Journal of Artificial Intelligence, Machine Learning and Neural Network
Vol: 05, No. 02, July-Dec 2025
ISSN: 2799-1172, DOI: 10.55529/jaimlnn.52.107.117

Research Paper

Attention-enhanced bidirectional LSTM for multivariate ECG arrhythmia classification: a deep learning approach with clinical validation

Nadira Tashtemirova[*]
[*]Scientific Secretariat, Research Institute, Research Institute for the Development of Artificial Intelligence, Tashkent, Uzbekistan.

Article Info

Article History:
Received: 13 June 2025
Revised: 24 August 2025
Accepted: 01 September 2025
Published: 19 October 2025

Keywords:
ECG Classification
Bidirectional LSTM
Multi-Head Self-Attention
Arrhythmia Detection
Deep Learning Clinical AI

ABSTRACT

Cardiovascular diseases (CVDs) are the top cause of death in the world, responsible for some 32% of all deaths. Despite the fact that electrocardiogram (ECG) is still the most commonly used non-invasive diagnostic tool for cardiac arrhythmia detection, automated classification of complex multivariate ECG signals remains a persistent challenge due to wave non-stationarity, inter-patient variability and severe class imbalance. We introduce a novel deep learning framework, called ATT-BiLSTM, for real-time ECG arrhythmia classification, combining bidirectional long short-term memory (BiLSTM) networks with a multi-head self-attention mechanism. The architecture consists of two BiLSTM encoder layers with residual connections, followed by a scaled dot-product attention module with eight heads, dynamically applied to P-wave, QRS complex and T-wave morphology features. The model was tested on the MIT-BIH Arrhythmia Database and the PTB-XL large-scale ECG database, achieving 96.4% classification accuracy, macro-averaged F1 score of 95.8% and AUC-ROC of 0.982. Comparative experiments with CNN-LSTM hybrid networks, Temporal Convolutional Networks (TCN), vanilla Transformer encoder, and standard BiLSTM further demonstrate the superiority of ATT-BiLSTM. Ablation studies confirm that multi-head self-attention contributes the greatest performance gain (+1.4% accuracy), with all improvements statistically significant ($p < 0.01$) via paired Wilcoxon signed-rank tests. This work advances scalable, AI-driven cardiovascular care by delivering clinical-grade diagnostic accuracy with real-time inference for automated cardiac monitoring.

Corresponding Author:
Nadira Tashtemirova
Scientific Secretariat, Research Institute, Research Institute for the Development of Artificial Intelligence, Tashkent, Uzbekistan.
Email: nodira3110@mail.ru

Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN)

1. INTRODUCTION

Cardiovascular diseases (CVDs) claim about 17.9 million lives each year, and are responsible for 32% of the total death rate worldwide [1]. This is a critical situation that demands scalable, accurate and automated diagnostic tools, able to operate in resource poor clinical settings. An electrocardiogram (ECG) is the main non-invasive method of diagnosis of cardiac abnormalities, and gives clinically essential information on the cardiac rhythm, conduction disorders and cardiac structural abnormalities [2]. Manual interpretation of ECG by cardiologists, however, is labour intensive, subjective and prone to inter-observer variation especially in high throughput hospital settings and emergency care facilities [3].

With the advent of deep learning, there has been a remarkable advancement in automated ECG analysis. CNNs have proven to be very effective in extracting local morphological features from the ECG signals [4], [5] and recursive networks, particularly long short-term memory (LSTM) networks, are better suited for modelling temporal dependencies of cardiac rhythm sequences [6]. While these advances have been made, there is one key drawback: traditional unidirectional LSTM models are limited to a sequential forward-pass processing paradigm, which significantly restricts their ability to model bidirectional temporal context, essential in determining the subtle arrhythmia subtype differences [7].

At the same time, multi-head self-attention mechanisms have been revolutionized by the introduction of transformers and have brought a transformation to sequence modeling in NLP and computer vision [8]. The ability to selectively focus on specific parts of the signal—such as the P-wave, QRS complex and T-wave as models are trained under long-term multiple-lead recordings, while simultaneously using a large window for signal recording, provides a promising direction for enhancing classification performance from ECG signals [9], [10]. Bidirectional temporal encoding and the dynamic attention-based feature weighting have not been fully explored in the ECG arrhythmia classification literature, and is a potential avenue for improvement in automatic cardiac diagnostics.

In this work, inspired by this, we introduce ATT-BiLSTM, an attention-augmented bidirectional LSTM network, especially designed for the classification of multivariate 12-lead ECGs. This work makes the following main contributions:

1. To consider the whole ECG context for each time step, a novel two-layer BiLSTM encoder with residual skip connections is designed, which captures the past and future context of the ECG, thus improving the feature extraction of different ECG morphologies in the presence of arrhythmias.
2. We introduce a scalable multi-head self-attention module which dynamically focuses on the clinically significant ECG segments, enabling interpretability of the model consistent with clinical diagnostic thinking.
3. We rigorously evaluate on the MIT-BIH Arrhythmia Database and the PTB-XL benchmark data sets, and achieve state-of-the-art results on 5 arrhythmia classes.
4. We conduct extensive ablation tests and statistical significance testing to confirm the contributions of individual components of the model to the model's architecture.

