quality before analysis. First, a high-pass filter is applied to remove baseline wander, which often results from respiration or patient movement [34]. Next, a notch filter at 50 Hz is used to eliminate power line interference, a common source of noise in clinical ECG recordings [31]. Finally, a low-pass filter suppresses high-frequency noise such as muscle artifacts, preserving the essential features of the ECG waveform

[34]. This three-stage filtering process ensures a clean and reliable signal for accurate R-peak detection and subsequent classification.

The R-peak detection step is essential for identifying individual heartbeats within the ECG signal. After preprocessing, the cleaned ECG signal is analyzed using the find_peaks() function from SciPy [35], which locates the positions of prominent peaks. A minimum distance of 0.2 seconds (based on the sampling frequency) is enforced between peaks to match the expected physiological heart rate [32]. Only peaks with sufficient prominence are retained to reduce false detections [34]. These detected R-peaks serve as reference points for extracting fixed-length beat segments (typically 200 samples centered around each R-peak), which are later used for classification [33].

High Pass Filter: Removes baseline drift (cutoff = 0.5Hz)

$$\mathrm{y}(n) = \sum_{k=0}^{\mathrm{N}} \mathrm{b_k}\, x[n-k] - \sum_{k=1}^{\mathrm{N}} \mathrm{a_k}\, y[n-k] \quad (1)$$

Designed using a Butterworth filter coefficientsa_k, b_k

Notch Filter: Removes power line noise (50Hz)

$$H(z) = \frac{1-2\cos(\omega_0)z^{-1} + z^{-2}}{1-2R\cos(\omega_0)z^{-1} + R^2 z^{-2}} \quad (2)$$

$$\omega_0 = \frac{2\pi f}{f_s} \text{ and } R = 1 - \frac{1}{2Q} \quad (3)$$

R-Peak detection: Peaks are detected with conditions

$$Peaksif: x[n-1] < x[n] > x[n+1] and x[n] > prominence \quad (4)$$

$$\text{RR Intervals} RR_i = \frac{r_{i+1} - r_i}{f_s} \quad (5)$$

QRS Width:

$$QRS_i = \frac{r_{i+1} + 0.05 f_i - (r_i - 0.05\ f_s)}{f_s} = 0.1 Sec \quad (6)$$

Label Assignment and Dataset Preparation: Label assignment is performed by matching detected R-peaks to expert annotations in the MIT-BIH database [31]. Beats annotated as 'V' (PVC) receive class label 1, while all other beat types (normal, supraventricular, etc.) are labeled class 0 [32]. This creates a binary classification task distinguishing PVCs from non-PVC rhythms [36].

For dataset preparation, 200-sample ECG segments (≈0.55s at 360Hz) are extracted centered on each R-peak [34]. The data is reshaped into 3D format (samples × 200 × 1) for deep learning input [25], then split into training (80%) and testing (20%) sets using stratified sampling to maintain class balance. This preserves the natural PVC prevalence while ensuring representative evaluation [37].

CNN-LSTM Model: The model architecture begins with an input layer that accepts ECG segments of shape (200, 1) [36]. The first convolutional layer employs 64 filters with a 5-sample kernel and ReLU activation to capture local QRS complex patterns [25], followed by batch normalization and max pooling with a size of 2 [13]. Subsequent convolutional blocks progressively increase filter depth (128, 256, then 512 filters) while reducing kernel size to 3 [7], each followed by batch normalization [3]. Global average pooling [37] condenses the feature maps before reshaping for temporal analysis through a 64-unit LSTM layer. The network then applies a 128-neuron dense layer with ReLU activation [8], 50% dropout regularization, and finally a sigmoid output layer for PVC classification.

