Review on hybrid deep learning-based cardiovascular prediction systems

Neeraj Varshney[1], *Pawan Kumar Verma*[2]
[1,2] Lincoln University College, Petaling Jaya, Malaysia ;
pdf.neeraj@lincoln.edu.my[1], abes.pawan@gmail.com[2]

Abstract: Cardiovascular diseases CVDs account for nearly one-third of the deaths of people around the world. For this reason, developing advanced diagnostic systems for the smart detection and prediction of the risk of onset of CVDs is paramount. The development of Artificial Intelligence (AI), Machine Learning (ML), and more recently, Deep Learning (DL), has enabled the automated analysis of ECG signals and other clinical data, which has the potential to radically change the diagnosis of CVDs. In particular, convolutional neural networks (CNN), recurrent neural networks (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Transformers and attention-based models, have all shown great promise in the classification of cardiac arrhythmias and the assessment of CVDs risk. Moreover, hybrid deep learning models, which consist of many of the above architectures, have been shown to have superior performance in the retrieval of spatial and temporal contextual features from large and complex biomedical datasets. This work focuses on review of hybrid deep learning-based systems for the prediction of CVDs. We also summarize the challenges that need to be addressed for the development of prediction systems of CVDs that are robust, explainable, and safe for use.

Keywords: Arrhythmia Detection; CVD; ECG; Deep Learning; CNN; LSTM; Transformer; Explainable;

Introduction

CVD is one of the most serious global health issues. As CVD account for millions of deaths on a yearly basis, the importance of early diagnosis and accurate detection cannot be overstated if mortality is to be decreased and outcomes for patients improved. In traditional diagnostic methods for cardiovascular disease, the diagnosis is mainly based on the highly skilled interpretation of the ECG recording, various laboratory tests, imaging techniques, and clinical judgment. These methods are inherently subjective and can be variable and difficult to scale.

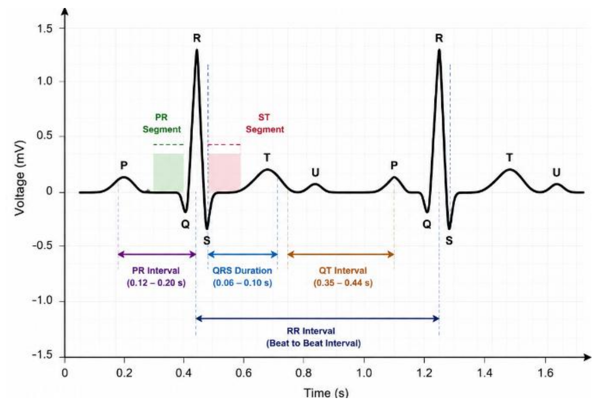

Fig.1: ECG Signal representation

ECGs are important clinical diagnostic tools that show a patient's cardiac activity throughout time. It graphs electrical impulses as waves. Wave interpretation may reveal impulse conduction bouts and cardiac contraction and recovery phases. ECG waveforms, intervals, and segments may be analysed to measure heart rhythm and diagnose cardiac diseases including heart failure. Artificial Intelligence and Machine Learning in ECG analysis have opened new avenues for early detection and clinical decision-making tools.

AI and DL have combined to create cutting-edge diagnostic systems that automatically analyse vast biological datasets and find hidden patterns. Due to DL's adaptability in signal processing, ECGs, and premature ventricular arrhythmia identification, clinical decision-making systems have improved rapidly. CNNs, LSTMs, Transformers, hybrid DL approaches, attentiveness, and explainable AI have improved classical methods.

ECG Signals and Arrhythmia Classification

ECGs are used to monitor cardiac activity. These signals are crucial because they reveal a heart's rhythm, electrical conduction routes, and circulatory system problems. The various temporal and structural aspects of ECG waveforms make them hard to define. DL is used to classify arrhythmia automatically. DL frameworks enable relevant characteristics to be selected and left to the model, unburdened by them, allowing even more sophisticated approaches to be employed on raw ECG waveforms. CNN models excel in identifying and classifying local and regional morphological traits. The collected ECG sequences deviated from simple recurrent patterns and are best fitted in latent space. More advanced deep learning and AI models have obtained greater performance and metrics for detecting complicated arrhythmias. These include atrial fibrillation, ventricular tachycardia, premature ventricular contractions, and bundle branch block classifications.

Related work

Studies show hybrid deep learning algorithms have improved detection of cardiovascular issues and arrhythmias from ECG signals. Recently, many studies have combined CNNs and LSTMs with Transformers, attention, and multi-scale feature extraction. The diagnostic accuracy and robustness of these newer models has improved as a result. Cardioformer, ECGBert, HCTG-Net, and several CNN-LSTM models have been validated on benchmark datasets like MIT-BIH, PTB-XL, MIMIC-IV, and datasets on PhysioNet, with many studies claiming classification rates higher than 99%. These studies provide evidence that automated prediction of cardiovascular disease and interpretation of ECGs greatly improve when local feature extraction is combined with temporal and contextual learning [4,8,9,11,13].

