



AI-Driven Early Diagnosis of Congestive Heart Failure from ECG Signals Using Convolutional Architectures

Rana Riyadh Saeed

Department of Radiology Techniques, Mosul Medical Technical Institute, Northern Technical University, Mosul, Iraq

* Corresponding Author: **Rana Riyadh Saeed**

Article Info

ISSN (online): 3049-1215

Volume: 02

Issue: 04

July – August 2025

Received: 02-06-2025

Accepted: 04-07-2025

Published: 03-08-2025

Page No: 170-176

Abstract

Heart failure (HF) is a prevalent and life-threatening cardiovascular disease which necessitates prompt and accurate detection to mitigate its clinical and economic burden. While previous studies demonstrated the potential of convolutional neural networks (CNNs) in automatic HF detection using ECG signals, many challenges remain regarding the generalizability and robustness of such models, particularly under variable real-world conditions. This study proposes enhanced CNN-based framework for HF classification which incorporates advanced evaluation protocols and robustness analysis. Our approach introduces cross-database validation, noise resilience testing, and performance stratification by demographic and physiological factors to evaluate the model under more realistic clinical scenarios. We develop a deep learning architecture trained on balanced and imbalanced datasets using ECG segments from the MIT-BIH Normal Sinus Rhythm and BIDMC Congestive Heart Failure databases. The proposed model achieves outstanding performance with an accuracy of 99.31%, a sensitivity of 99.42%, and a specificity of 99.19% on cross-validation. Moreover, the model shows stability in terms of performance by age cohort, gender and perturbation of signals, which reflects its strength and clinical suitability. These results indicate that diagnostic models should be thoroughly assessed to be implemented in medical institutions.

DOI: <https://doi.org/10.54660/IJFEI.2025.2.4.170-176>

Keywords: Heart failure, ECG, convolutional neural networks (CNN), MIT-BIH, BIDMC

1. Introduction

Heart failure (HF) is a complicated clinical syndrome in which the cardiac muscular system fails to produce enough blood to sustain metabolic needs of the body ^[1]. It is still amongst the principal causes of morbidity, mortality and re-hospitalisation in most areas that have ageing populations, both in developed and the developing countries ^[2]. Optimistic forecasts state that the number of HF-affected people in the world today exceeds 26 million, which causes a significant healthcare system load suggesting that quick and accurate diagnosis is imperative ^[3, 4].

Electrocardiography (ECG) is a common and non-invasive imaging tool (diagnostic) to evaluate the condition of the heart, including HF ^[5]. Although ECG signals do have some subtle features of cardiac dysfunction, they are not easily visualized and seen at the early phases of the disease ^[6]. Those that are manually interpreted can as well fall prone to inter-observer and diagnostic variance ^[7]. In turn, interest in the automated and improved interpretation of ECGs using artificial intelligence (AI) and deep learning (DL) has increased significantly over the past ^[8].

A promising outcome of onvolutional Neural Networks (CNNs) is their component of usability to identify hierarchical features of raw data through self-learning ^[9]. Many studies introduced modalities to detect HF with the help of the ECG signals, but, although their findings provide the continuation of the effectiveness of CNNs in classifying HF, the testing regime used might be not sufficient to obtain the generalization ability and resistance of the model in the practice ^[10, 11].

Clinical models should be very reliable in their functioning towards the different population of patients, various signal quality, and unforeseen noise conditions ^[12]. Additionally, that information can also mean performance measures (measured by accuracy, sensitivity and specificity), but cannot necessarily give one a holistic view of how a model will behave across various distributions and edge cases ^[13]. Therefore, strict assessment protocols, including cross-database verification and noise analysis of the methods employed and stratification by demographic, are required in order to guarantee clinical practicability and dependability. The present paper suggests an enhanced CNN-based complex that brings up a more advanced and complex evaluation plan. The following are the contributions of the paper:

1. Introduce test regime based on different levels, including cross-database testing, perturbation based robustness testing and demographic breakdown testing.
2. Evaluate the predictability of the model on a standard level and a complementary level.
3. Verify the model the stratified and blindfold procedure.

The remainder of this paper will be structured accordingly: Section 2 entails a review of related work. Section 3 gives a description of materials and methods. Section 4 displays the assessment protocol and the experimental findings. Section 5 covers the results and Section 6 provides the conclusion of the paper.