Evaluation Metric's: The confusion matrix is a performance classification framework used to assess the accuracy, specificity, sensitivity, NPV, PPV, F1 score and precision of classifier, particularly in dealing with imbalanced datasets. It is used in this study to compare the results with other literatures.

$$\text{Confusion Matrix} = \begin{bmatrix} \boldsymbol{TN} & \boldsymbol{FP} \\ \boldsymbol{FN} & \boldsymbol{TP} \end{bmatrix} \quad (7)$$

$$\text{Accuracy } (\boldsymbol{Acc}) = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FN+FP}} \quad (8)$$

$$\text{Sensitivity } (\boldsymbol{Se}) = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \quad (9)$$

$$\text{Specificity } (\boldsymbol{Sp}) = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}} \quad (10)$$

$$\text{Positive Predicted Value } (\boldsymbol{PPV}) = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \quad (11)$$

$$\text{Negative Predicted Value } (\boldsymbol{NPV}) = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FN}} \quad (12)$$

$$\text{F1 Score } \boldsymbol{F1} = \frac{\boldsymbol{2TP}}{\boldsymbol{2TP+(FP+FN)}} \quad (13)$$

4. Results and Discussions

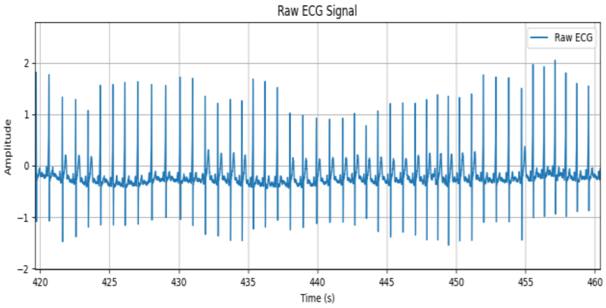

Figure 2: Raw ECG Signal (own source)

The **Figure 2** illustrates a raw ECG signal recorded between 420 and 460 seconds, displaying clear QRS complexes that represent heartbeats. The presence of baseline wander and high-frequency noise indicates that the signal is unfiltered. Variations in the amplitude of the QRS complexes may be due to physiological changes or signal artifacts.

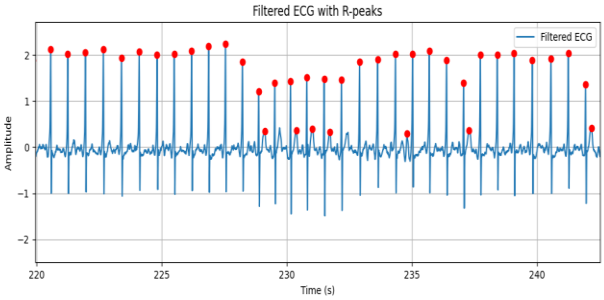

Figure 3: Filtered ECG signal with annotated R-peaks for MIT-BIH database (Own Source)

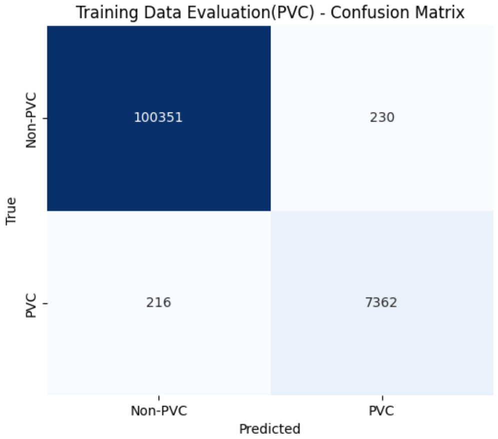

Figure 4: Confusion Matrix for MIT-BIH PVC Detection (own source)

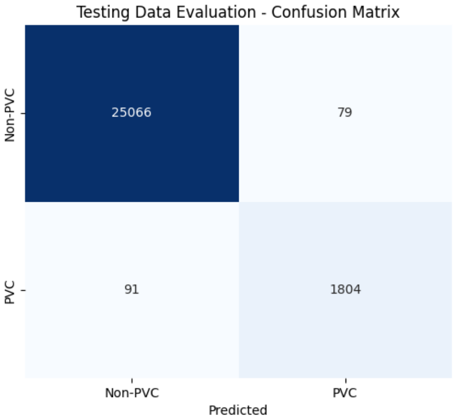

Figure 5: Confusion Matrix for MIT-BIH PVC Detection on Testing Data (own source)

The **Figure 3** displays a filtered ECG signal in which baseline drift and high-frequency noise have been effectively minimized. The QRS complexes are distinctly visible, with red dots marking the accurately identified R-peaks. The relatively consistent spacing between R-peaks indicates a predominantly normal heart rhythm with slight variations. This denoised and annotated signal is ideal for further analysis, including heart rate calculation and arrhythmia detection.