Another trend emerging from the literature is a focus on explainable and lightweight AI. Explainable AI (XAI) like SHAP, Grad-CAM, Lead-wise Grad-CAM, and multimodal visual frameworks have been used to enhance model prediction understanding and build clinical trust. LightX3ECG and ArrhythmiNet are examples of lightweight AI where architecture has been simplified, and have shown that highly performing classifiers are achievable with low computational expense, meaning these models can run on wearable and limited resources. Many studies have employed transfer learning, synthetic EC generating, augmentation with GANs, and innovative preprocessing to mitigate issues of sparse data, class imbalance, and overall model robustness within the field of cardiovascular health [6, 12, 16, 18, 19].

The review suggests that there is a growing focus on the demand for real-time cardiovascular monitoring and smart health solutions. Research that has employed YOLO-based detection systems, gated fusion networks, and the analysis of multi-channel physiological signals to achieve low inference latency has succeeded in high precision diagnosis. Furthermore, the combination of ECG signals with other physiological signals and the analysis of non-linear features has improved disease detection and lowered the false alarm rate. These findings indicate that hybrid deep learning systems have the ability to support continuous patient monitoring, telemedicine, and wearable health systems [5, 10, 14, 15, 20].

Though much progress has been made, many challenges remain in the area of research. Many models that exist today have high computational complexity, low generalization across datasets, poor performance on the less common arrhythmia classes, and have not been validated in real-life clinical settings. In addition, finding the right balance between prediction accuracy, interpretability, computational consumption, and ease of deployment remains one of the greatest challenges. Thus, it is evident that a hybrid deep learning system is needed that focuses on robust feature extraction, attention mechanisms, explainability, and model optimization in order to meet the demand for accurate, interpretable, and clinically actionable cardiovascular disease prediction systems [1, 2, 3, 7, 17, 21,22].

Table 1. Comparative Review of State-of-the-Art Cardiovascular Disease Prediction Models

Ref.	Proposed Model / Technique	Dataset Used	Performance	Limitation
[1]	Cardioformer (ResNet + Transformer + Multi-granularity patching)	PTB, PTB-XL, MIMIC-IV	AUROC: 96.34%	Sensitive to patch-size selection

[2]	ECGBert (1D-CNN + BERT)	MIT-BIH	High Accuracy, Precision, Recall, F1-score	Lower performance in ventricular ectopic beat classification
[3]	CNN-LSTM-FCL Hybrid Model	MIT-BIH	Accuracy: 99.43%	Requires extensive training and real-time validation
[4]	LightX3ECG (1D-SEResNet + Lead-wise Attention + Grad-CAM)	Chapman, CPSC-2018	F1-score: 97.18%, 80.04%	Increased training complexity
[5]	HCTG-Net (CNN + Transformer + Gated Fusion)	MIT-BIH	Accuracy: 99.46%, F1-score: 97.11%	High computational complexity
[6]	CNN + Bi-LSTM + Transformer + SHAP	PhysioNet CinC 2020	Accuracy: 91.24%, F1-score: 83.01%	Reduced performance for similar arrhythmias
[7]	DMSFNet (Multi-scale CNN + Attention)	CPSC_2018, CinC_2017	F1-score: 82.8%, 84.1%	Increased computational complexity
[8]	1D-CNN for VT Alarm Classification	VTaC Benchmark Dataset	Accuracy: 83.22%, AUC: 0.901	Limited noisy ICU validation
[9]	2D-CNN-LSTM with CWT	MIT-BIH, CHF, NSR	Accuracy: ~99%	High computational cost
[10]	CNN/DNN + XAI (SHAP, LIME)	MITDB, PTBDB, NSTDB	Accuracy: 99.74%	Higher memory and resource requirement
[11]	LightX3ECG + Grad-CAM + DSConv	Chapman, CPSC-2018	F1-score: 0.9718	Limited to 3 ECG leads
[12]	CNN + Attention-based RNN (GRU)	MIT-BIH, CPSC	F1-score: 0.9110	Requires balanced subject-specific training
[13]	ArrhythmiNet V1/V2 (Lightweight 1D-CNN)	MIT-BIH	Accuracy: 99%, 98%	Limited to single-lead ECG
[14]	HeartNet (CNN + Self-Attention + GAN)	MIT-BIH, PTB, AF Dataset	Accuracy: 99.67%	Difficulty with highly imbalanced datasets
[15]	HARDC (BiGRU-BiLSTM + Dilated CNN + Attention + CGAN)	MIT-BIH	Accuracy: 99.60%	High architectural complexity
[16]	DeepArrNet (Lightweight CNN)	MIT-BIH, PTB	Accuracy: 99.28%	Depends on accurate R-peak detection
[17]	BRDC (BiGRU + BiLSTM + Dilated CNN + GAN)	MIT-BIH	Accuracy: 99.90%	Needs larger dataset validation
[18]	Modified YOLOv8 for ECG Classification	MIT-BIH	Accuracy: 99.5%	High GPU requirement

