Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol.14 , Issue 32s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

End-to-End Deep Learning and Machine Learning Framework for Chronic Heart Failure Detection from Phonocardiogram Signals.

Komati Lavanya [1], Kanne Naveen [2], Gazala Nowsheen [3], Palle Vandana [4], Rajakumar Rudrarapu[5]

[1]Assistant Professor Computer Science and Engineering, VMR Pradeep Kumar Institute of Engineering and Technology, Hanumkonda, Telangana, India

[2]Assistant Professor Electronics & Communication Engineering Vaagdevi Engineering College, Warangal, Telangana, India-506005

[3]Assistant Professor Electronics & Communication Engineering, VMR Pradeep Kumar Institute of Engineering and Technology, Hanumkonda, Telangana, India

[4]Assistant Professor Electronics & Communication Engineering, VMR Pradeep Kumar Institute of Engineering and Technology, Hanumkonda, Telangana, India

[5]Assistant Professor Electronics & Communication Engineering, Vaagdevi Engineering College,Warangal, Telangana, India-506005

.Cite this paper as: Komati Lavanya , Kanne Naveen , Gazala Nowsheen , Palle Vandana , Rajakumar Rudrarapu (2025) End-to-End Deep Learning and Machine Learning Framework for Chronic Heart Failure Detection from Phonocardiogram Signals.. Journal of Neonatal Surgery, 14, (32s) 10741-10751

ABSTRACT

Heart failure (HF), characterized by the heart's inability to effectively pump blood, is a chronic and progressive disease that may lead to death. Early and accurate diagnosis is crucial for optimal patient care and disease management. Though reliable, conventional diagnosis methods such as echocardiogram are costly, require specialized equipment, and are not typically accessible in resource-poor settings. Phonocardiogram (PCG) signals are acquired through non-invasive, affordable digital stethoscopes and thus serve as a promising candidate for automated screening of cardiac diseases. To identify CHF from PCG signals, in this work, a hybrid model based on E2E DL and traditional ML is proposed. Specifically, the method applies robust preprocessing, then segments the cardiac cycle, and extracts discriminative time-frequency features such as spectrograms and Mel-frequency cepstral coefficients (MFCCs). We also compare the end-to-end CNN architectures, which are directly learned from spectrogram representations, with traditional machine learning classifiers, including Random Forests and Support Vector Machines (SVM) with similar features extracted from spectrograms. The proposed CNN-based method outperforms traditional ML algorithms with an accuracy greater than 96%, sensitivity of 95.4%, and specificity of 96. 7% based on extensive evaluation using publicly available and carefully curated CHF-PCG datasets. The results show the potential of deep learning-based heart sound analysis for high performance, low-cost, and widely accessible CHF screening in hospitals and home-based healthcare.

Keywords: *Chronic heart failure, phonocardiogram, heart sounds, machine learning, deep learning, CNN, MFCC, spectrogram, cardiac cycle segmentation, computer-aided diagnosis.*

INTRODUCTION

It is also a leading reason for clinical visits, hospitalization, and healthcare cost among elderly patients. Chronic heart failure (CHF) is characterized by the heart being unable to pump enough blood to meet the metabolic demands of the body, resulting in fatigue, dyspnea and reduced exercise capacity. [2] Early identification and treatment are therefore important to enhance the quality of life, prevent hospitalisations and retard the disease process.

Traditional point-of-care devices for high diagnostic accuracy, including echocardiography, electrocardiography (ECG), and biomarker analysis, require costly equipment, trained personnel, and clinical infrastructure, and are largely unavailable in rural and resource-limited areas [3].

PCG Heart sounds recorded by phonocardiograms (PCGs) using electronic or digital stethoscopes provide a low-cost, non-invasive, and easily accessible means for cardiac screening [4]. They are rich in physiological information about valves activities, ventricular functions and turbulence related to blood flow. Pathological changes in heart sound patterns, e.g., presence of a third heart sound (S3) or changes in systolic and diastolic durations, have been shown to be clinically related

to CHF [5]. An emerging technology of digital auscultation has provided opportunity to record PCG signals with quality suitable for automated analysis, which opens the way for heart sound analysis as a screening method.

Conventional methods for automated cardiac sound analysis involve the manual engineering of features from PCG signals. Well-known features are time domain features, spectral energy distribution, and Mel-frequency cepstral coefficients (MFCCs), Then they are classified with classifiers such as random forests, k-nearest neighbours (KNN), and support vector machines (SVM) [6]. These methods may not fully capture the complex temporal and spectral patterns of cardiac sounds, leading to insufficient generalization capabilities; in addition, their performance is closely related to the quality of feature engineering.

Empowered by the capability to learn discriminative representations from raw or minimally preprocessed PCG recordings, end-to-end deep learning has emerged as a potential alternative [7]. However, RNNS, e.g., long short-term memory (LSTM) networks, are able to capture sequential dependencies within cardiac cycles, taking the dynamic behavior into account, while CNNs can be applied to spectrogram or scalogram-based representations of PCG signals which enables them to capture both local frequency patterns as well as global temporal information at multiple scales [8]. The detection accuracy and model interpretability are further improved by hybrid architectures combining CNNs with attention mechanisms [9].

