Detection and Classification of Electrocardiogram Signals to Identify Congestive Heart Failure Based on Machine Learning Techniques and Grasshopper Optimization Algorithm

Naser Safdarian[1], Parisa Eghbal Kiani[1], Nader Jafarnia Dabanloo[1], Saman Parvaneh[2]

[1] Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[2] Edwards Lifesciences, Irvine, California, USA

Abstract

Congestive Heart Failure (CHF) is a life-threatening cardiovascular condition that requires early detection to enhance patient outcomes. This study introduces a hybrid diagnostic system for the automated classification of ECG signals into CHF, Normal Sinus Rhythm (NSR), and Arrhythmia (ARR). To evaluate the performance of the proposed method, 162 ECG signals for cases with ARR, CHF, and NSR were used. Then we use filtering to remove noises, followed by R-peak detection using the Pan-Tompkins algorithm. From the cleaned signals, two key physiological time series are extracted: the Heart Rate Time Series (HRTS) and ECG-Derived Respiration (EDR). A novel feature extraction approach is employed where both the HRTS and EDR signals are transformed into two-dimensional (2D) images using Recurrence Plots (RP), which capture complex nonlinear dynamics. The feature maps from the CNN are then concatenated into machine learning classifiers for the final classification. The performance of Random Forest and SVM classifiers was evaluated. To maximize accuracy, the SVM kernel parameters were fine-tuned using the Grasshopper Optimization Algorithm, a metaheuristic approach. This integrated system, combining advanced signal processing, deep feature learning via CNNs, and metaheuristic optimization, demonstrates high accuracy (>95%) and provides a robust framework for cardiac diagnostics.

1. Introduction

Congestive Heart Failure (CHF) occurs when the heart is unable to adequately pump blood throughout the body without increasing the pressure inside the cardiac chambers. It represents one of the most serious manifestations of cardiovascular disease, contributing substantially to global morbidity and mortality rates [1]. According to the European Society of Cardiology (ESC) guidelines, heart failure can result from structural or functional abnormalities that impair ventricular filling or ejection of blood, leading to insufficient tissue perfusion and fluid accumulation [2]. Early and accurate diagnosis of CHF is therefore crucial for improving clinical outcomes and preventing disease progression. Traditional diagnostic methods rely heavily on clinical assessment, echocardiography, and ECG interpretation, but these approaches often require expert supervision and are prone to inter-observer variability [3].

The ECG signal, as a non-invasive and cost-effective diagnostic tool, plays a vital role in detecting cardiac abnormalities. Recent advances in machine learning and signal processing have made it possible to automatically analyze ECG signals to distinguish between Normal Sinus Rhythm (NSR), Cardiac Arrhythmia (ARR), and CHF [3,4]. Heart failure has been linked to autonomic nervous system dysfunction, reflected in altered heart rate dynamics and irregular ECG patterns [5,6]. Moreover, Heart Rate Time Series (HRTS) analysis and ECG-derived respiration (EDR) techniques have been widely explored to quantify these physiological changes [7,8]. To effectively capture such nonlinear and nonstationary dynamics, modern signal processing frameworks, such as recurrence plots and time–frequency analysis, have been increasingly utilized [9].

In recent years, machine learning and deep learning techniques have demonstrated remarkable performance in biomedical signal classification, offering robust feature extraction and improved generalization across patient populations [10, 11]. However, optimizing the parameters of such algorithms remains a challenging task that significantly influences classification accuracy. Therefore, this study aims to design a fully automated diagnostic framework for detecting and classifying ECG signals to identify

 ISSN: 2325-887X DOI: 10.22489/CinC.2025.338

CHF from other cardiac conditions using machine learning algorithms optimized by the Grasshopper Optimization Algorithm (GOA). The proposed approach aims to enhance diagnostic reliability and computational efficiency, ultimately contributing to the development of intelligent, data-driven systems for early detection of CHF.