[19]	Transfer Learning-based 1D-CNN	PTB-XL, PTB-XL Plus, DeepFake ECG	AUC: 0.906	Limited transferability in synthetic ECG
[20]	MM-GradCAM (Multimodal Explainable AI)	Chapman ECG	Accuracy: 97.44%	Increased multimodal processing complexity
[21]	LSTM + FSST-based ECG Labeling	QT Database	Segmentation Accuracy: 86.6%	Moderate segmentation accuracy
[22]	AlexNet + HRV + Chaos Theory Features	MIT-BIH, BIDMC	Accuracy: 98.75%	Requires handcrafted HRV features

Challenge

Significant progress has been made in the hybrid deep learning-based prediction of cardiovascular diseases. However, some fundamental issues remain, limiting their practical use. Most of the current research is based on ECG signals and does not include other clinically important data, such as health records, lab data, images, demographics, and data on lifestyle. Many deep learning models are black box solutions, and their unsolved problem of severe class imbalance in cardiovascular datasets, especially the prediction of rare arrhythmias, makes them less useful and more difficult to trust for clinicians.

Many models report good performance on benchmark datasets. However, they do not generalize when applied to other datasets from different clinical contexts, including different patient populations and different data collection methods. The hybrid frameworks built on a combination of CNNs, LSTMs, Transformers, and attention, for example, have a high computational cost. Even with promising results in research, the lack of extensive clinical research studies has been a significant limiting factor for their use in healthcare. The need for hybrid deep learning models for the prediction of cardiovascular diseases that are interpretable, efficient, and clinically validated is evident.

Research Gap

A detailed hybrid framework, encompassing all elements of multi-scale CNN feature extraction, Transformer contextual awareness, BiLSTM temporal analysis, XAI, and the fusion of multimodal healthcare data along with easy deployment, is yet to appear in the literature with respect to the prediction of cardiovascular diseases. The combinations of modules, as showcased in the aforementioned frameworks, lack the ability to capture the complex spatial-temporal-contextual relationships and do so within the constraints of explainability and computational efficiency. Thus, the integrated frameworks are bound to increase the prediction accuracy and the trust level of the end users (clinicians). It will also allow the use of different types of healthcare data and promote the use of online (real time) clinical and wearable healthcare tools. All these factors will make the prediction models of cardiovascular diseases more robust, reliable, and practical.

Conclusion

The use of CNNs, recurrent networks, Transformers, and deep learning architectures, together with explainable AI and different attention mechanisms, has opened new avenues to enhance the predictive capabilities and diagnostic accuracy in the automated analysis of ECGs. Hybrid architectures combining these elements will help augment automated ECG analysis. As promising classification results are beginning to emerge with AI solutions on a plethora of benchmark datasets in the literature, the future gains in healthcare systems will shift the focus to solving the remaining roadblocks. Multimodal integration, explainability, and computational consideration, in addition to clinical justification, will limit the impact of these systems if not developed in conjunction with the systems. Further work in hybrid architectures with clinically justified, scalable, and intelligent explainable AI on integrated healthcare systems will play an important role to augment future cardiovascular and precision medicine systems.