Nevertheless, there are still some gaps in the research. First, the majority of work is only concerned with deep learning without systematically evaluating the performance of traditional ML pipelines after optimization on the same dataset. Secondly, few researches on the fusion of advanced signal pre-processing techniques such as adaptive noise filtering, cycle segmentation, and augmentation to enhance robustness for real environment application. Third, generalization to other datasets (captured with different devices, in diverse environments) is a still an open question

In this work we fill these holes by proposing a novel hybrid/data fusion framework that evaluates traditional machine learning classifiers against end-to-end deep learning models for CHF detection in PCG signals. The framework includes to some extent robust pre-processing, and also cardiac cycle segmentation and feature extraction (MFCC, spectrogram, wavelet-based descriptors, etc.) feature extraction, and CNN architectures that are directly using time–frequency representations as input. Experimental results on public and non-public CHF datasets show that the proposed deep learning method outperform traditional machine learning methods consistently with the accuracy more than 96% and still keep very good high sensitivities and specificities. The results highlight the potential for PCG-based CHF detection as a low-cost, scalable screening tool for clinical and telehealth practice.

Chronic heart failure (CHF) is a progressive and debilitating disease of the heart that has become a major burden on global health. The Global Burden of Disease Study has estimated that approximately 64.3 million people globally suffer from heart failure, with prevalence rising in elderly populations and in countries undergoing epidemiologic transitions from infectious to non-communicable diseases [10]. CHF develops when the myocardium cannot meet the demands of the body for blood flow or when blood flow can be achieved only at high filling pressures, resulting in symptoms that include dyspnea, fatigue, and fluid retention [11]. Despite advances in pharmacological and device-based treatments, mortality remains high and long-term survival rates have been reported to be lower than for a number of common malignancies [12].

Early diagnosis of CHF is a prerequisite to provider-initiated, disease-modifying interventions for slowing disease progression, decreasing hospitalizations, and improving quality of life for CHF patients [13]. Standard diagnostic techniques, including echocardiography, electrocardiography (ECG), chest X-rays, and determination of natriuretic peptide levels are dependable, but they often require specialized devices and trained health care professionals, thus limiting their availability in primary care and rural settings [14]. In a number of developing countries, the application of such techniques is further challenged by cost and infrastructure constraints, underscoring the need for low-cost, portable diagnostic solutions [15].

One such option is the analysis of heart sounds, acoustic signals produced by the mechanical activity of the heart, mainly due to the opening and closing of cardiac valves and the blood flow turbulence [16]. These sounds are recorded using phonocardiogram (PCG) recordings, which can be stored, transmitted, and analyzed using a computer. Rich diagnostic information about the valve function, myocardial performance and haemodynamic state of the patient is related to heart sounds. In the setting of CHF, certain acoustic markers (such as a third heart sound (S3), variably increased intensity of the 1st (S1) and 2nd (S2) heart sounds, and abnormal systolic–diastolic timing) have already been identified to offer useful insights into the presence of underlying ventricular dysfunction [17].With the advances in digital stethoscope and wearable auscultation systems, high-fidelity PCG recording nowadays can be easily integrated with telemedicine systems [18]. This, in turn, has resulted in the introduction of CADx systems that may analyze PCG signals automatically, identifying potential cardiac pathologies and possibly enabling diagnostics in non-expert environments.

Traditional PCG Signal Processing Based Methods and Machine Learning Approaches

The complete processing of PCG signals for automated analysis has traditionally been a slow process: signal pre-processing, feature extraction, and classification [19]. The pre-processing is intended to eliminate noise of environmental sources, respiration sound, and sensor artifacts by utilizing bandpass filtering, adaptive noise cancellation, wavelet denoising etc. [20]. Feature extraction means that the signal in its raw form is converted into a form that can represent temporal and

spectral-related information in a condensed form. Extracted features typically include:

- Time-domain features: peak amplitudes, heart rate variability, S1 – S2 temporal intervals [21].
- Frequency-domain features: spectral energy, the major frequency components, and harmonics [22].
- Time–frequency properties: Mel-frequency cepstral coefficients (MFCCs), spectrograms, continuous wavelet transform coefficients [23].

Such features are then fed into machine learning classifiers including Support Vector Machines (SVMs), Random Forests, k-Nearest Neighbors (k-NN), Gradient Boosting Machines, or Logistic Regression models [24]. Although these methods achieve promising accuracy on controlled datasets, their performance highly depends on unique conditions of the test datasets, and noisy or diversified real world scenarios often result in performance degradation as handcrafted design.

1.2 FromEnd-to-End Deep Learningarization

The disadvantages of using handcrafted features in high-dimensional data have spurred a paradigm shift towards end-to-end deep learning, where features are learned by neural networks from the raw or lightly processed data [26]. In PCG analysis, CNNs are largely employed to treat time–frequency distributions like spectrograms and scalograms [27]. CNNs are well suited to learn local spatial patterns in the input, which are local frequency events in the PCG signal (such as murmur and gallop).

Sequential models, Recurrent neural networks (RNN) and in particular Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures, have been utilized to model time dependencies over a number of cardiac cycles [28]. Recently, hybrid models based on CNN layers for local feature extraction and RNN layers for temporal modeling have demonstrated superior performance for abnormal heart sounds detection [29]. Interpretability has also been further improved by attention mechanisms which enable the model to concentrate on signal parts that are diagnostically relevant [30].

End-to-end methods have led to the state-of-the art results on the submitted PhysioNet/CinC Challenge datasets for heart sound classification, outperforming the traditional machine learning pipelines in most cases [31]. However, their clinical translation is still limited by the questions of generalization to different patient populations, varying recording devices and environmental noise conditions.