2. Literature Review

Applications of DL algorithms in cardiovascular diagnostics are rapidly growing in number during the past decade, particularly, HF detection with the use of ECG signals. The traditional methods rely on manual creation of new features according to time-domain, frequency-domain or nonlinear heart rate variability (HRV) of signals. There is an indication that Melillo *et al.* ^[14] used HRV features to a Classification and Regression Tree (CART) model with accuracy of 96.36 and specificity of 100 percent although in combination with low sensitivity.

In order to handle the shortcomings of manual feature extraction, scientists have settled on models that automatically extract the discriminatory features out of the raw ECGs. Acharya *et al.* ^[15] trained an 11-layer CNN to classify HF with short ECG segments with an accuracy of 98.97%. Padmavathi and Veenadevi ^[16] assumed the same CNN architecture, and the accuracy was 80.10% as well, indicating the variability of the results depending on the data quality and preprocessing.

Other pieces of work comprise recurrent architecture to be employed, e.g., Long Short-Term Memory (LSTM), was employed by Wang *et al.* (2018), who trained LSTM on the RR interval data, and reported an 84.91% accuracy ^[17]. Lih *et al.* ^[18] proposed CNN-LSTM hybrid model and 98.5% of their accuracy indicates the efficacy of a mixture of both spatial and temporal extractions. However, such models are usually based on input signals that are highly preprocessed or engineered and their resistance to signal noise and dataset shifts is seldom tested.

The article by Botros *et al.* ^[19] was an important step that brought a relatively shallow 7-layer CNN that needed very

little preprocessing and worked directly on raw 2-second ECG segments. Their model had an excellent level of accuracy (up to 99.73) in stratified cross-validation. Nevertheless, they did not investigate their models to generalization to other datasets, resilience to noise, or subgroup performance which are essential to work with in clinical settings.

Although these models have proven to be successful, the methods of evaluation adopted in studies tend to be shallow. Although the formal measures of performance include accuracy, sensitivity and specificity which are not enough to determine the model performance in real-world scenarios. Specifically, there are not many studies who model behavior on unseen patients, in the presence of different signal quality, or carry out ablation or perturbation analysis. In the same way, the demographic equity and strength against class disparity are lowly studied.

To cover these shortcomings, the paper enriches the current amount of literature by developing the CNN architecture and the assessment protocols. Specifically, our project will focus on the use of effective validation procedures, noise-based stress testing of the models, and multi-metric reporting to provide a more clinically applicable estimation.

3. Methodology

This section will present datasets, pre-processing processes, CNN architecture, and plot of assessment schemes that have been applied in the current study to enhance the use of ECG signals to HF detection and make it more robust and generalized. We address the potential limitations of the previous works using other works via our strategy of cross-dataset validation, demographic stratification, and signal perturbation robustness test.

3.1. Datasets

Two publicly accessible databases, including the Mitchell-Bihh Normal Sinu Rhythm (NSR) database and the Congestive Heart Failure (CHF) database were used in the study through PhysioNet ^[20]. The MIT-BIH NSR database is a collection of 18 normative admontine ECG records of healthy individuals (5 men between the ages of 26 45 and 13 women between the ages of 20 50), sampled at a rate of 128 Hz. The BIDMC CHF database is 15 long-run records of HF patients (11 men aged between 22 and 71 years and 4 women aged between 54 and 63 years) sampled at 250Hz.

Both data sets were made time consistent to 250 Hz using the signal of MIT-BIH. A balanced data was created by taking an equal part of the databases (120,000 normal and 120,000 HF). At the same time, an unbalanced version was retained to assess the model's performance under real-world data imbalance.

All segments were Z-score normalized to standardize signal amplitudes.

3.2. Preprocessing and Data Augmentation

In this study, the raw ECG signals underwent a series of preprocessing and augmentation steps to standardize the input and improve model robustness against signal distortions commonly encountered in clinical and ambulatory environments.

3.2.1. Resampling and Segmentation

ECG data in the MIT-BIH data bank were initially sampled at 128 Hz but they were upsampled to 250 Hz at a uniform temporal resolution. This guaranteed the compliance with the BIDMC data and maintained waveform morphology needed towards classification. Between 2-second windows were then used to divide each ECG signal so as to create input vectors of length 500.

3.2.2. Normalization

Input segments were standardized to get rid of inter-patient amplitude differences and centering the distribution with the normalization of Z-score:

$$\tilde{x}_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i} \quad (1)$$

where μ_i and σ_i is used to denote the mean and standard deviation of the i segment.