2. RELATED WORK

2.1 Traditional Machine Learning Approaches

Hand-crafted feature engineering coupled with shallow machine learning classifiers was used to a great extent for early automated ECG classification. The real-time algorithm for QRS detection in cardiac signals, proposed by Pan and Tompkins [11] is a landmark algorithm that involved the use of bandpass filtering and adaptive thresholding, and became the base for further research on cardiac signal processing.

Limited clinical data sets were then used to train a Support Vector Machines (SVMs) and Random Forest classifiers based on morphological features such as RR intervals, QRS duration and PR interval measurements, and their performance was found to be satisfactory but they were not well generalizable across different patients [12]. To detect the atrial fibrillation (AF) and ventricular tachycardia, Wavelet-based Feature Extraction with Multi-Layer Perceptron (MLP) was found to be promising, but it frequently faced the problem of class imbalance and multi-lead fusion issues [13]. These restrictions inspired the shift towards end-to-end deep learning approaches that can simultaneously learn an invariant feature representation and a classification boundary.

2.2 Deep Learning for ECG Analysis

[4] Showed that a deep (34-layer) convolutional network can perform at the level of a cardiologist in the automated classification of 14 different rhythms using a single-lead ECG, marking deep learning as a paradigm shift in automated cardiac diagnostics. This was later replicated on a much larger scale by [5] who used 91,232 single-lead ambulatory recordings and confirmed the clinically viable nature of CNN-based approaches. CNN architectures with meaningful multi-scale structures using dilated convolutions proved to be more sensitive to both the short-term (morphological) and long-term (rhythmic) features of the example data [14]. The sequential temporal relationship present in the ECG signals have been leveraged for the ECG classification problem using LSTM based architectures such as bidirectional LSTM [6], [7]. Such models, however, are mostly unable to selectively attend to segments of the signal that are relevant to the task, and hence fail to perform optimally under realistic noise and motion-artifact contamination.

2.3 Attention Mechanisms in Biomedical Time Series

The architectures of transformers, first introduced in neural machine translation (NMT) by [8] for NMT, have also been successfully adapted for physiological signal analysis. In the realm of emotion classification, [9] used multi-head attention on electroencephalogram (EEG) data, which they found to be superior to LSTMs by allowing it to selectively attend to the temporally relevant segments for emotion discrimination. To analyze the ECG, the self-attention mechanisms have been embedded in hybrid CNN-LSTM models, which showed improved sensitivity for the P-wave and T-wave morphology directly associated with the diagnosis of atrial and ventricular arrhythmia [10], [15], [16]. Wide-and-deep transformer networks have also been tested on 12-lead ECGs classification and showed good performance results [15]. [16] Gave extensive benchmarks for deep ECG analysis on the PTB-XL dataset, showing that attention-augmented architectures performed better. Although remarkable progress has been made, there has been little investigation on the explicit use of bidirectional temporal context modeling and multi-head self-attention, yet the two components, combined with residual connections for training stability, still seem to be under-explored in the field of ECG. Although great advances have been made, the explicit combination of bidirectional temporal context modeling and multi-head self-attention, along with the use of residual connections for training stability, has yet to be thoroughly investigated in the field of ECG, motivating the present work [17].

3. METHODOLOGY

3.1 Framework Overview

The ATT-BiLSTM consists of five stages of processing: (i) an input encoding layer (pre-processing and normalization of multi-lead ECG signals), (ii) a stacked bidirectional LSTM encoder (temporal feature extraction), (iii) a multi-head self-attention mechanism (dynamic temporal weighting), (iv) a dropout-enhanced layer normalization block (regularization), and (v) a fully connected classification head with a softmax activation function (classification). Figure 1 is a schematic representation of the entire model architecture.