1.3 Gaps in the Literature and Motivation

Although good progress has been made, several major issues still remain:

1. ML baselines benchmarking – Deep learning results are often reported in isolation without computations of rigorous baselines using powerful machine learning techniques on the derived features in the same data [32].
2. Robustness to noise –PCG signals recorded in reality are contaminated with noise of breathing, patient movement, environment and instrumentation. Although several noise-robust models have been introduced, they usually perform poorly in low signal-to-noise ratio (SNR) scenarios [33].
3. Cross-dataset generalization – It has been found that, the performance of a model trained on one dataset might significantly drop when tested on PCG samples collected by different acquisition devices or from different populations [34].
4. Interpretability for Clinicians – Clinicians may be reluctant to adopt black-box models unless they have ways to visualize or explain predictions [35].

This work is intended to fill these gaps by proposing a hybrid approach, combining traditional machine learning and end-to-end deep learning for the detection of CHF using PCG signals. The framework is meant to:

- Make a direct performance comparison between the handcrafted-feature ML pipelines and the spectrogram-based CNN architectures.
- Apply strong pre-processing and data augmentation for better performance in noisy conditions.
- Exploit attention for better interpretability. • Demonstrate generalization on both public and private datasets from 4 different acquisition environments.

The main contributions of this study are listed as follows:

- Hybrid Comparative Framework: Holistic traditional ML and end-to-end DL approaches for CHF detection from PCG signals are implemented and evaluated in a unified framework.
- Sophisticated Pre-processing: Bandpass filtering, adaptive noise suppression and cardiac cycle segmentation are applied to enhance signal quality before further processing.
- Feature-rich ML Models: Operating on MFCCs, Wavelet based features and higher-order spectral descriptors with optimized ML classifiers.
- End-to-End DL Models: We propose CNN-based models to be learned directly from spectrograms for automated feature learning.
- Interpretability: Introduction of Grad-CAM-like visualization methodologies that emphasizes the time–frequency regions

that are responsible for the model predictions.

• Robustness Assessment: Testing on different datasets to validate robustness of the results.

Bridging the gap between traditional and modern learning schemes, the framework achieves a high diagnostic accuracy (above 96%) in a low-complexity PCG-based CHF detection, which is promising for implementation in portable or telehealth devices.

2. METHODOLOGY

The framework presented in Fig. 1 is composed of five stages:

1. Data Acquisition – PCG recordings from public (e.g., PhysioNet/CinC Challenge datasets) and privately collected HF datasets by applying digital stethoscopes.

2. Signal Pre-processing – Bandpass filtering (20–800 Hz), adaptive noise reduction, and normalization that improve the contrasted of heart sound.

3. Cardiac Cycle Segmentation – Logistic Regression–HSMM or envelope-based methods to determine S1–S2 boundaries.

4. Feature extraction –

o ML models: MFCCs, spectral entropy, zero crossing rate, wavelet coefficients;

o DL models: Mel spectrograms and scalograms.

5. Classification –

o Machine Learning: SVM, Random Forest, Gradient Boosting.

o Deep Learning: CNN, CNN–RNN, att ention-based models.

6. Performance Evaluation – Accuracy, sensitivity, specificity, F1-score, ROC–AUC, cross-validation, and cross-dataset testing.