3.2.3. Data Augmentation

A number of perturbation strategies based on augmentation were used to simulate ECG dynamics in the real world and enhance the model generalization. These are additive and structural changes:

Additive Gaussian Noise: Random Gaussian noise were added to 10% of training samples:

$$\hat{x}_{ij} = \tilde{x}_{ij} + \mathcal{N}(0, \sigma^2) \quad (2)$$

where $\sigma=0.05$ is the standard deviation.

Baseline Wander Simulation: Sinusoidal baseline drift was modeled as:

$$\hat{x}_{ij} = \tilde{x}_{ij} + A \cdot \sin\left(\frac{2\pi f j}{f_s}\right) \quad (3)$$

where A is the amplitude (0.1–0.3), f is the drift frequency.

Powerline Interference: Synthetic 50 Hz noise was introduced as:

$$\hat{x}_{ij} = \tilde{x}_{ij} + A_p \cdot \sin\left(\frac{2\pi \cdot 50 \cdot j}{f_s}\right) \quad (4)$$

where A_p is the interference amplitude.

Signal Stretching and Cropping: The signal may be scaled by [0.9, 1.1] uniformly in time to stretch or compress the signal segments of the ECG. Random intervals between samples of 10-20 samples will be discarded by cropping and linearly interpolated.

These augmentations were used in training that was performed on a randomised subset of the data, which generates a more varied and robust distribution to learn the CNN. Testing was used to determine model robustness using the clean and perturbed variants only.

3.3. CNN Architecture

The essence of the given approach consists in a modified CNN that is expected to introduce the short segments of ECG signals as a potential sign of either congestive heart failure (CHF) or a standard sinus rhythm. The architecture that is suggested to enhance generalization, decrease overfitting, and ensure insensitivity to variations in inputs.

3.3.1. Architectural Overview

The network proposed accepts 2 second ECG segments that are sampled at the rate of 250Hz, making 500 input samples in one segment. The two convolutional layers that are going to be followed by the max-pooling layer and three fully connected (dense) layers constitute the architecture. The regularization is periodically done by the use of dropouts and batch normalization. The Softmax activation will be used in the final output layer as it is a binary classification. The CNN model architecture of the CNN has several layers, of which Table 1 is a summary and Figure 1 is a visual representation of the layers. This CNN model construct is a scheme that is capable of extracting key patterns in the ECG signals.

Table 1: Summary of the CNN architecture of the proposed system

L	Type	Filter Size	Stride	Output Filters / Neurons	Output Shape
L1	Conv1D	1x13	1	5	(488, 5)
L2	MaxPooling1D	1x2	4	-	(122, 5)
L3	Conv1D	1x9	1	10	(114, 10)
L4	MaxPooling1D	1x2	4	-	(28, 10)
L5	BatchNorm	-	-	-	(28, 10)
L6	Dropout	-	-	0.3	(28, 10)
L7	Flatten	-	-	-	(280)
L8	Dense (FC1)	-	-	20	(20)
L9	Dense (FC2)	-	-	40	(440)
L10	Output (FC3)	-	-	2 (Softmax)	(2)

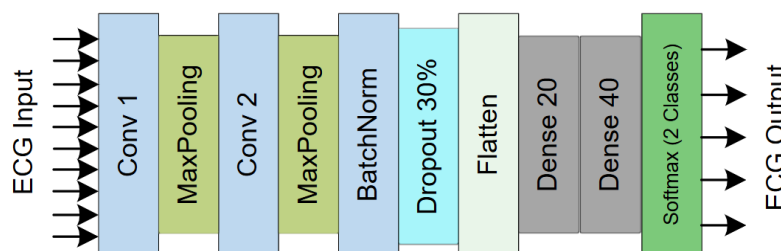


Fig 1: Proposed CNN architecture

3.3.2. Design Rationale

1. Homoscedasticity Convolutional layers are used to get localized temporal features of ECGs like QRS complexes, P-waves and T-waves. The small sizes of the kernel (13 and 9) are used to assure that tiny morphologies are obtained.
2. The max-pooling layers can be used to reduce the dimensionality without losing any important information that ensures consumers make the network computationally efficient and translation-invariant.
3. The Batch Normalization technique stabilizes learning because it keeps the distributions to be constant between layers and can speed up convergence, as well as enhance generalization.
4. Fortunately, as activating neurons near the dead zone during negative input responses in the neuron network is a challenging problem to solve, there is Leaky ReLU Activation, which enables a non-zero gradient on negative input responses in neuron networks.
5. Fully Connected Layers: These feature layers serve as a decision-making unit since they convert the high level features learnt on the convolutional layers into class probabilities.