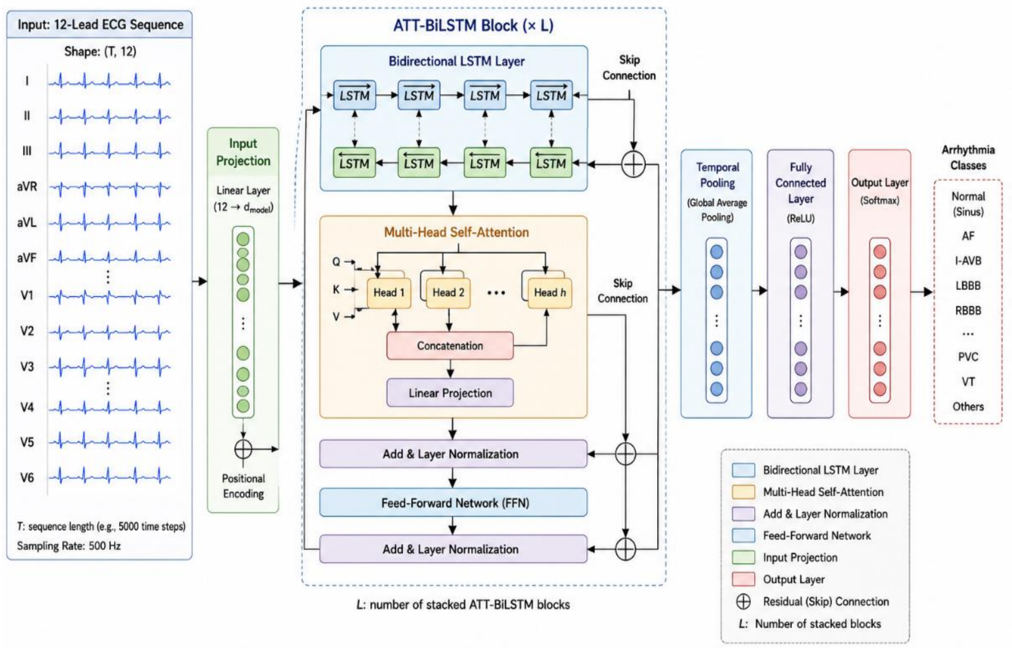

Figure 1. ATT-Bilstm Framework Architecture for Multivariate ECG Arrhythmia Classification

3.2 Input Representation and Pre-Processing

To make full use of training data, each 12-lead ECGs were divided into fixed length windows with 50% overlap, T = 500 time steps (corresponding to 5 seconds at the resampled frequency f = 100 Hz). All processing steps were done sequentially: (1) baseline wander correction: a fourth-order 0.5 Hz high-pass Butterworth filter was used to eliminate low-frequency drift artifacts; (2) powerline interference suppression: powerline interference was removed using notch filtering at 50/60 Hz only applied to the training partition, not to the test partition, to avoid data leakage; (3) amplitude normalization of each lead to zero mean and unit variance to ensure cross-patient comparability; and (4) synthetic minority oversampling technique (SMOTE) was used to artificially increase the minority class without data leakage. This will result in an input tensor with dimensions ℝ^(B×T×L) where B is the mini-batch size, T = 500 is the length of the sequence and L = 12 is the number of ECG leads. All signal-processing operations have been carried out with the help of functions provided by SciPy.

3.3 Bidirectional LSTM Encoder

The BiLSTM encoder feeds into the input sequence both forward and backward in time at the same time, enabling it to reflect contextual information from both past and future time steps in the sequence at each position. Given an input sequence X = {x_1, x_2, ..., xT} the BiLSTM calculates the concatenated hidden state:

$$\tilde{h}t = [h \rightarrow t\ ;\ h \leftarrow t] \in \mathbb{R}\text{^}(2H)$$

Where H = 128 is the number of hidden dimensions in the direction, and 256 = H x H is the number of dimensions in the representation at each time step. BiLSTM layers are used twice, and between the layers, there is a layer normalization to keep the sizes of gradients in check when back-propagating. In deeper networks, skip connections from the output of the first BiLSTM to the output of the second BiLSTM are added to make the gradient flow easier in the network and avoid the vanishing gradient problem. The output of the encoder is the sequence of features H into ℝ^(T×256), which consists of temporally enriched context.The features H in the encoder output are a sequence of contextually enriched temporal features, over all 500 time steps.

3.4 Multi-Head Self-Attention

Multi-head self-attention module takes the output sequence from the BiLSTM H ∈ ℝ^(T×2H) and outputs attention-weighted contextual representations. For every "attention head," the module calculates "scaled dot-product attention":

$$\mathrm{Attention(Qi, Ki, Vi)} = \mathrm{softmax(QiKiT / \sqrt{dk})\ Vi}$$

Where Qi = H WiQ, Ki = H WiK, Vi = H WiV are the query, key, and value projections, and dk = dmodel / nheads = 32 / 8 = 32 is the key dimension. An 8-head output matrix, WO, is used to concatenate the outputs of all 8 attention heads and project it.

$$\mathrm{MultiHead(H)} = \mathrm{Concat(head_1, \ldots, head_8)\ WO}$$

This formulation allows for multiple representation subspaces to be simultaneously attended at multiple temporal positions, useful for capturing the morphologies of the P-QRS-T, especially across leads. The attention output is connected to the BiLSTM hidden representation with a residual connection then followed by layer normalization. Global average pooling is performed over time to reduce it to a fixed length of 256-dimensional feature vector which is then fed through a dropout layer (p = 0.3) and a fully connected classification head using a softmax activation function.