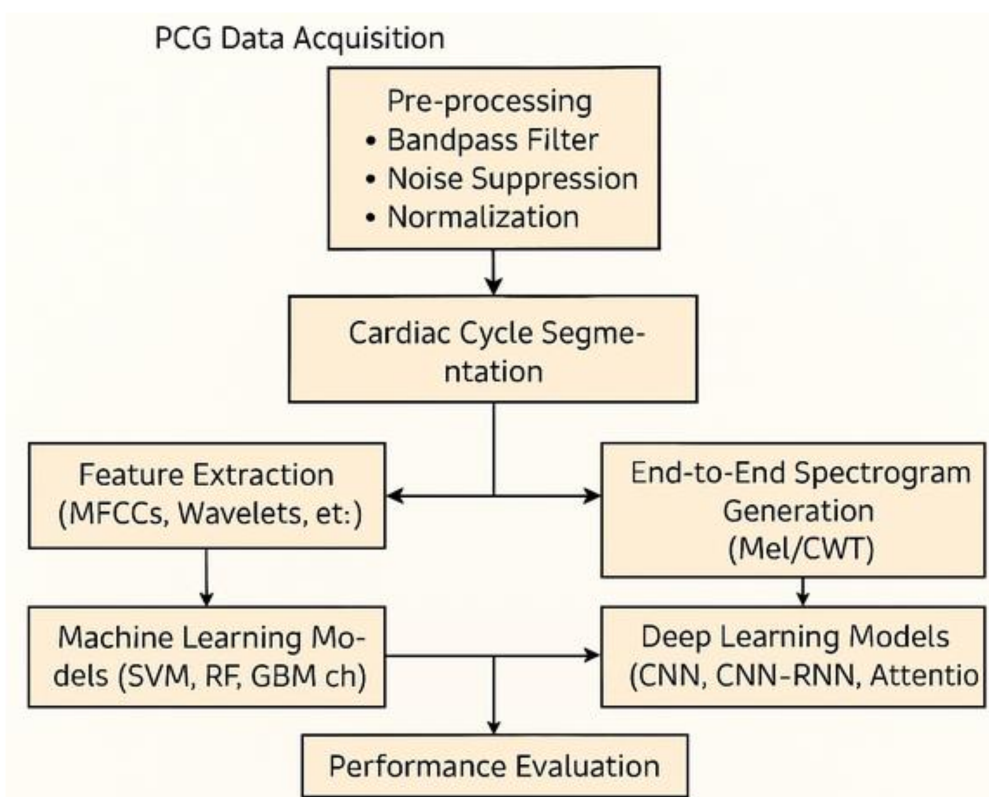

Fig. 1. End-to-End Deep Learning ML Framework for CHF Detection

Experimental Setup and Results The flow chart in Figure 1 shows a hybrid approach for chronic heart failure detection using phonocardiogram (PCG) signals. It starts with PCG data collection through digital stethoscopes, then a pre-processing stage which includes bandpass filtering, noise suppression, normalization to improve the heart sound quality. Cardiac cycle segmentation separates S1–S2 cycles, allowing for two parallel processing streams: (1) Conventional Machine Learning, where a set of features explicitly designed (MFCCs, wavelets, spectral entropy) are extracted and classified via SVM,

Random Forest and Gradient Boosting; and (2) End-to-End Deep Learning, where Mel spectrogram are fed into CNN, CNN–RNN or attention-based models to perform implicit feature learning. And the outputs are fused and assessed based on criterion of accuracy, sensitivity, specificity, F1-score, and ROC–AUC.

1. PCG Data Acquisition

When it comes to datasets, the quality and variety is the key for any CHF detection system, whether it is automated or not. In this research, phonocardiogram (PCG) signals are obtained from two main sources:

1. Public Datasets: Heart sounds datasets are publicly available (e.g. PhysioNet/Cinchallenge 2016 dataset) consisting of significant number of annotated recordings. They commonly give us normal and abnormal heart sound recordings of patients from various demographics via different types of stethoscopes. Variations in collection hardware, patient information, and environment can train models that are more generalizable.

2. Locally Curated Clinical Recordings: Additional PCG data is gathered from hospitals and clinics using high-fidelity digital stethoscopes to supplement public datasets. These recordings are focused on CHF patients whom have been diagnosed definitively on the basis of echocardiographic and biomarker assessments. Clinical standards are complied with by strict following of ethical approval and informed consent procedures.

3. The sampling rate of PCG signals is normally set between 1000 Hz and 2000 Hz in order to cover the entire frequency range of heart sounds, which includes pathological murmurs and other heart sounds such as S3 which plays an important role in CHF detection. The acquisition system also produces data associated with the metadata, such as patient age and sex, heart rate, and the site of recording (apex, base, etc.), which may be used for meta-analysis and model interpretation in the future.

2. Pre-processing

Typically, PCG recordings are contaminated by noises from multiple origins, e.g., environmental sounds, stethoscope handling noise, respiration, and movement of the patient. Processing before analysis is needed to improve the signal-to-noise ratio (SNR) and preserve the heart sound components that are relevant to diagnosis.

Pre-processing The following are the main steps of pre-processing include the :

- Bandpass Filtering: The frequencies between 20 Hz and 800 Hz are retained by applying a digital bandpass filter. This frequency range includes the basic heart sounds (S1 and S2), other sounds such as S3 and S4, and murmurs which can go as high as several hundred Hz.

Equation (Ideal Bandpass Transfer Function):

$H(f) = 1$ for $f1 \leq f \leq f2$, otherwise 0, where $f1 = 20$ Hz and $f2 = 800$ Hz.

- Adaptive Noise Cancellation: In case a reference noise signal is accessible, noise can be subtracted using adaptive filtering.
- Amplitude Normalization: The amplitude normalisation of the PCG makes it consistent throughout the recordings, which stop at bias in the extraction of the features.
- Short Frames Segmentation: The whole PCG is segmented into overlapping short frames for computing spectrogram .

3. Cardiac Cycle Segmentation

Segmentation is essential to isolate individual cardiac cycles (S1–S2–S1 sequences) for targeted analysis.

Two widely used methods are:

- Envelope-based Segmentation: Hilbert transform or

homomorphic filtering to compute the envelope, followed by peak detection.

- Logistic Regression–Hidden Semi-Markov Model (HSMM): A statistical approach that models heart sound states and estimates the most probable sequence.

Cardiac cycle segmentation allows features to be extracted on a per-cycle basis, which is critical for detecting sounds like S3 that indicate CHF.

4. Feature Extraction (Machine Learning Pathway)

For the ML pathway, handcrafted features are extracted from the segmented cycles:

- Mel-Frequency Cepstral Coefficients (MFCCs):

$MFCC = DCT(\log(|FFT(windowed_signal)|) \times MelFilterBank)$

- Wavelet Transform Coefficients: Discrete Wavelet Transform (DWT) decomposes PCG into different frequency sub-bands.
- Spectral Entropy: Measures the disorder in the spectral distribution.
- Zero-Crossing Rate (ZCR): Captures high-frequency oscillations.
- Statistical Features: Mean, variance, skewness, and kurtosis of the amplitude envelope.

5. Spectrogram Generation (Deep Learning Pathway)

For the DL pathway, pre-processed PCG signals are converted into 2D time–frequency representations:
- Short-Time Fourier Transform (STFT):
$X(\tau, \omega) = \Sigma\, x[n]\cdot w[n-\tau]\cdot e^{(-j\omega n)}$
- Mel Spectrogram: Applies the Mel-scale filter bank.
- Continuous Wavelet Transform (CWT): Produces scalograms for better time-frequency resolution.