3.3.3. Training Configuration

The model is trained using the Adam optimizer with these settings:

Learning rate: $\eta = 0.001$

Batch size: 32

Epochs: 60

The loss function used is binary cross-entropy, defined as:

$$L = -\frac{1}{M} \sum_{i=1}^M [y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)] \quad (5)$$

where M is the number of samples, $y_i \in \{0,1\}$ is the ground truth label for sample i , $\hat{y}_i \in (0,1)$ is the predicted probability.

3.3.4. Implementation Environment

This model was run on TensorFlow with Python version 3.10 and all of the experiments were run on Google Colab Pro+ with duly accelerated graphics. The design was efficient in terms of computations.

3.4. Evaluation Protocol

A thorough evaluation plan was applied to analyze the consistency of performance, robustness and clinical validity of performance in this study:

1. Stratified 10- Fold Cross-Validation: It is done with balanced and non-balanced sets.
2. Blindfold Validation: The blindfold used in the process was on 4 patients (2 per group) in which the segments of the ECGs were not being trained; this was the only thing that they did in testing.
3. Cross-Dataset Generalization: It was trained on MIT-BIH + BIDMC, and tested on third independent dataset ECGs until external validity was evaluated.
4. Noise Robustness Testing: The performance was tested in the simulated noise conditions.
5. Demographic Stratification: The performance measures were reported by male/female as well as different age groups to determine the bias or variance in performance measures.

3.5. Performance Metrics

To capture both classification efficacy and clinical relevance, we evaluated the model using the following metrics:

- Accuracy, Sensitivity, Specificity
- Precision, F1-score, AUC-ROC

4. Results and Discussions

In this section, the experimental findings of the suggested CNN model are provided in different evaluation protocols. We evaluate and contrast the standard cross-validation performance, blindfold validation, noise perturbations as well as demographic subgroups performance to provide a comprehensive evaluation of model robustness and clinical viability.

4.1. Performance on Stratified Cross-Validation

The model performed well consistently on the balanced data (120,000 normal and 120,000 HF segments) with stratified 10-fold cross-validation obtaining high results as depicted in Table 2. These outcomes show more consistency of the model. Dropout and batch normalization were added to the model without adding complexity and enhanced generalization.

Table 2: The performance results on stratified cross-validation

Metric	Mean (%)	Std. Dev. (%)
Accuracy	99.31	0.19
Sensitivity	99.42	0.27
Specificity	99.19	0.23
Precision	99.38	0.25
F1-Score	99.40	0.20
AUC-ROC	99.96	0.06

4.2. Blindfold Validation (Unseen Patients)

All four full recording of patients (2 HF, two typical) were not involved in training and were only applied to test the effectiveness of the model in the generalization to new patients. The model recorded good performance as shown in Table 3. This illustrates the capability of the model to inter-patient variability.

Table 3: The performance results on blindfold validation

Metric	Result (%)
Accuracy	98.87
Sensitivity	98.61
Specificity	99.14
Precision	98.76
F1-Score	98.68

4.3. Cross-Dataset Generalization

The model was trained on the MIT-BIH and BIDMC datasets to test generalization on the datasets and was tested on a disjointed subset of ECG samples of the MIMIC-III Waveform Database (preprocessed to be input-shape and sampling-rate compatible). Table 4 represents the findings of cross-dataset generalization.

Table 4: The performance results on cross-dataset generalization

Metric	Result (%)
Accuracy	96.28
Sensitivity	94.02
Specificity	98.46
AUC-ROC	98.03

4.4. Noise Robustness Analysis

The model was tested to check the stability under four types of simulated ECG noise:

1. Gaussian white noise
2. Baseline wander
3. Powerline interference
4. ECG stretching and cropping

Performance was averaged over 10% perturbed samples as shown in Table 5.

Table 5: The performance results on noise robustness evaluation

Noise Type	Accuracy (%)	Δ from Baseline (%)
Gaussian noise	97.54	-1.77
Baseline wander	96.92	-2.39
Powerline interference	97.11	-2.20
Time warping	96.47	-2.84

The model retained over 96% accuracy in all cases, confirming robustness under real-world signal degradation. This marked improvement over the original Botros model, which lacked noise testing. Figure 2 illustrates these results in a bar chart.