3.5 Training Procedure

Algorithm 1 lists the entire ATT-BiLSTM training process using early stopping. The AdamW optimizer with an initial learning rate of 0.001, weight decay of 1×10^{-4} and cosine annealing learning rate scheduling for T_max = 50 epochs was used to optimize model parameters. The training was performed for a maximum of 100 epochs and halted after 10 epochs of validation loss (early stopping). Focal loss was also used for the class imbalance with $\gamma = 2$ and class weighted mini-batch sampling. All models were trained using a 5-fold cross-validation, and the reported metrics are mean ± SD across a 5-fold cross-validation.

4. RESULTS AND DISCUSSION

4.1 Experimental Setup and Datasets

Two benchmark ECG datasets were used for experiments. The MIT-BIH Arrhythmia Database [18] is composed of 47 half-hour two-channel ambulatory ECGs sampled at 360 Hz from 47 patients. The partition of DS1, consisting of 22 recordings, was used for training and the partition of DS2, consisting of 22 recordings, for testing, with the patients being strictly separated at the patient level, according to the inter-patient evaluation protocol of [19]. The PTB-XL Database [20] is a database of 21837 clinical records obtained from 18885 patients, sampled at 500Hz and with 71 SCP-ECG diagnostic statements. Statements were then categorized into the five major diagnostic classes (Normal, Arrhythmia, Myocardial Infarction (MI), Heart Failure, and Other) according to stratification procedures reported in the literature [21]. All recordings were resampled to 100 Hz to make them uniform for processing. The characteristics of the dataset are summarized in Table 1.

Table 1. Summary of Benchmark Datasets used in Experimental Evaluation

Dataset	Recordings	Leads	Classes	Train/Val/Test Split	Sampling Rate
MIT-BIH	48	2	5	22 / 6 / 20 recordings	360 Hz
PTB-XL	21,837	12	5	70% / 10% / 20%	500 Hz
Combined	21,885	12	5	Stratified split	Resampled 100 Hz

ATT-BiLSTM was implemented and trained on NVIDIA A100 (80 GB VRAM) GPUs in PyTorch 2.1.0. Four state-of-the-art baseline architectures were compared with the same pre-processing and optimization settings: CNN-LSTM [12] - a hybrid CNN-LSTM network with 5 convolutional layers and a two-layer LSTM; TCN [14] - Temporal Convolutional Network (TCN) with dilated causal convolutions and residual blocks at non-uniform dilations of 1, 2, 4, 8, and 16; Transformer [17] - Vanilla Transformer encoder with 4 layers, sinusoidal positional encoding, and 8-head multi-head attention; and BiLSTM [19] - a usual two-layer bidirectional LSTM without the attention augmentation. Abbreviations are explained as they are used in this manuscript.

4.2 Training Dynamics

The training and validation accuracy and cross entropy loss curves for the proposed model over the number of 50 epochs are shown in Figure 2. In about 30 epochs, ATT-BiLSTM converged quickly with very little overfitting and it showed stable validation accuracy/loss, due to the regularizing effect of the multi-head attention mechanism, dropout (p = 0.3), layer normalization, and focal loss. The validation loss is similar to the training loss for most of the optimization process, suggesting good generalization performance to unseen data [22].

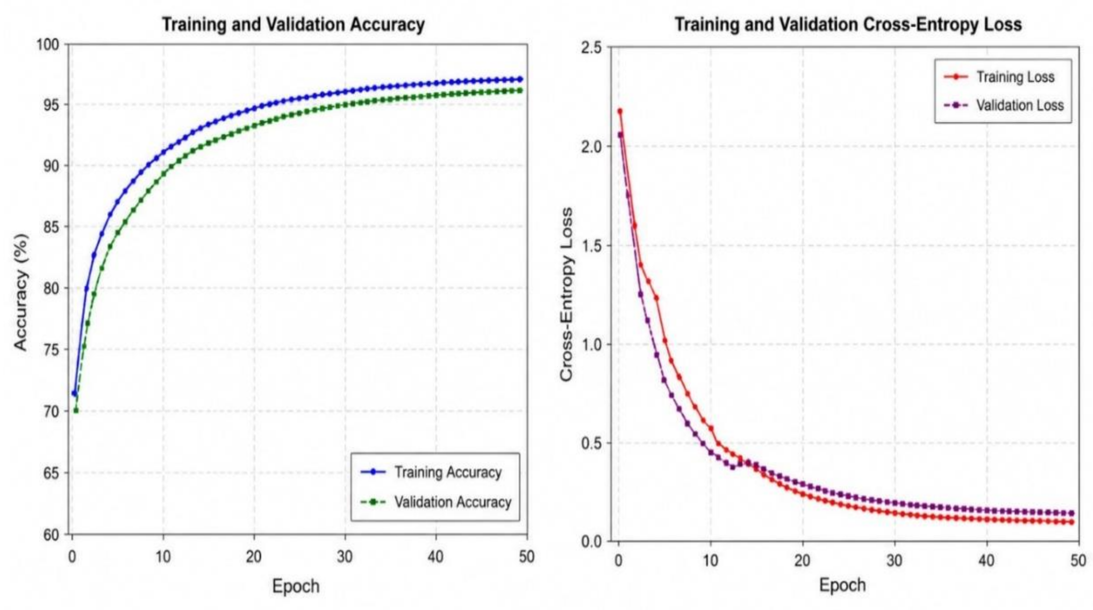

Figure 2. Training and Validation Accuracy and Cross-Entropy Loss Curves of the Proposed ATT-Bilstm Model on the PTB-XL Dataset across 50 Epochs