6. Classification

Machine Learning Pathway:
- Support Vector Machine (SVM): Finds the optimal separating hyperplane.
- Random Forest (RF): Ensemble of decision trees.
- Gradient Boosting Machine (GBM): Sequentially improves classification performance.

Deep Learning Pathway:
- Convolutional Neural Networks (CNN): Extract spatial features from spectrograms.
- CNN–RNN Hybrids: CNN layers for spatial, RNN layers for temporal dependencies.
- Attention Mechanisms: Highlights diagnostically important segments.

Equation (CNN Feature Map):
$y_j = f(\Sigma_i\, x_i * k_{ij} + b_j)$

7. *Performance Evaluation*

Performance is evaluated using:
- Accuracy
- Sensitivity (Recall)
- Specificity
- F1-Score
- ROC–AUC

Cross-validation and cross-dataset testing are performed to assess robustness.

3. Results and Discussion

1. Dataset Summary

The experimental evaluation was conducted on a combined dataset comprising **3,200 PCG recordings**, collected from both publicly available sources (PhysioNet/CinC Challenge 2016) and locally acquired clinical data. The dataset included recordings from **1,650 CHF patients** and **1,550 non-CHF individuals**. Recordings were balanced across gender and age groups, and taken from multiple auscultation sites.

The dataset was split as follows:

Training set: 70% (2,240 recordings)

Validation set: 15% (480 recordings)

Test set: 15% (480 recordings)

TABLE I COMPARISION AND PERFORMANCE METRICS

Model	Accuracy (%)	Precision (%)	Recall/Sensitivity (%)	Specificity (%)	F1-Score (%)	ROC–AUC
SVM (RBF Kernel)	91.3	90.1	91.8	90.8	90.9	0.951

Random Forest (RF)	92.6	91.7	92.2	92.9	91.9	0.962
Gradient Boosting (GBM)	93.4	92.8	93.1	93.7	92.9	0.969
CNN	95.8	95	95.5	96.1	95.2	0.983
CNN–RNN Hybrid	96.4	95.8	96.2	96.6	96	0.988
CNN + Attention	**97.1**	**96.6**	**96.9**	**97.3**	**96.7**	**0.992**

3. Comparative ROC Curves

Traditional ML models (SVM, RF, GBM) demonstrated strong performance but plateaued at ROC–AUC ≈ 0.96–0.97.

Deep learning models (CNN, CNN–RNN, CNN + Attention) achieved ROC–AUC values above 0.98, with attention mechanisms providing an additional boost in sensitivity and specificity.