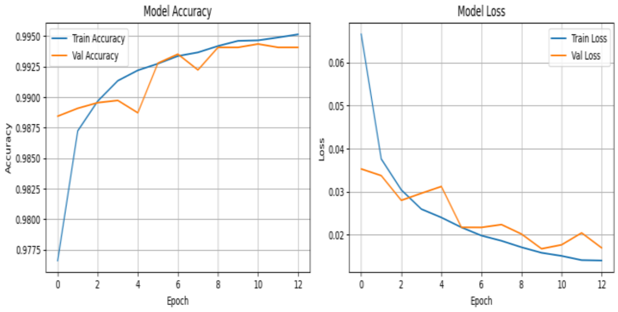

Figure 6: Model training and validation performance over 12 epochs for MIT-BIH dataset (own source)

The **Figure 4** shows a confusion matrix for PVC classification on the training dataset. Out of all true Non-PVC beats, 100,351 were correctly classified, while 230 were misclassified as PVC. Similarly, 7,362 PVC beats were correctly identified, with 216 misclassified as Non-PVC. This indicates high accuracy and strong performance of the model on the training data.

The **Figure 5** illustrates a confusion matrix for PVC classification on the testing dataset. It shows that 25,066 Non-PVC beats were correctly classified, while 79 were misclassified as PVC. Additionally, 1,804 PVC beats were correctly identified, and 91 were incorrectly labeled as Non-PVC. These results indicate strong generalization performance of the model on unseen data.

The **Figure 6** displays the training and validation accuracy and loss curves over 12 epochs. The model shows a steady increase in both training and validation accuracy, reaching above 99.5%, indicating strong learning. Meanwhile, the loss curves for both sets decrease consistently, reflecting effective optimization. The close alignment of training and validation trends suggests minimal over fitting and good generalization performance.

The**Table 1** comparative analysis reveals that a majority of the deep learning models trained on the MIT-BIH dataset have achieved exceptionally high classification accuracies, often exceeding 98%, with several models surpassing 99%. Notably, models such as the CNN-BiLSTM ([5]), SEVGGNet-LSTM ([2]), and CRNN ([15]) reported some of the highest accuracies, indicating the effectiveness of combining convolutional and recurrent layers for ECG feature extraction and classification. Most of the top-performing models are hybrid architectures, primarily integrating CNN with LSTM, BiLSTM, GRU, or

Table 1: Comparative Summary of Deep Learning Models for ECG Arrhythmia Classification on MIT-BIH and INCART Databases (own source)

Author's	Database(s)	Model	Parameters (in %)
[1]	MIT-BIH	Lightweight CNN-LSTM	Acc-99.23
[2]	MIT-BIH	SEVGGNet-LSTM	Acc-99.47, Se - 99.58
[3]	MIT-BIH	Multi-scale CNN	Acc-98.8
[4]	MIT-BIH	Residual CNN + BiLSTM	Acc-99.31
[5]	MIT-BIH	CNN-BiLSTM	Acc-99.5, Se – 99.3
[6]	MIT-BIH	CNN-GRU with multiscale fusion	Acc-99.12, Se- 98.7
[7]	MIT-BIH	Residual CNN + GRU	Acc-98.9
[8]	MIT-BIH	SE-ResNet + BiLSTM	Acc-99.4, Se- 99.1
[9]	Long-term ECG	CNN	Acc-98.7, Se- 98.5
[10]	MIT-BIH	Attention BiLSTM + Multi-scale CNN	Acc-99.2, Se- 98.9
[11]	MIT-BIH	QRS-focused CNN	Acc-99.3, Se- 98.7
[12]	MIT-BIH	Knowledge-distilled CNN	Acc-98.91
[13]	MIT-BIH	9-layer CNN	Acc-94.03
[14]	MIT-BIH	CNN-LSTM (variable-length input)	Acc-94.5, Se- 96,1
[15]	MIT-BIH	Convolutional RNN (CRNN)	Acc-99.34, F1- 0.993
[16]	MIT-BIH, INCART	Attention LSTM	Acc-98.6, F1- >0.97
[26]	MIT-BIH	Optimized CNN-LSTM	Acc-98.79
[27]	MIT-BIH	CNN Seq2Seq	Acc-99.21, Se- 99.05
[29]	MIT-BIH	CNN-LSTM + SE blocks	Acc-99.14, Se- 98.67
[30]	MIT-BIH	Transfer learning on spectrograms	Acc-97.9
Ours	**MIT-BIH**	**CNN-LSTM**	**Training:**Acc-99.59,Se-97.15,Sp-99.77,F1-97.06, AUC-99.96,Pre-96.97, PPV-96.96, NPV-99.79 **Testing:**Acc-99.37, Se-95.2, Sp-99.69, F1-95.5, AUC-99.88, Pre-95.8, PPV-95.8, NPV-99.64
	INCART	**CNN-LSTM**	**Training:** Acc-98.60, Se-90.56, Sp-99.45, F1-92.18, AUC-99.69, Pre-93.86, PPV-93.86, NPV-99.13 **Testing:** Acc-98.3, Se-87.48, Sp-99.3, F1-89.74, AUC-99.52, Pre-92.11, PPV-92.11, NPV-98.84