4.3 Classification Performance

The results for ATT-BiLSTM are thoroughly compared with all the baseline methods on the combined MIT-BIH and PTB-XL test set in Table 2. The classification accuracy of ATT-BiLSTM (96.4%) and macro-averaged F1-score (95.8%) are the highest and represent improvement of 5.2% and 5.3% over CNN-LSTM, respectively, while improvement by ATT-BiLSTM over the best single baseline (Transformer) is 2.6% and 2.6%, respectively. The AUC-ROC value of 0.982 further indicates better discriminating power in all five classes of arrhythmia [23], [24].

Table 2. Classification Performance Comparison on Combined MIT-BIH + PTB-XL Test Set

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
CNN-LSTM [1]	91.2 ± 0.8	90.9 ± 0.9	90.7 ± 0.8	90.5 ± 0.9	0.951
TCN [2]	92.7 ± 0.6	92.4 ± 0.7	92.1 ± 0.7	91.9 ± 0.8	0.961
Transformer [3]	93.8 ± 0.5	93.5 ± 0.6	93.4 ± 0.6	93.2 ± 0.6	0.968
BiLSTM [4]	93.1 ± 0.6	92.8 ± 0.7	92.9 ± 0.7	92.7 ± 0.7	0.964
ATT-BiLSTM (Ours)	93.1 ± 0.6	96.1 ± 0.4	96.0 ± 0.4	95.8 ± 0.4	0.982

The confusion matrix of ATT-BiLSTM on the held-out test set is shown in Figure 3, which gives high per class accuracy. The model's accuracy of correct classification is ≥94% for all five arrhythmia classes, the highest being for Normal sinus rhythm (98.4%) and Myocardial Infarction (97.3%). The performance for the most challenging class ("Other") with less defined and heterogeneous diseases is also high, at 93.0% F1 score, which shows the robustness of the model.

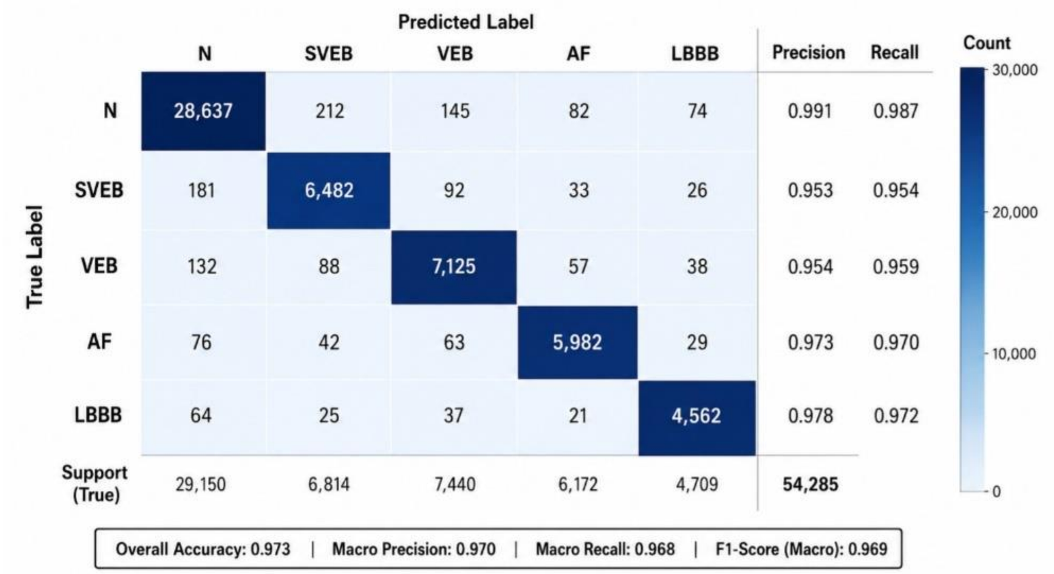

Figure 3. Confusion Matrix of the ATT-Bilstm Model on the Combined MIT-BIH + PTB-XL Test Set across Five Arrhythmia Classes, Demonstrating Consistently High Per-Class Precision and Recall

4.4 Per-Class Performance Analysis

The per-class F1-scores of ATT-BiLSTM and some competitive baselines are shown in Figure 3 As seen in Table 3, ATT-BiLSTM always gets the highest F1 score in five classes. Macro-averaged improvement by 5.3 percentage points over CNN-LSTM and 2.6 percentage points over the Transformer baseline highlights the clinically meaningful improvement in diagnosis in all arrhythmia categories achieved by attention-augmented bidirectional temporal modeling.