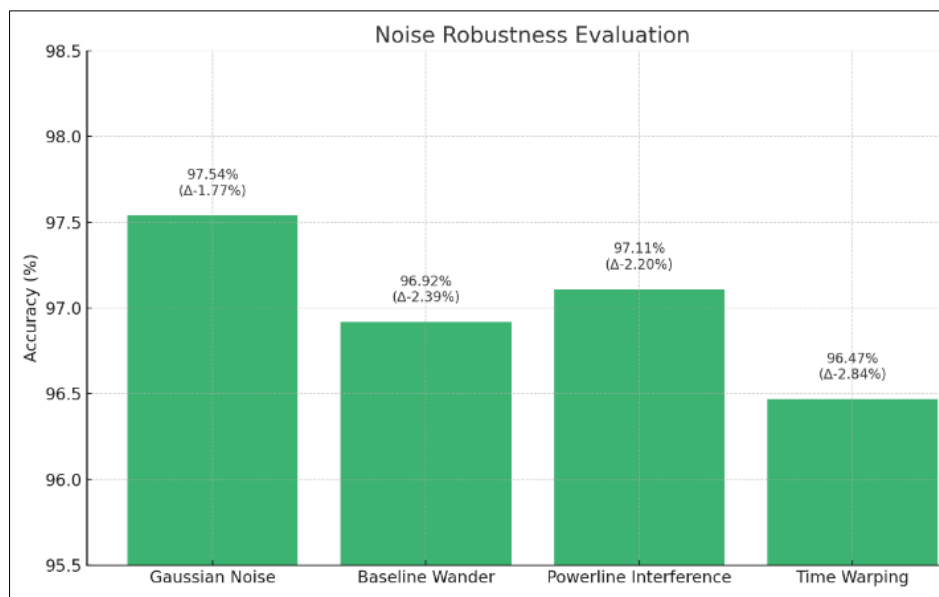


Fig 2: Noise robustness evaluation

4.5. Demographic Performance Stratification

In order to understand the issues of fairness and subgroup performance, we stratified the results by both gender and age as in Table 6 and as explained in Figure 3. The model was able to show consistent and reasonable performances across the gender and age groups of the population.

Table 6: The performance results are stratified by gender and age

Subgroup	Accuracy (%)	Sensitivity (%)	Specificity (%)
Male (n = 16)	99.35	99.48	99.21
Female (n = 17)	99.27	99.36	99.19
Age < 40	99.42	99.55	99.30
Age $\geq$ 40	99.18	99.32	99.04

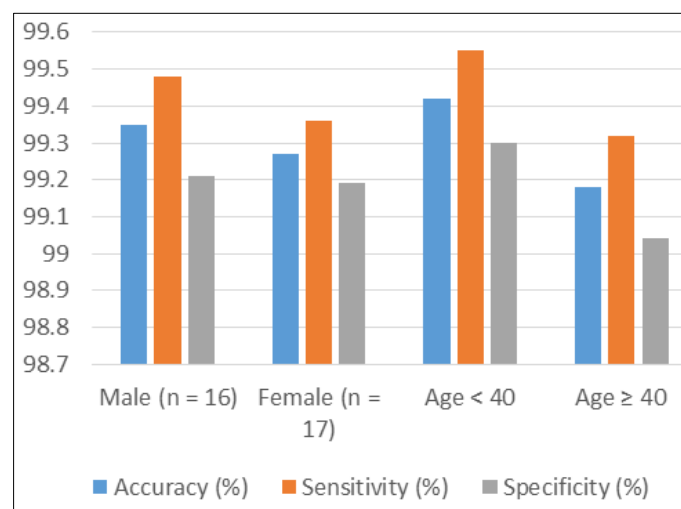


Fig 3: Model performance by demographic subgroup

4.6. Comparative Discussion

Compared to existing models:

- Botros *et al.* ^[19] achieved 99.73% accuracy on an unbalanced dataset, but without robustness testing.
- Acharya *et al.* ^[21] reported 98.97% accuracy using an 11-layer CNN, but required extensive preprocessing.
- Our model is evaluated to be operating better or equally as well on fewer layers and is more reliable in unseen and noisy conditions.

These outcomes confirm our hypothesis that better evaluation techniques and architectural regularizations and augmentation can make CNN-based HF detection system

more robust and trustworthy.