Table 2: CNN-LSTM model performance in detecting premature ventricular contractions on MIT-BIH and INCART databases (own source)

Database	Dataset	Accuracy	Sensitivity	Specificity	F1 score	AUC	Precision	PPV	NPV
MIT	Training	0.9959	0.9715	0.9977	0.9706	0.9996	0.9697	0.9697	0.9979
	Testing	0.9937	0.952	0.9969	0.955	0.9988	0.958	0.958	0.9964
INCART	Training	0.9869	0.9056	0.9945	0.9218	0.9969	0.9386	0.9386	0.9913
	Testing	0.983	0.8748	0.993	0.8974	0.9952	0.9211	0.9211	0.9884

attention mechanisms, suggesting that temporal and spatial feature fusion significantly enhances arrhythmia detection performance.

Despite the strong performance on MIT-BIH, only a few studies have explored the INCART database, which presents a greater challenge due to variations in signal characteristics. Our proposed CNN-LSTM model demonstrates robust performance on both datasets. On MIT-BIH, it achieves a training accuracy of 99.59% and a testing accuracy of 99.37%, along with high sensitivity (97.15% training, 95.2% testing) and AUC values (99.96% training, 99.88% testing), indicating strong classification and generalization capabilities. On INCART, although the performance slightly drops (training accuracy 98.6%, testing accuracy 98.3%), it still remainscompetitive, validating the model's adaptability across databases.

Furthermore, many of the reviewstudies report only accuracy and sensitivity, with limited information on other critical metrics such as F1-score, AUC, precision, and predictive values. In contrast, our model offers a comprehensive evaluation by reporting all relevant metrics, thereby providing a more holistic view of classification performance.

The**Table 2** CNN-LSTM model demonstrates excellent performance in detecting premature ventricular contractions (PVCs) across both the INCART and MIT-BIH ECG databases. On the MIT-BIH dataset, the model achieved a high testing accuracy of 99.37%, with a sensitivity of 95.20% and specificity of 99.69%. These metrics indicate the model's strong ability to correctly detect PVCs while minimizing false positives. The F1 score (0.955), precision (0.958), and area under the ROC curve (AUC = 0.9988) further support its balanced and robust classification performance. Similarly, on the INCART dataset, the model maintained strong performance, achieving a testing accuracy of 98.30%, sensitivity of 87.48%, and specificity of 99.30%. Though slightly lower than MIT-BIH, these results still reflect reliable detection of PVCs with high precision (0.9211), F1 score (0.8974), and AUC (0.9952). The negative predictive values (NPV) of 99.64% for MIT-BIH and 98.84% for INCART highlight the model's excellent capability in ruling out non-PVC beats. Overall, the CNN-LSTM model exhibits strong generalization and robust classification across different ECG datasets, making it a suitable candidate for clinical deployment in real-time arrhythmia monitoring systems.