2. Proposed Method

The proposed framework aims to automatically detect and classify CHF using ECG signals by integrating signal processing, feature extraction, and optimization-based machine learning techniques. The block diagram of the proposed system is shown in Figure 1, which illustrates the main steps: (1) signal preprocessing, (2) feature extraction, and (3) classification and optimization.

2.1. Datasets

To evaluate the effectiveness and performance of the proposed method, 162 recordings of ECG signals for cases with ARR, CHF, and NSR from three open-access databases (MIT-BIH ARR, MIT-BIH NSR, and BIDMC for CHF data) were used [12–14]. Among the 162 ECG recordings, 96 recordings are related to the ARR group, 30 recordings are associated with CHF, and 36 recordings are linked to NSR. All ECG recordings were measured from leads II and VI, then analyzed and labeled by cardiologists. Considering the differences in sampling frequency among the three groups, all signals were resampled at a rate of 128 Hz.

2.2. Signal Processing

Signal preprocessing plays a crucial role in ensuring that the ECG signals are clean, consistent, and suitable for subsequent feature extraction and classification stages. Raw ECG signals are often contaminated with various types of noise, including baseline wander, power-line interference, and muscle artifacts, which can significantly affect diagnostic accuracy. To address these issues, several preprocessing operations were applied sequentially.

First, all ECG recordings were resampled to a uniform sampling frequency to ensure temporal consistency across all recordings. Then, baseline wander removal was performed to eliminate low-frequency components (typically below 0.5 Hz) caused by respiration and patient movement. This was achieved using a high-pass filter. To further suppress high-frequency noise, a low-pass filter was employed to remove unwanted components above 40 Hz without distorting the QRS complex.

After filtering, R-peak detection was performed using the Pan–Tompkins algorithm, a widely accepted and robust technique for identifying the QRS complex. The algorithm combines digital band-pass filtering, differentiation, squaring, and adaptive thresholding to precisely determine the R-wave positions in the ECG signal. Accurate detection of R-peaks is crucial, as it provides the foundation for the subsequent computation of inter-beat intervals, which are essential for creating HRTS and EDR.

2.3. Feature Extraction

Feature extraction is the core step in transforming the one-dimensional ECG signals into meaningful two-dimensional representations suitable for pattern recognition and classification. In this work, two types of physiological signals were extracted from the preprocessed ECG data: HRTS and EDR.

HRTS analysis quantifies the temporal variation between consecutive R-peaks, which reflects the influence of the autonomic nervous system on cardiac function. HRTS is a well-established indicator for assessing cardiovascular health, particularly in CHF patients, who typically exhibit reduced variability due to impaired autonomic regulation. EDR analysis was also conducted to capture the respiratory modulation present in the ECG signal. The EDR signal was estimated using amplitude and frequency modulation patterns of the ECG waveform, which indirectly reflect the breathing cycle. This measure is valuable because cardiorespiratory coupling is often altered in patients with heart failure.

Once HRTS and EDR signals were extracted, both were transformed into two-dimensional images using the Recurrence Plot (RP) technique. Recurrence plots are nonlinear dynamical system tools that visualize the recurrence of specific states within a time series. The resulting recurrence plot encodes temporal dependencies and hidden chaotic patterns within the physiological signal as a two-dimensional matrix.

These 2D recurrence plots, corresponding to HRTS and EDR signals, were then used as input to a two-dimensional Convolutional Neural Network (2D-CNN) to extract high-level spatial features. A CNN architecture, which consists of multiple convolutional layers followed by pooling and fully connected layers, enables hierarchical representation learning. The use of two parallel inputs allows the model to capture both cardiac rhythm irregularities via HRTS and respiratory coupling dynamics via EDR, thus improving the discriminative power for identifying CHF patients from those with NSR and ARR.