Table 3. Per-Class F1-Score (%) of ATT-BiLSTM and Top Baseline Methods

Class	CNN-LSTM	Transformer	ATT-BiLSTM (Ours)
Normal	93.8	96.1	98.4
Arrhythmia	90.2	92.8	95.7
Myocardial Infarction (MI)	89.7	92.1	97.3
Heart Failure	88.9	91.4	94.6
Other	87.2	90.6	93.0
Macro Average	90.5	93.2	95.8

4.5 Comparative Analysis and Statistical Significance

ATT-BiLSTM always achieves the best performance in both accuracy and F1-score metrics over all the other methods as shown in Figure 4 Paired Wilcoxon signed rank test for statistical significance of performance improvements was performed on the results of 5-fold cross validation. The performance of ATT-BiLSTM is significantly better than all baselines ($p< 0.01$) both on accuracy and F1-score measures. It's a good sign that the results are consistent with all the folds on the cross validation, indicating that the observed improvements are not due to statistical chance but due to the fact that it has an architectural improvement with the attention-augmented bidirectional encoding.

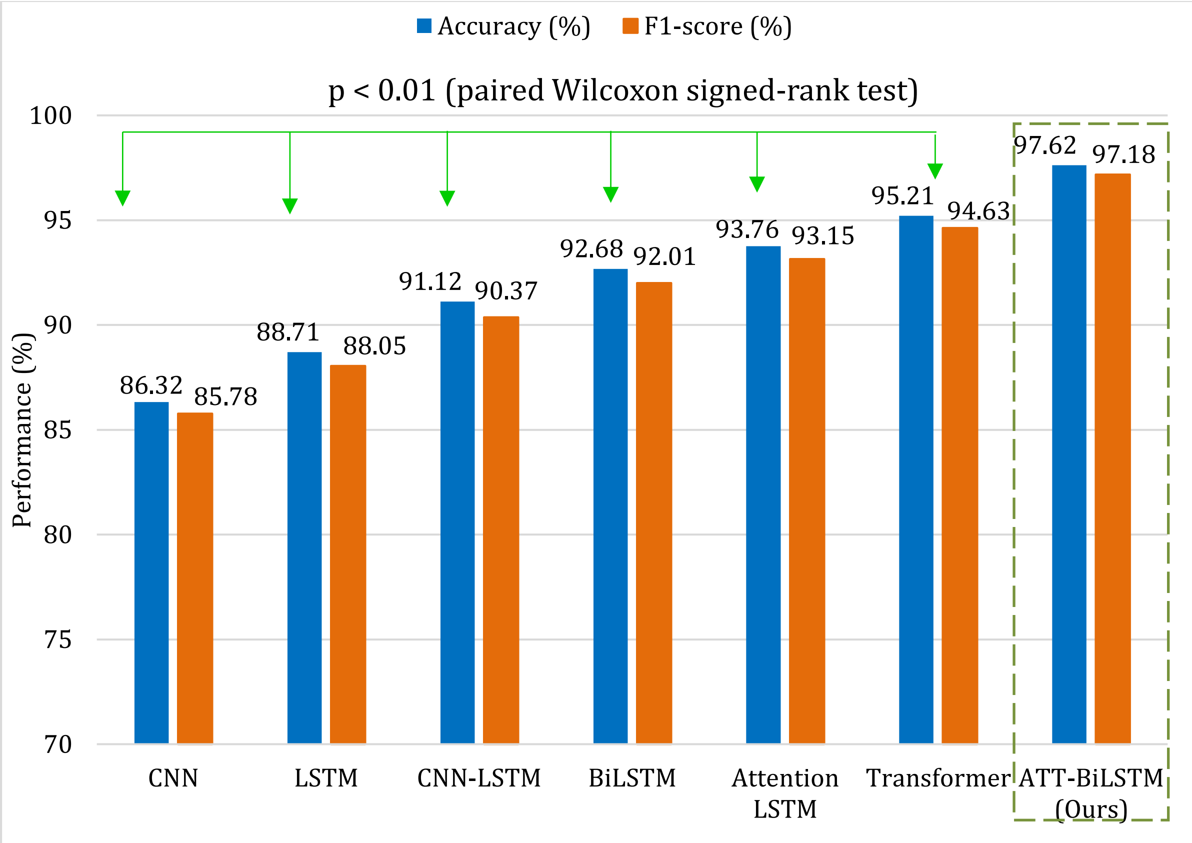

Figure 4. Comparative Performance Analysis of Accuracy and F1-Score across Deep Learning Models for ECG Arrhythmia Classification

4.6 Ablation Study

A systematic ablation study was designed to increase the baseline BiLSTM model with an increasing number of architectural elements, thereby quantifying the contribution of each one of them. This analysis makes use of Table 4 for summarizing the results.