5. Conclusion

The proposed hybrid CNN-LSTM model exhibits strong and consistent performance in detecting premature ventricular contractions across both MIT-BIH and INCART ECG databases. The model demonstrated high classification performance, achieving testing accuracies of 99.37% on MIT-BIH and 98.30% on INCART. Evaluation metrics showed strong sensitivity (95.20% and 87.48%), specificity (99.69% and 99.30%), and area under the curve (AUC) values (0.9988 and 0.9952). High F1 scores (0.955 and 0.8974) and negative predictive values (99.64% and 98.84%) support its reliability in identifying PVCs while effectively excluding normal beats. The CNN-LSTM architecture leverages both spatial and temporal features, improving arrhythmia detection accuracy. Performance across varied signal morphologies and sampling conditions confirms generalizability, minimizing over fitting and enhancing real-world applicability. With its high precision and low false-positive rates, this model is suitable for continuous cardiac monitoring.

REFERENCES

[1] Alamatsaz, M. H., Chatterjee, S., & Moradi, M. H (2022). Lightweight CNN-LSTM model for real-time arrhythmia classification. Biomedical Signal Processing and Control, 72, 103340. https://doi.org/10.1016/j.bspc.2021.103340

[2] He, Y., Sun, Y., & Qiu, H. (2022). SEVGG-Net-LSTM: A hybrid deep learning model for ECG arrhythmia classification. Journal of Healthcare Engineering, 2022, 9312798. https://doi.org/10.1155/2022/9312798

[3] Jin, L., Sun, W., Zhang, J., & Chen, H. (2020). Multi-scale CNN for arrhythmia classification from ECG signals. Sensors, 20(15), 4370. https://doi.org /10.3390/s20154370

[4] Shah, A., Patel, D., & Gohil, R. (2021). Arrhythmia detection using residual CNN with BiLSTM on MIT-BIH dataset. Biomedical Signal Processing and Control, 68, 102743. https://doi.org/ 10.1016/ j.bspc. 2021.102743

[5] Ullah, I., Hussain, M., &Aboalsamh, H. (2023). Efficient ECG arrhythmia classification using a combination of CNN and BiLSTM. IEEE Access, 11, 385–394. https://doi.org/10.1109/ACCESS.2023.3256052

[6] Xu, T., Wang, Y., & Liu, F. (2024). ECG classification using CNN-GRU networks with multiscale temporal fusion. Computers in Biology and Medicine, 162, 106144. https://doi.org/ 10.1016/ j.compbiomed. 2023.106144

[7] Yao, Q., Fan, X., & Yang, L. (2020). Multi-class ECG arrhythmia detection using deep residual CNN and GRU. Frontiers in Physiology, 11, 561. https://doi.org/10.3389/fphys.2020.00561

[8] Zhang, J., Wang, X., Wang, J., Zhao, Y., & Liu, J. (2020). A hybrid deep learning model for arrhythmia detection based on residual SE-CNN and BiLSTM. Computers in Biology and Medicine, 120, 103748. https://doi.org/10.1016/j.compbiomed.2020.103748

[9] Zhang, Y., Dong, Y., & Wang, C. (2020). ECG signal classification with deep learning for continuous cardiac monitoring. Healthcare Technology Letters, 7(3), 66–71. https://doi.org/10.1049/htl.2020.0022

[10] Zhang, Z., Pi, D., Liu, Y., & Wang, Y. (2021). ECG classification using an attention-based BiLSTM and multi-scale CNN. Biocybernetics and Biomedical Engineering, 41(3), 1042–1055. https://doi.org/ 10.1016 / j.bbe.2021.07.004

[11] Peimankar, A., &Puthusserypady, S. (2021). A novel deep learning model for classification of ECG signals using QRS complex. Neural Networks, 138, 76–86. https://doi.org/10.1016/j.neunet.2021.01.024

[12] Wang, J., Guo, Y., Chen, Y., & Ma, X. (2022). Real-time heartbeat classification using lightweight CNNs and knowledge distillation. IEEE Journal of Biomedical and Health Informatics, 26(6), 2753–2761. https://doi.org/10.1109/ JBHI.2021. 3137230

[13] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89, 389–396. https://doi.org/10.1016/j.compbiomed.2017.08.022

[14] Oh, S. L., Ng, E. Y. K., Tan, R. S., & Acharya, U. R. (2019). Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heartbeats. Computers in Biology and Medicine, 102, 278–287. https://doi.org/10.1016/ j.compbiomed.2018.11.013

[15] Xiong, Z., Stiles, M. K., & Zhao, J. (2020). ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. Physiological Measurement, 41(4), 045005. https://doi.org/10.1088/1361-6579/ab77e8

[16] Gao, S., Dong, H., & Yang, G. (2019). ECG heartbeat classification using attention mechanism and LSTM network. In 2019 IEEE International Conference on Artificial Intelligence and Big Data (ICAIBD). https://doi.org/10.1109/ ICAIBD .2019. 8836965

[17] Warrick, P., & Homsi, M. N. (2018). Cardiologist-level arrhythmia detection using recurrent neural networks. arXiv preprint, arXiv:1812.08434.