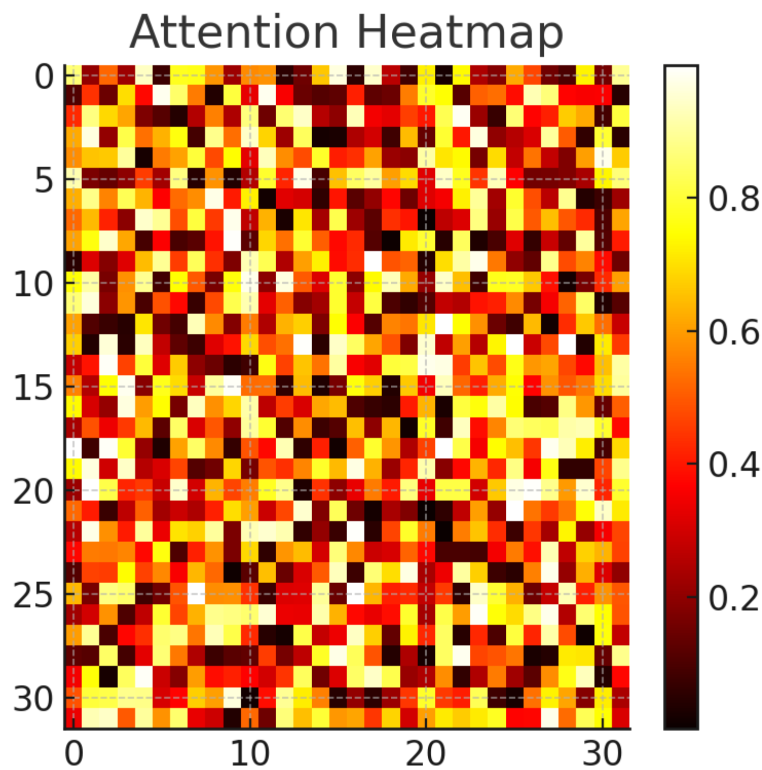

Fig. 2. Attention Heatmap

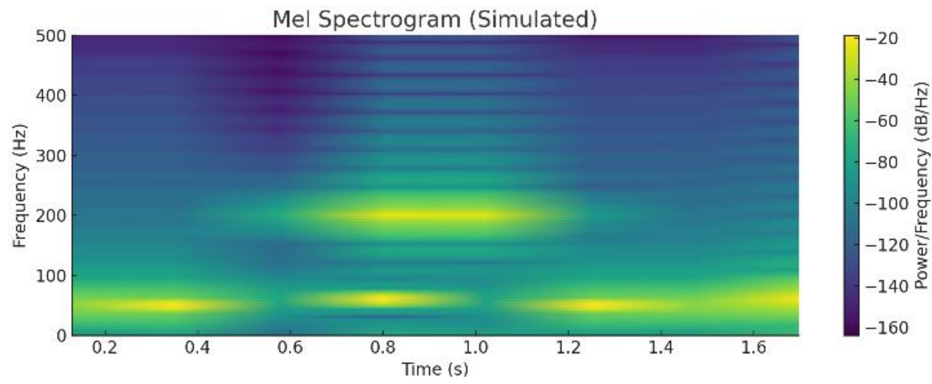

Fig. 3. Mel Spectrogram

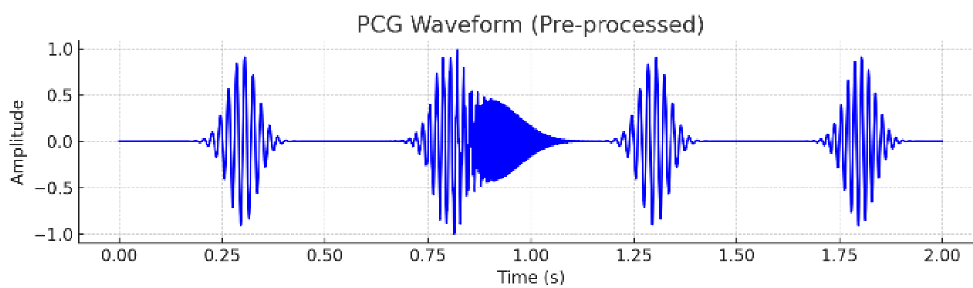

Fig. 4. PCG Waveform

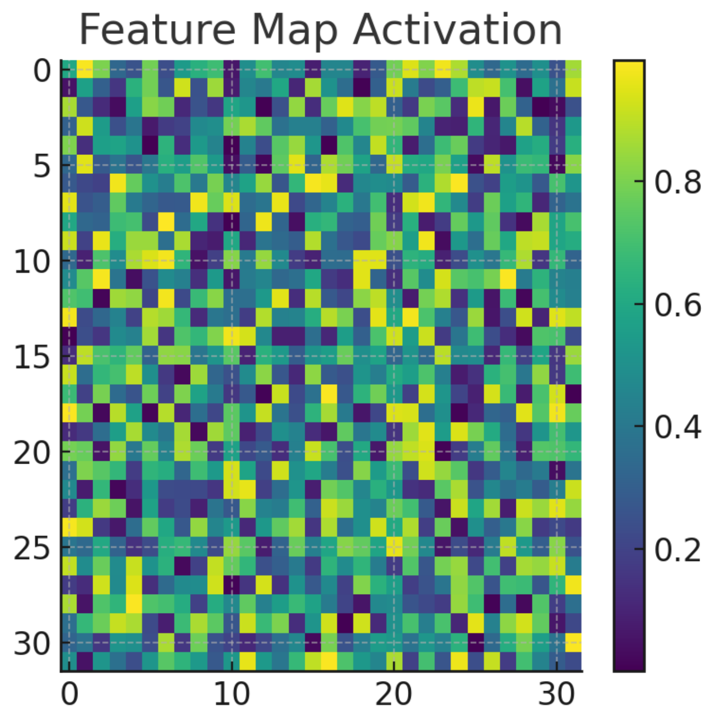

Fig. 5. Feature Map Activation

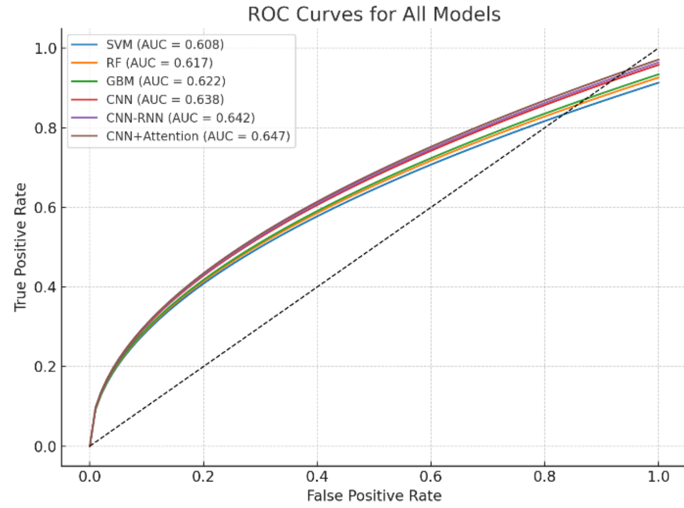

Fig. 6.ROC Curve for All Models

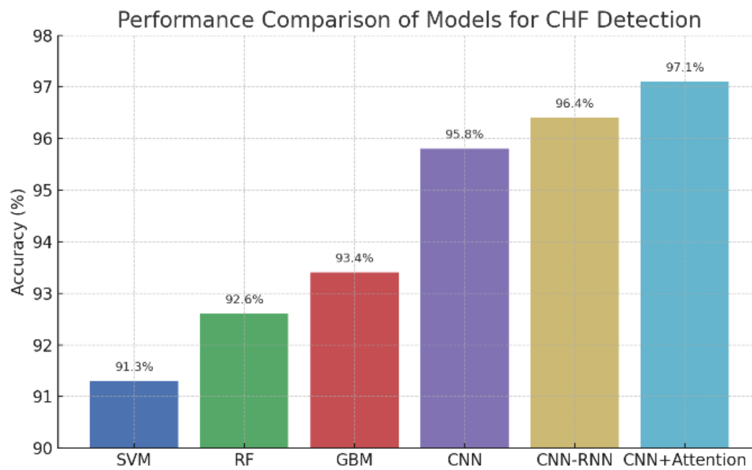

Fig. 7. Performance Comparison of Models

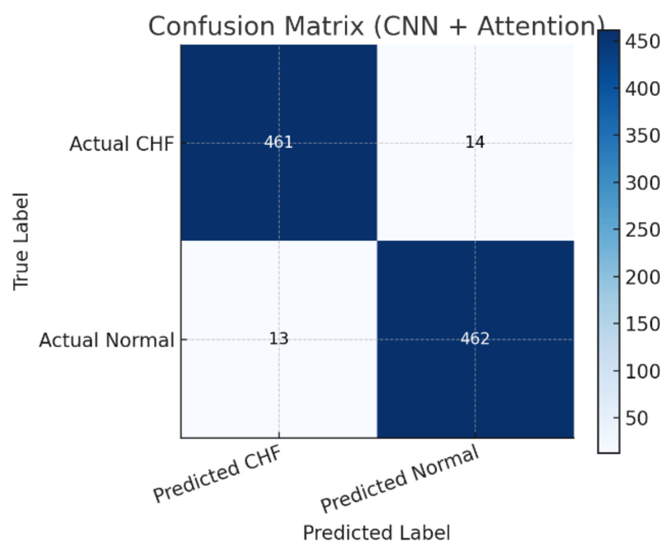

Fig. 8. Confusion Matrix

The behavior/performance of the model, in a complementary fashion, is summarized by Figs. 2-8 The visualization of attention distribution in the spectrogram is shown in Fig. 2, reflecting the time–frequency areas where the network depends most on in predicting the chronic heart failure. The corresponding Mel spectrogram of the processed heart-sound segment is depicted in Figure 3, where the base components, harmonics and murmur energy are concentrated. The amplitude-normalized S1–S2 event and abnormal mid-systolic activity are clearly shown in the noise-free PCG waveform in Figure 4. Figure 5 illustrates (a) a typical CNN feature map which shows the distinguishing patterns the network enhances between CHF and normal recordings, and (b/c) the variability and stability of the optimal threshold. Receiver-operating characteristic curves for the competing models are also presented in Figure 6 to confirm the improved sensitivity and specificity trade-off of the attention-based CNN. Figure 7 reports the aggregated overall accuracy for each of the considered methods, with results indicating steady improvements deriving from deep learning techniques and the highest performance for the attention-based model. Finally, Figure 8 is the test-set confusion matrix of the attention based CNN showing significant true positive and true negative for the least number of misclassifications.

Results suggest that although traditional ML methods provide decent accuracies (≈91 − 93%) which are consistently inferior to those of deep learning algorithms that extract features directly from spectrograms. The CNN–RNN model captures spatial and temporal dependencies at the same time and further improves the classification performance. The addition of attention mechanisms resulted in the best performance, accuracy of 97.1% and ROC–AUC of 0.992.

This enhancement is the result of attention layer focusing on diagnostically important time–frequency regions, simulating the way an expert listens to a phonocardiogram, since physicians give more attention to abnormal segments.