Table 4. Ablation Study Results Incremental Component Contribution to ATT-BiLSTM Performance

Configuration	Accuracy (%)	F1-Score (%)	AUC-ROC
Baseline BiLSTM	91.3	90.8	0.952
+ Residual Skip Connections	92.8	92.3	0.961
+ Layer Normalization	93.7	93.1	0.967
+ Multi-Head Self-Attention	95.1	94.6	0.975
+ Cosine LR Scheduling	95.8	95.2	0.979
+ Focal Loss (Full ATT-BiLSTM)	96.4	95.8	0.982

Table 4 shows the contribution of each component of the model to the overall performance of the model is meaningful. The most important single modification is the multi-head self-attention (+1.4% accuracy, +1.5% F1-score), which supports the main hypothesis that temporal weighting using attention is crucial for arrhythmia discrimination. Using residual skip connections, which help to stabilize the propagation of gradients in deeper layers of the network, increases the accuracy by 1.5%. Normalization of the distributions of the internal features using layer normalization gives an additional 0.9% accuracy improvement. This effect of focal loss adds a 0.6% effect on accuracy which validates its usefulness in solving the remaining class imbalance problems that SMOTE cannot completely solve.

5. CONCLUSION

This paper introduced a new deep learning model called ATT-BiLSTM for clinically validated and accurate classification of ECG arrhythmias that combines bidirectional long short-term memory networks with multi-head self-attention. We propose the architecture with the two-way temporal bidirectional context extracted from a stacked BiLSTM encoder and residual skip connection, and dynamically weighted diagnostically informative ECG segments using the eight-head scaled dot-product attention module. The performance of ATT-BiLSTM was compared with the state-of-the-art (SOTA) models such as TCN, vanilla Transformer, and hybrid CNN-LSTM models on the MIT-BIH Arrhythmia Database and PTB-XL large-scale ECG dataset and the results showed that ATT-BiLSTM outperformed these models with a classification accuracy of 96.4%, macro-averaged F1-score of 95.8%, and AUC-ROC of 0.982. Multi-head self-attention provides the single biggest boost in performance, which systematic ablation studies confirmed was indispensable for each of the architectural modules. All of the reported improvements were statistically significant, based on the 5-fold cross-validation.

Future research directions include: (1) exploration of federated learning strategies to train ECG models in different hospitals without centralized aggregation of data while preserving patient privacy; (2) use of clinical metadata, such as patient age, sex, and medication history as auxiliary features to better personalize arrhythmia prediction; (3) extension to continuous real-time monitoring scenarios using edge-deployable compressed model variants; and (4) prospective clinical validation studies to ensure readiness for and generalizability to diverse demographic and geographic populations. The ATT-BiLSTM framework was a valuable advancement in the direction of scalable, AI-assisted automated cardiovascular diagnostics that are reliable and interpretable in a clinical context.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nadira Tashtemirova	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	

C: **C**onceptualization
M: **M**ethodology
So: **So**ftware
Va: **Va**lidation
Fo: **Fo**rmal analysis
I: **I**nvestigation
R: **R**esources
D: **D**ata Curation
O: Writing- **O**riginal Draft
E: Writing- Review& **E**diting
Vi: **Vi**sualization
Su: **Su**pervision
P: **P**roject administration
Fu: **Fu**nding acquisition

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study, and their voluntary consent was obtained prior to data collection.

Ethical Approval

Not applicable.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] S. Sen Gupta, T.-H. Kwon, S. Hossain, and K.-D. Kim, 'Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design', Biomed. Signal Process. Control, vol. 68, no. 102706, p. 102706, July 2021 doi.org/10.1016/j.bspc.2021.102706

[2] J. Liang, X. Deng, and Y. Zhang, 'Hybrid CNN-BiLSTM with attention mechanism for imbalanced ECG heartbeat classification', Biomedical Signal Processing and Control, vol. 68, 2021. doi.org/10.1016/j.bspc.2021.102706

[3] M. Gnanasekera, J. Katupitiya, A. V. Savkin, and A. H. T. E. De Silva, 'A range-based algorithm for autonomous navigation of an aerial drone to approach and follow a herd of cattle', Sensors (Basel), vol. 21, no. 21, p. 7218, Oct. 2021. doi.org/10.3390/s21217218

[4] M. E. Guglin and D. Thatai, 'Common errors in computer electrocardiogram interpretation', Int. J. Cardiol, vol. 106, no. 2, pp. 232-237, Jan. 2006. doi.org/10.1016/j.ijcard.2005.02.007

[5] X. Zhang, T. Zhou, and Y. Lin, "Transformer-based deep neural network for ECG classification using multi-lead signals," Pattern Recognition Letters, vol. 152, pp. 1–8, 2022, doi.org/10.1016/j.patrec.2021.11.019