2.4. Classification and Optimization

The two feature maps generated from HRTS and EDR recurrence plots using 2D-CNN were concatenated to form a unified feature vector. For the classification task, two machine learning algorithms were evaluated: Random For-

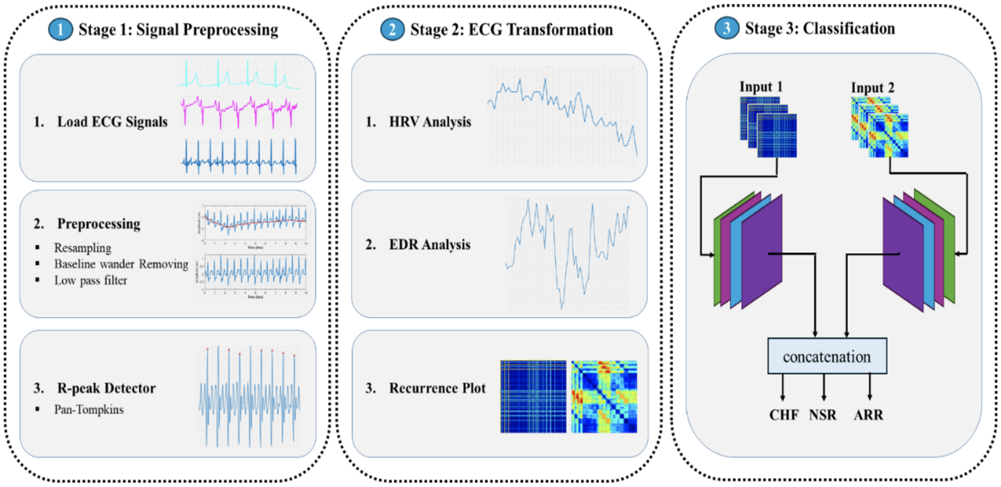

Figure 1. The block diagram of the proposed method.

est (RF) and Support Vector Machine (SVM). The SVM is particularly effective for high-dimensional data, as it constructs an optimal hyperplane that maximizes the margin between different classes. The RF classifier, on the other hand, is an ensemble method that combines multiple decision trees to achieve better generalization and robustness.

To enhance the performance of the SVM classifier, the parameters of its kernel functions were optimized using the GOA. The GOA is a population-based metaheuristic inspired by the swarming and foraging behavior of grasshoppers in nature. It balances exploration and exploitation by modeling the attraction and repulsion forces between individuals in the search space. In this study, three grasshoppers were initialized as search agents, and optimization was performed over three kernel types: linear, Radial Basis Function (RBF), and polynomial kernels. The objective function aimed to minimize the classification error while optimizing kernel parameters such as the penalty term (C) and kernel width (γ).

The optimized SVM parameters were then used to classify ECG signals into three distinct classes: CHF, NSR, and ARR.

3. Results

The proposed system was implemented and tested to evaluate its performance in detecting and classifying CHF from ECG signals. The data were divided into training and testing sets in an 80:20 ratio to ensure reliable evaluation. For performance assessment, standard metrics, including accuracy, sensitivity, and specificity, were calculated.

Classification accuracies of 94.25% and 95.8% were achieved using an SVM classifier without optimization and RF, respectively. The proposed model, utilizing an optimized SVM with GOA, achieved an overall classification accuracy of 98.35%, demonstrating its strong capability in distinguishing CHF from NSR and ARR signals. The sensitivity and specificity values were found to be 97.8% and 98.7%, respectively, indicating that the system not only correctly identifies CHF patients but also effectively minimizes false alarms. The confusion matrix revealed that most misclassifications occurred between the ARR and CHF classes, which is expected due to overlapping waveform characteristics in certain pathological cases.

The experimental results demonstrate that the proposed hybrid method, which combines recurrence plot analysis, 2D-CNN-based feature extraction, and SVM classifier optimized by GOA, achieves superior accuracy and robustness compared to conventional approaches.