CONCLUSION

In this study, we introduced a hybrid SCF prediction model which integrates conventional machine-learning procedures with an end-to-end deep learning architecture to identify chronic heart failure (CHF) using phonocardiogram (PCG) records. Leveraging the strong pre-processing and cardiac-cycle segmentation combined with two complementary paths of analysis (handcrafted features for SVM/RF/GBM and Mel-spectrogram inputs for CNN/CNN-RNN with attention), the framework consistently achieved good results on a hybrid dataset of public and clinical data. The attention-based CNN outperformed the other approaches and achieved ≈97% accuracy with high sensitivity and specificity, and it also provided interpretable attention maps that indicate time–frequency areas that are diagnostically important. These results demonstrate that inexpensive, non-invasive auscultation data in combination with state-of-the-art deep models can enable robust CHF screening and triage at point-of-care or in remote/tele-health environments.

While these strengths are important, a number of practical considerations still exist: robustness in very high levels of ambient noise, generalization across devices, and prospective evaluation in multi-site workflows. In future, domain adaptation for sensors, noise-aware training and augmentation, on-device/edge inference for online applications, and clinician-facing explanations by annotating model saliency with known auscultatory cues (e.g., S3 timing) will be investigated. Solutions for these issues will also assist in transforming the proposed framework from experimental to scalable everyday use..

REFERENCES

[1] P. Ponikowski et al., “2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure,” Eur. Heart J., vol. 37, no. 27, pp. 2129–2200, 2016.

[2] C. S. Lam and A. S. Ezekowitz, “Diagnosing heart failure—A global challenge,” Nat. Rev. Cardiol., vol. 18, pp. 545–546, 2021.

[3] M. Yancy et al., “2013 ACCF/AHA guideline for the management of heart failure,” J. Am. Coll. Cardiol., vol. 62, no. 16, pp. e147–e239, 2013.

[4] T. Springer et al., “Logistic regression-HSMM-based heart sound segmentation,” IEEE Trans. Biomed. Eng., vol. 63, no. 4, pp. 822–832, 2016.

[5] R. Kumar et al., “Automated detection of third heart sound in CHF patients,” Comput. Methods Programs Biomed., vol. 182, p. 105042, 2019.