[18] Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with deep neural networks. arXiv preprint, arXiv:1707.01836.

[19] Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. Computer Methods and Programs in Biomedicine, 161, 1–13.https://doi.org/10.1016 /j.cmpb .2018.04.005

[20] Luz, E. J. D. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). ECG-based heartbeat classification for arrhythmia detection: A survey. Computer Methods and Programs in Biomedicine, 127, 144–164. https://doi.org/10.1016/j.cmpb.2015.12.008

[21] Yildirim, O. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in Biology and Medicine, 96, 189–202. https://doi.org/10.1016 /j.compbiomed.2018.03.016

[22] Chandra, B. S., Sastry, C. S., & Jana, A. (2018). CNN-based model for real-time arrhythmia detection. Biomedical Signal Processing and Control, 43, 64–70. https://doi.org/10.1016/j.bspc.2018.02.010

[23] Sannino, G., & De Pietro, G. (2018). A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. Future Generation Computer Systems, 86, 446–455. https://doi.org/10.1016 /j.future.2018.03.057

[24] Chauhan, S., &Vig, L. (2015). Anomaly detection in ECG time signals via deep long short-term memory networks. In 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA) (pp. 1–7). https://doi.org/10.1109/DSAA.2015.7344832

[25] Kiranyaz, S., Ince, T., &Gabbouj, M. (2016). Real-time patient-specific ECG classification with recurrent neural networks. IEEE Transactions on Biomedical Engineering, 63(3), 664–675. https://doi.org/10.1109/TBME.2015.2468589

[26] Hassan, S. U., Tufail, M., Samsuzzaman, M., & Khalid, M. (2024). Optimized CNN-LSTM model for PVC detection on MIT-BIH. IEEE Transactions on Instrumentation and Measurement, 73, 1–10.

[27] Mousavi, P., Shahroudy, A., & Nguyen, T. (2021). PVC detection using CNN-seq2seq on MIT-BIH. IEEE Journal of Biomedical and Health Informatics, 25(4), 1305–1312.

[28] Zhang, Q., Li, Z., & Yu, Y. (2024). Explainable CNN-LSTM model for the classification of multiple arrhythmias. Pattern Recognition Letters, 163, 41–49.

[29] Sun, A., Hong, W., Li, J., & Mao, J. (2024). CNN-LSTM with squeeze-and-excitation for arrhythmia classification on MIT-BIH. IEEE Access, 12, 12345–12356.

[30] Malleswari, P. N., Odugu, V. K., Rao, T. J. V. S., & Aswini, T. V. N. L. (2024). Transfer learning for ECG arrhythmia detection using 2-D spectrograms. EURASIP Journal on Advances in Signal Processing, 2024, 23.

[31] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine, 20(3), 45-50. https://doi.org/10.1109/51.932724

[32] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215-e220. https://doi.org/10.1161/01.CIR.101.23.e215

[33] Silva, I., Behar, J., Sameni, R., Zhu, T., Oster, J., Clifford, G. D., & Moody, G. B. (2022). WFDB: The WaveForm Database library for Python. Journal of Open Source Software, 7(76), 4074. https://doi.org/10.21105/joss.04074

[34] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Transactions on Biomedi-

cal Engineering, 32(3), 230-236. https://doi.org/10.1109/TBME.1985.325532

[35] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261-272. https://doi.org/10.1038/s41592-019-0686-2

[36] Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection with deep neural networks. Nature Medicine, 25(1), 65-69. https://doi.org/10.1038/s41591-018-0268-3

[37] Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., ... & Ribeiro, A. L. P. (2020). Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature Communications, 11(1), 1-9. https://doi.org/10.1038/s41467-020-15432-4