4. Discussion and Conclusion

The integration of recurrence plot-based features, CNN-based deep representations, and GOA-based parameter optimization for the SVM classifier yielded a hybrid diagnostic system that achieved high classification accuracy, robustness to noise, and efficient computation. A comparative analysis with other state-of-the-art methods, such as conventional SVM, RF, and CNN-only approaches, reveals that integrating Recurrence Plot (RP)-based features and GOA optimization significantly enhances model performance. For instance, a baseline SVM classifier without optimization achieved an accuracy of approximately 94.2%, while the GOA-optimized SVM improved this accuracy by over 4%. Similarly, the RF classifier achieved an accuracy of around 95.8%, highlighting that parameter optimization and nonlinear feature representation were key factors in the superior performance of the proposed method.

The implementation of this algorithm for automated CHF diagnosis demonstrates the practical potential of machine learning in clinical decision support systems. By transforming ECG signals into recurrence-based visual patterns and employing optimized classifiers, the system effectively captures complex temporal and nonlinear dynamics that are often overlooked by traditional signal processing methods. Moreover, the high accuracy and low computational cost make the proposed approach suitable for real-time or embedded applications in clinical and remote monitoring environments.

References

[1] The WHO CVD Risk Chart Working Group. World health organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health 2019;7(10):e1332–e1345.

[2] Ponikowski P, Voors A, Anker S, Bueno H, Cleland J, Coats A, Falk V, González-Juanatey J, Harjola V, Jankowska E, Jessup M. 2016 esc guidelines for the diagnosis and treatment of acute and chronic heart failure. Kardiologia Polska Polish Heart Journal 2016;74(10):1037–1147.

[3] Gacek A. An introduction to ecg signal processing and analysis. In ECG Signal Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence. London: Springer London, 2011; 21–46.

[4] Faezipour M, Saeed A, Bulusu S, Nourani M, Minn H, Tamil L. A patient-adaptive profiling scheme for ecg beat classification. IEEE Transactions on Information Technology in Biomedicine June 2010;14(5):1153–1165.

[5] Kishi T. Heart failure as an autonomic nervous system dysfunction. Journal of Cardiology 2012;59(2):117–122.

[6] Moharreri S, Rezaei S, Parvaneh S. Using heart rate fragmentation and heart rate asymmetry to discriminate congestive heart failure. Computers in Cardiology 2024;51(1):1–4.

[7] ChuDuc H, NguyenPhan K, NguyenViet D. A review of heart rate variability and its applications. APCBEE Procedia 2013;7:80–85.

[8] Moody G, Mark R, Bump M, Weinstein J, Berman A, Mietus J, Goldberger A. Clinical validation of the ecg-derived respiration (edr) technique. In Computers in Cardiology, volume 13. October 1986; 507–510.

[9] Eckmann J, Kamphorst S, Ruelle D. Recurrence plots of dynamical systems. World Scientific Series on Nonlinear Science Series A September 1995;16:441–446.

[10] Lee I, Kim D, Kang S, Lee S. Ensemble deep learning for skeleton-based action recognition using temporal sliding lstm networks. In Proceedings of the IEEE International Conference on Computer Vision. 2017; 1012–1020.

[11] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521(7553):436–444.

[12] Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, Mietus J, Moody G, Peng C, Stanley H. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation June 2000;101(23):e215–e220.

[13] Moody G, Mark R. The impact of the mit-bih arrhythmia database. IEEE Engineering in Medicine and Biology Magazine May 2001;20(3):45–50.

[14] Baim D, Colucci W, Monrad E, Smith H, Wright R, Lanoue A, Gauthier D, Ransil B, Grossman W, Braunwald E. Survival of patients with severe congestive heart failure treated with oral milrinone. Journal of the American College of Cardiology March 1986;7(3):661–670.

Address for correspondence:

Naser Safdarian
Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Daneshgah Square, Sattari Avenue, Tehran, Iran
E-mail address: naser.safdarian@iau.ac.ir