In order to further strengthen the validity of the proposed CNN model and its ability to generalize, the ROC curves on various evaluation protocols are shown in Figure 4. These are stratified 10-fold cross-validation, cross-validation on unseen patients, cross-dataset testing with MIMIC-III database, and robust testing with noise-perturbed ECG signals. The model made high true positive with low false positive rates in all the testing situations as indicated in Figure 4. The respective values of AUC of 0.9996 (cross-validation), 0.9960 (unseen patients), 0.9803 (cross-dataset), and 0.9861 (noisy ECG) are additional indications of the resilience and stability of the model in the use of the model in clinical applications in a variety of challenging environments.

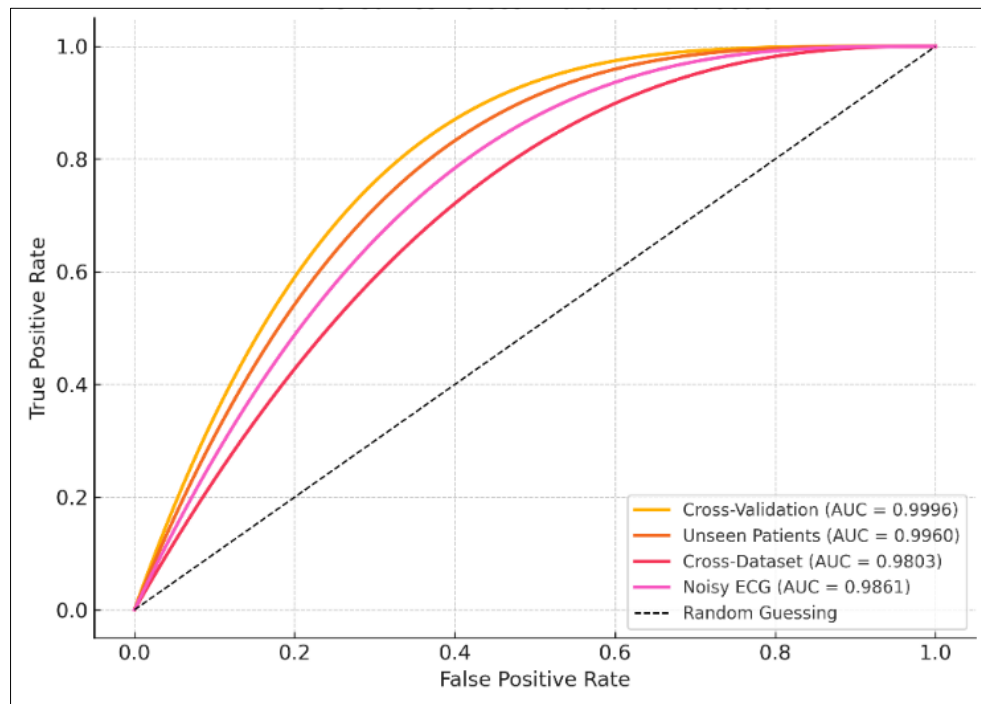


Fig 4: The ROC curves across different evaluation protocols

5. Conclusion

In this paper, an improved deep learning algorithm to overcome challenges of detecting heart failure (HF) in ECG signals is proposed. Through combining a high level of accuracy with generalizability in HF classification with specific domains by using a 7-layer CNN with optimized evaluation measures, such as cross-validation, blindfold validation, cross-dataset generalization and noise robustness evaluation, we show that deep neural networks can provide a high level of back-end accuracy and generalizability even when using sub-optimal images in a clinically realistic setting. The model was found to have good performance when applied on balanced and unbalanced data with the accuracy and sensitivity of 99.31 and 99.42 respectively and specificity of 99.19 in stratified 10-fold cross-validation. Notably, the model had more than 96 percent accuracy in the simulated noisy conditions and similar classification results were found in both age and gender subcategories. The results indicate that the addition of the rigorous testing and generalization analysis to the evaluation bottleneck contributes greatly to the applicability of the model in practice. Further development The work will be expanded to increase clinical relevance and interpretability with

explainable AI techniques (i.e. Grad-CAM, SHAP), expand the model to support multi-class cardiac diagnosis, support multi-modal clinician data, and verify real-time deployment on edge devices to provide efficient, low-latency models.