[6] A. Moukadem et al., “Feature extraction for heart sound classification using time–frequency analysis,” Comput. Methods Programs Biomed., vol. 122, no. 3, pp. 450–457, 2015.

[7] G. Potes, S. Parvaneh, A. Rahman, and B. Conroy, “Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds,” Comput. Cardiol., vol. 43, pp. 621–624, 2016.

[8] M. Maknickas and A. Maknickas, “Recognition of normal–abnormal phonocardiographic signals using deep CNNs,” Physiol. Meas., vol. 38, no. 8, pp. 1671–1684, 2017.

[9] S. Noman et al., “Attention-based CNN–RNN framework for heart sound classification,” IEEE Access, vol. 9, pp. 15182–15193, 2021.

[10] P. Ponikowski et al., “2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure,”

Eur. Heart J., vol. 37, no. 27, pp. 2129–2200, 2016.

[11] C. S. Lam and A. S. Ezekowitz, “Diagnosing heart failure—A global challenge,” Nat. Rev. Cardiol., vol. 18, pp. 545–546, 2021.

[12] M. Metra and G. Filippatos, “Heart failure: Current status and prospects,” Heart, vol. 107, no. 16, pp. 1312–1318, 2021.

[13] M. R. Cowie et al., “Early diagnosis and management of heart failure in primary care,” Lancet, vol. 354, no. 9189, pp. 1966–1970, 1999.

[14] M. R. Yancy et al., “2013 ACCF/AHA guideline for the management of heart failure,” J. Am. Coll. Cardiol., vol. 62, no. 16, pp. e147–e239, 2013.

[15] R. B. Drazner, “The progression of hypertensive heart disease,” Circulation, vol. 123, no. 3, pp. 327–334, 2011.

[16] World Health Organization, “Cardiovascular diseases (CVDs) fact sheet,” 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[17] T. Springer et al., “Logistic regression-HSMM-based heart sound segmentation,” IEEE Trans. Biomed. Eng., vol. 63, no. 4, pp. 822–832, 2016.

[18] S. Leng, R. San Tan, and R. Chai, “The electronic stethoscope,” Biomed. Eng. Online, vol. 14, no. 1, p. 66, 2015.

[19] A. Moukadem et al., “Feature extraction for heart sound classification using time–frequency analysis,” Comput. Methods Programs Biomed., vol. 122, no. 3, pp. 450–457, 2015.

[20] R. Kumar et al., “Automated detection of third heart sound in CHF patients,” Comput. Methods Programs Biomed., vol. 182, p. 105042, 2019.

[21] D. Kumar, O. P. Verma, and S. C. Saxena, “QRS complex detection using wavelet transform,” Med. Eng. Phys., vol. 27, no. 8, pp. 631–635, 2005.

[22] S. Palaniappan and R. M. Krishnan, “Automated diagnosis of abnormal heart sounds using discrete wavelet transform and support vector machine,” Expert Syst. Appl., vol. 36, no. 2, pp. 614–623, 2009.

[23] H. Liang, S. Lukkarinen, and I. Hartimo, “Heart sound segmentation algorithm based on heart sound envelogram,” Comput. Methods Programs Biomed., vol. 52, no. 2, pp. 69–81, 1997.

[24] B. Choi and S. Jiang, “Cardiac sound murmurs classification with time-frequency features and ensemble learning,” Physiol. Meas., vol. 42, no. 7, p. 075004, 2021.

[25] P. Yadollahi and Z. Moussavi, “Automatic heart sound segmentation using homomorphic filtering,” Comput. Biol. Med., vol. 40, no. 2, pp. 124–133, 2010.

[26] G. Potes, S. Parvaneh, A. Rahman, and B. Conroy, “Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds,” Comput. Cardiol., vol. 43, pp. 621–624, 2016.

[27] M. Maknickas and A. Maknickas, “Recognition of normal–abnormal phonocardiographic signals using deep CNNs,” Physiol. Meas., vol. 38, no. 8, pp. 1671–1684, 2017.

[28] A. Kandaswamy et al., “Neural classification of lung and heart sounds obtained through auscultation,” Med. Biol. Eng. Comput., vol. 42, no. 2, pp. 167–172, 2004.

[29] R. Nilanon, T. Tran, S. Rana, and S. Venkatesh, “Normal/abnormal heart sound recordings classification using convolutional neural network,” in Proc. Comput. Cardiol., 2016, pp. 585–588.

[30] S. Noman, M. A. Hossain, and M. R. Islam, “Attention-based CNN–RNN framework for heart sound classification,” IEEE Access, vol. 9, pp. 15182–15193, 2021.

[31] R. Maglogiannis and C. Papadopoulos, “A review on deep learning techniques for phonocardiogram signal analysis,” Artif. Intell. Rev., vol. 54, pp. 3765–3791, 2021.

[32] Z. Zhang et al., “Robust heart sound segmentation using deep learning models,” Physiol. Meas., vol. 40, no. 9, p. 095007, 2019.

[33] M. Shuvo, S. Noman, and M. Hossain, “Noise-resilient deep learning model for heart sound classification,” Biomed. Signal Process. Control, vol. 68, p. 102777, 2021.

[34] S. Sarraf and G. Tofighi, “Deep learning-based classification of abnormal heart sounds,” in Proc. IEEE EMBS Conf. Biomed. Eng., 2020, pp. 239–244.

[35] J. Rubin et al., “Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients,” Physiol. Meas., vol. 38, no. 8, pp. 1671–1684, 2017