References

1. Kamrujjaman Mobin, M., Saiful Islam, M., Al Barid, S., & Masum, M. (2025). Cardioformer: Advancing AI in ECG Analysis with Multi-Granularity Patching and ResNet. *arXiv e-prints*, arXiv-2505.
2. Liu, H. (2025). A hybrid model combining 1D-CNN and BERT for intelligent ECG arrhythmia classification. *Scientific Reports*.
3. Ramachandran, D., Kumar, R. S., Alkhayyat, A., Malik, R. Q., Srinivasan, P., Priya, G. G., & Gosu Adigo, A. (2022). Classification of electrocardiography hybrid convolutional neural network-long short term memory with fully connected layer. *Computational Intelligence and Neuroscience*, *2022*(1), 6348424.
4. Le, K. H., Pham, H. H., Nguyen, T. B., Nguyen, T. A., Thanh, T. N., & Do, C. D. (2023). Lightx3ecg: A lightweight and explainable deep learning system for 3-lead electrocardiogram classification. *Biomedical signal processing and control*, *85*, 104963.
5. Xiong, N., Wei, Z., Wang, X., Wang, Y., & Wang, Z. (2025). HCTG-Net: A Hybrid CNN–Transformer Network with Gated Fusion for Automatic ECG Arrhythmia Diagnosis. *Bioengineering*, *12*(11), 1268.
6. Al-Bairmani, M. T., Yazdchi, M., & Nasimi, F. (2026). A joint CNN-Bi-LSTM-transformer architecture with SHAP explanations for multi-label arrhythmia detection from 12-lead ECGs. *Scientific Reports*.
7. Wang, R., Fan, J., & Li, Y. (2020). Deep multi-scale fusion neural network for multi-class arrhythmia detection. *IEEE journal of biomedical and health informatics*, *24*(9), 2461-2472.
8. Khanolkar, U., Yadav, A., Mann, A., Pappada, S., Deshmukh, A., & Maan, A. (2026). Role of a Deep-Learning Based Convolutional Neural Network Model for Real-Time Ventricular Tachycardia Alarm Classification. *Journal of Cardiovascular Electrophysiology*.
9. Madan, P., Singh, V., Singh, D. P., Diwakar, M., Pant, B., & Kishor, A. (2022). A hybrid deep learning approach for ECG-based arrhythmia classification. *Bioengineering*, *9*(4), 152.
10. Talukder, M. A., Talaat, A. S., Muna, N. J., Alazab, A., Kazi, M., & Das, U. K. (2025). An explainable deep learning framework for trustworthy arrhythmia detection from ECG signals. *Scientific Reports*, *15*(1), 39496.
11. Le, K. H., Pham, H. H., Nguyen, T. B., Nguyen, T. A., Thanh, T. N., & Do, C. D. (2023). Lightx3ecg: A lightweight and explainable deep learning system for 3-lead electrocardiogram classification. *Biomedical signal processing and control*, *85*, 104963.
12. Sun, J. (2023). Automatic cardiac arrhythmias classification using CNN and attention-based RNN network. *Healthcare Technology Letters*, *10*(3), 53-61.
13. Baig, Z., Nasir, S., Khan, R. A., & Haque, M. Z. U. (2025). Arrhythmiavision: Resource-conscious deep learning models with visual explanations for ecg arrhythmia classification. *arXiv preprint arXiv:2505.03787*.
14. Rafi, T. H., & Ko, Y. W. (2022). HeartNet: Self multihead attention mechanism via convolutional network with adversarial data synthesis for ECG-based arrhythmia classification. *IEEE Access*, *10*, 100501-100512.

15. Islam, M. S., Hasan, K. F., Sultana, S., Uddin, S., Quinn, J. M., & Moni, M. A. (2023). HARDC: A novel ECG-based heartbeat classification method to detect arrhythmia using hierarchical attention based dual structured RNN with dilated CNN. *Neural Networks*, *162*, 271-287.
16. Mahmud, T., Fattah, S. A., & Saquib, M. (2020). DeepArrNet: An Efficient Deep CNN Architecture for Automatic Arrhythmia Detection and Classification From Denoised ECG Beats. *IEEE Access*, *8*, 104788-104800.
17. Islam, M. S., Islam, M. N., Hashim, N., Rashid, M., Bari, B. S., & Al Farid, F. (2022). New hybrid deep learning approach using BiGRU-BiLSTM and multilayered dilated CNN to detect arrhythmia. *IEEE Access*, *10*, 58081-58096.
18. Ang, G. J. N., Goil, A. K., Chan, H., Lee, X. C., Mustaffa, R. B. A., Jason, T., ... & Shen, B. (2023). A novel application for real-time arrhythmia detection using yolov8. *arXiv preprint arXiv:2305.16727*.
19. Jayasundara, R., Fernando, I., Fernando, A., Ragel, R., Thambawita, V., & Nawinne, I. (2025). Inductive transfer learning from regression to classification in ECG analysis. *arXiv preprint arXiv:2508.11656*.
20. Murat Duranay, F., Murat, E., Yıldırım, Ö., Demir, Y., Tan, R. S., Sampathila, N., & Acharya, U. R. (2026). MM-GradCAM: an improved multimodal GradCAM method with 1D and 2D ECG data for detection of cardiac arrhythmia. *Scientific Reports*.
21. Hussein, O., Jameel, S. M., Altmemi, J. M., Abbas, M. A., Uğurenver, A., Alkubaisi, Y. M., & Sabry, A. H. (2025). Improving automated labeling with deep learning and signal segmentation for accurate ECG signal analysis. *Service Oriented Computing and Applications*, *19*(2), 93-106.