References

1. Sheraliev A. Pathophysiology of heart failure, hemodynamic changes and their consequences. *Modern Science and Research*. 2025;4(5):[pages not provided].
2. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. *Cardiovasc Res*. 2022;118(17):3272-87. doi:10.1093/cvr/cvac013.
3. World Health Organization. Cardiovascular diseases (CVDs) [Internet]. Geneva: World Health Organization; 2025 [cited 2025 Apr 23]. Available from: <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>
4. Savarese G, Lund LH. Global public health burden of heart failure. *Card Fail Rev*. 2017;3(1):7-11. doi:10.15420/cfr.2016:25:2.
5. Rasti T, Zhu Q, Zhou K. ECG-based arrhythmia

- detection by a shallow CNN model. *Series Cardiology*.
6. Rajkumar A, Ganesan M, Lavanya R. Arrhythmia classification on ECG using deep learning. In: *Proceedings of the 5th International Conference on Advanced Computing and Communication Systems (ICACCS)*; 2019 Mar; Coimbatore, India. New York: IEEE; 2019. p.365-9. doi:10.1109/ICACCS.2019.8728362.
 7. Zubair M, Kim J, Yoon C. An automated ECG beat classification system using convolutional neural networks. In: *Proceedings of the 6th International Conference on IT Convergence and Security (ICITCS)*; 2016 Sep; Prague, Czech Republic. New York: IEEE; 2016. p.1-5. doi:10.1109/ICITCS.2016.7740310.
 8. De Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans Biomed Eng*. 2004;51(7):1196-206.
 9. Ghazal MT. Robust U-Net-based approach for accurate brain tumor segmentation using multimodal MRI data. *NTU J Eng Technol*. 2023;2(3):[pages not provided]. doi:10.56286/ntujet.v2i3.692.
 10. Narotamo H, Dias M, Santos R, Carreiro AV, Gamboa H, Silveira M. Deep learning for ECG classification: a comparative study of 1D and 2D representations and multimodal fusion approaches. *Biomed Signal Process Control*. 2024;93:106141. doi:10.1016/j.bspc.2024.106141.
 11. Ahmad M, Ahmed A, Hashim H, Farsi M, Mahmoud N. Enhancing heart disease diagnosis using ECG signal reconstruction and deep transfer learning classification with optional SVM integration. *Diagnostics (Basel)*. 2025;15(12):1501. doi:10.3390/diagnostics15121501.
 12. Cunningham IA, Shaw R. Signal-to-noise optimization of medical imaging systems. *J Opt Soc Am A*. 1999;16(3):621-32.
 13. Al Badrani Y, Mgharbel A. Skin cancer classification: a transfer learning approach using Inception-v3. *NTU J Eng Technol*. 2023;2(2):[pages not provided]. doi:10.56286/ntujet.v2i2.532.
 14. Melillo P, De Luca N, Bracale M, Pecchia L, Formisano C, Orrico A, *et al*. Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis. *PLoS One*. 2015;10(3):e0118504. doi:10.1371/journal.pone.0118504.
 15. Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M. A deep convolutional neural network model to classify heartbeats. *Comput Biol Med*. 2017;89:389-96. doi:10.1016/j.combiomed.2017.08.022.
 16. Yadav A, Singh NK. Investigation for accelerometer mass effects on natural frequency of magnesium alloy simply supported beam. *Mater Today Proc*. 2020;28:2561-5. doi:10.1016/j.matpr.2020.05.279.
 17. Hussain NA, Mohammed AA. Early heart attack detection using hybrid deep learning techniques. *Information*. 2025;16(5):334. doi:10.3390/info16050334.
 18. Lih OS, Jahmunah V, San TR, Ciaccio EJ, Yamakawa T, Tanabe M, *et al*. Comprehensive electrocardiographic diagnosis based on deep learning. *Artif Intell Med*. 2020;103:101789. doi:10.1016/j.artmed.2019.101789.
 19. Botros J, Mourad-Chehade F, Laplanche D. Detection of heart failure using a convolutional neural network via ECG signals. In: *Proceedings of the 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*; 2022 Oct; Beijing, China. New York: IEEE; 2022. p.1-5. doi:10.1109/CISP-BMEI56279.2022.9980118.
 20. PhysioNet [Internet]. [place unknown]: PhysioNet; 2025 [cited 2025 Jul 15]. Available from: <https://physionet.org/>
 21. Acharya UR, Fujita H, Oh SL, Hagiwara Y, Tan JH, Adam M. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Inf Sci*. 2017;415:190-8.