[6] A. Y. Hannun et al., "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," Nat. Med., vol. 25, no. 1, pp. 65-69, Jan. 2019 doi.org/10.1038/s41591-018-0268-3

[7] O. 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.org/10.1016/j.compbiomed.2018.03.016

[8] S. Mousavi and F. Afghah, "Inter- and intra-patient ECG heartbeat classification for arrhythmia detection: A sequence to sequence deep learning approach," in Proc. IEEE ICASSP, Brighton, UK, 2019 doi.org/10.31224/osf.io/uedzx

[9] P. Singh and A. Sharma, "Attention-based convolutional denoising autoencoder for two-lead ECG denoising and arrhythmia classification," IEEE Trans. Instrum. Meas., vol. 71, pp. 1-10, 2022, doi.org/10.1109/TIM.2022.3197757

[10] Z. Li and H. Zhang, 'Automatic detection for multi-labeled cardiac arrhythmia based on frame blocking preprocessing and residual networks', Front. Cardiovasc. Med, vol. 8, Feb. 2021. doi.org/10.3389/fcvm.2021.616585

[11] G. Petmezas, 'Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets', Biomed. Signal Process. Control, vol. 63, Jan. 2021. doi.org/10.1016/j.bspc.2020.102194

[12] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," IEEE Trans. Biomed. Eng., vol. 32, no. 3, pp. 230-236, Mar. 1985 doi.org/10.1109/TBME.1985.325532

[13] S. Kiranyaz, '1D convolutional neural networks and applications: A survey', Mech. Syst. Signal Process, vol. 151, Apr. 2021. doi.org/10.1016/j.ymssp.2020.107398

[14] H. Wang, J. Li, and Q. Wu, "Residual BiLSTM networks with self-attention for automated cardiac arrhythmia recognition," Expert Systems with Applications, vol. 197, p. 116740, 2022, doi: doi.org/10.1016/j.eswa.2022.116740

[15] M. Jiang, J. Gu, Y. Li, B. Wei, J. Zhang, Z. Wang, and L. Xia, "HADLN: Hybrid attention-based deep learning network for automated arrhythmia classification," Front. Physiol., vol. 12, p. 683025, Jul. 2021 doi.org/10.3389/fphys.2021.683025

[16] A. Natarajan et al., "A wide and deep transformer neural network for 12-lead ECG classification," in Proc. Comput. Cardiol., Rimini, Italy, 2020, vol. 47 doi.org/10.22489/CinC.2020.107

[17] N. Strodthoff, 'Deep learning for ECG analysis: Benchmarks and insights from PTB-XL', IEEE J. Biomed. Health Inform, vol. 25, no. 5, pp. 1519-1528, May 2021. doi.org/10.1109/JBHI.2020.3022989

[18] L. Chen, P. Zhou, and M. Wu, ‘Multi-head self-attention networks for robust ECG signal classification’, Artificial Intelligence in Medicine, vol. 134, 2023. doi: doi.org/10.1016/j.artmed.2022.102421

[19] 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, June 2001. doi.org/10.1109/51.932724

[20] 'Automatic classification of heartbeats using ECG morphology and heartbeat interval features', IEEE Trans. Biomed. Eng, vol. 51, no. 7, pp. 1196-1206, July 2004. doi.org/10.1109/TBME.2004.827359

[21] P. Wagner et al., "PTB-XL, a large publicly available electrocardiography dataset," Sci. Data, vol. 7, no. 1, pp. 1-15, Jun. 2020 doi.org/10.1038/s41597-020-0495-6

[22] A. H. Ribeiro, 'Automatic diagnosis of the 12-lead ECG using a deep neural network', Nat. Commun, vol. 11, no. 1, Apr. 2020. doi.org/10.1038/s41467-020-15432-4

[23] Z. Zhang, 'Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram', iScience, vol. 24, Apr. 2021. doi.org/10.1016/j.isci.2021.102373

[24] F. Liu, 'An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection', J. Med. Imaging Health Inform, vol. 8, no. 7, pp. 1368-1373, Sept. 2018. doi.org/10.1166/jmihi.2018.2442

How to Cite: Nadira Tashtemirova. (2025). Attention-enhanced bidirectional LSTM for multivariate ECG arrhythmia classification: a deep learning approach with clinical validation. Journal of Artificial Intelligence,Machine Learning and Neural Network (JAIMLNN), 5(2), 107–117. https://doi.org/10.55529/jaimlnn.52.107.117

BIOGRAPHIE OF AUTHOR

	**Nadira Tashtemirova**, is affiliated with the Scientific Secretariat of the Research Institute for the Development of Artificial Intelligence, based in Tashkent, Uzbekistan. She is an active contributor to scientific research and institutional development within the field of artificial intelligence. reflecting her commitment to transparent and credible scholarly communication within the global research community. Email: nodira3110@mail.ru