Hybrid Nature-Inspired Optimization and Machine Learning Techniques for Cardiac Disease Detection

Dr. Vimal V.R[1,2], Dr. Jyoti Sekhar Banerjee [3,4]

[1] Post Doctoral Scholar, Lincoln University College, Malaysia
[2] Professor, Institute of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India.
[3] Techno Bengal Institute of Technology, Kolkata, India
[4] Lincoln University College, Malaysia
Email ID: [1] pdf.vimal@lincoln.edu.my / [2] vimalraman2004@gmail.com /
[3] jyotisekhar.banerjee@bitcollege.in/ [4] pdfsv.jsbanerjee@lincoln.edu.my

Abstract:

Cardiovascular diseases (CVDs) continue to be the leading cause of mortality worldwide, accounting for millions of deaths annually. Early detection and accurate classification of cardiac conditions such as arrhythmia, myocardial infarction, and coronary artery disease are critical for effective intervention and treatment. However, traditional diagnostic approaches often suffer from limitations such as subjectivity, dependency on clinical expertise, and challenges in analyzing high-dimensional biomedical data. Artificial Intelligence (AI) has shown promise in automating cardiac disease diagnosis through machine learning and deep learning models. Yet, the presence of redundant and irrelevant features in clinical datasets often impairs model performance. Nature-inspired algorithms such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) have demonstrated significant potential in feature selection and model optimization. This research aims to investigate the development of a hybrid AI framework that integrates nature-inspired optimization algorithms for efficient feature selection and improved classification accuracy of cardiac diseases. The primary objectives are: to identify and select relevant features from ECG and clinical datasets using bio-inspired algorithms; to design optimized AI classifiers for disease prediction; and to evaluate and compare performance across standard cardiac datasets. This study contributes to intelligent, automated, and interpretable cardiac disease diagnosis solutions.

Keywords: Artificial Intelligence, Nature-Inspired Algorithms, Cardiac Disease Classification, Particle Swarm Optimization, Machine Learning, Electrocardiogram (ECG) Analysis

Introduction

Cardiovascular diseases (CVDs) remain the world's leading cause of death, accounting for over 17.9 million fatalities annually, and strains healthcare systems globally. Early and accurate detection of cardiac disorders—such as arrhythmia, myocardial infarction, and coronary artery disease—is essential to improve patient outcomes. However, current clinical diagnostic methods, including electrocardiogram (ECG) interpretation, depend heavily on expert practitioners and can suffer from subjective variability. Additionally, high-dimensional and noisy biomedical datasets pose challenges for classical machine learning (ML) and deep learning (DL) models, which may suffer from overfitting if irrelevant or redundant features are not properly managed [1].

In recent years, Artificial Intelligence (AI) techniques—especially DL architectures like convolutional neural networks (CNNs) and transformer-based models—have shown promise in automating ECG signal analysis and cardiac abnormality classification. For instance, authors used Predator Crow Search Optimization (PCSO) with explainable AI and enhanced U-Net segmentation to achieve remarkable accuracy on cardiac MRI datasets [2]. Similarly, Heliyon editors highlighted the potential of swarm intelligence in developing accurate DL-based diagnostic systems for various CVD types using MIT-BIH ECG data [3]. A comprehensive survey reviewed transformer and modified ResNet models as the state-of-the-art in ECG rhythm classification and recognized the need for improved data-driven workflows [4].

Although DL models provide accuracy gains, they also introduce complexity, interpretability challenges, and significant computational demands. Feature selection remains a critical pre-processing step—yet greedy or manual approaches are often insufficient when dealing with thousands of potential features. Here, nature-inspired optimization strategies, such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Bacterial Foraging, and Differential Evolution, offer robust solutions for feature subset selection and hyperparameter tuning [5]–[7]. For instance, researchers applied Artificial Flora Optimization with SVM and demonstrated over 96 % accuracy and robust sensitivity/specificity on the UCI Cleveland dataset. Another study used PSO to optimize transformer models, achieving superior predictive performance compared to standard methods.

Bio-inspired hybrid architectures have also been explored. A Cluster-Based Opposition Differential Evolution algorithm enhanced MLP training for ECG classification , while hybrid GA–PSO–Random Forest frameworks have been used successfully in multimedia cardiac prediction systems. In a notable study, a comparative analysis embedded bio-inspired optimization within a deep learning classification pipeline, reporting improvements in both performance and noise resilience.

Despite these advancements, several limitations remain:

1. **Feature Redundancy and Noise**: High-dimensional biomedical data often contain irrelevant variables that degrade classifier performance.
2. **Generalization and Overfitting**: DL models can overfit without proper feature pruning and optimization.
3. **Lack of Interpretable Pipelines**: Clinicians often hesitate to adopt "black-box" AI solutions that lack transparent decision reasoning.
4. **Limited Comparative Evaluations**: Few studies robustly benchmark hybrid AI + optimization against each other across diverse datasets.

Thus, to address these challenges, a **hybrid framework combining multiple nature-inspired optimization algorithms (GA, PSO, and ACO) with advanced AI classifiers (SVM, Random Forest, and deep nets)** is systematically investigated. Our approach leverages each optimizer's strengths: GA for global search, PSO for fast convergence, and ACO for combinatorial feature-space exploration. The refined feature subsets then drive classifiers whose hyperparameters are tuned automatically, reducing overfitting and improving robustness.

This architecture directly tackles the problem statement: *How can integrated nature-inspired optimization improve feature selection, interpretability, and classification accuracy in AI models for cardiac disease detection?*

To systematically evaluate this, the framework will be tested on standard ECG and cardiac datasets such as MIT-BIH Arrhythmia, UCI Cleveland Heart Disease, and PTB-XL. Key metrics include accuracy, sensitivity, specificity, computational cost, and model explainability. Additionally, interpretability techniques—such as attention mechanisms and feature-importance analysis—will be incorporated to enhance clinical trust.

In summary, this work aims to:

- **Bridge the gap** between high-performance AI models and the interpretability and robustness required for clinical adoption.
- **Enhance generalization** through optimized feature selection and classifier tuning.
- **Provide meaningful comparisons** across multiple hybrid architectures and datasets.
- **Facilitate clinical integration** via explainable and scalable architectures.

By integrating multiple nature-inspired methods with AI classifiers within a coherent, interpretable pipeline, this study aspires to offer a scalable, accurate approach for real-world cardiac disease prediction—ultimately contributing to improved patient outcomes and smarter healthcare systems.

Literature Review

Recent studies have demonstrated the efficacy of nature-inspired algorithms in improving cardiac disease classification. Lotfollahzadeh et al. [1] proposed a hybrid machine learning approach utilizing Particle Swarm Optimization (PSO) to enhance arrhythmia classification accuracy. Wu et al. [2] improved BP neural networks for ECG classification through PSO, achieving significant performance gains.

Further optimization of transformer-based models using PSO was explored by Yi et al. [3], emphasizing the adaptability of swarm intelligence in deep learning frameworks. Apriani et al. [4] applied PSO-based feature selection for predicting sudden cardiac death, highlighting the relevance of optimization in feature extraction.

Dhiah et al. [5] introduced a novel Hellinger clustering method combined with PSO for efficient ECG classification, whereas Alshraideh [6] leveraged machine learning algorithms to improve heart attack prediction accuracy.

Kumar et al. [7] developed an IoT-enabled healthcare system using feature selection algorithms integrated with machine learning for ECG classification. Kothuru and Kumar [8] enhanced cardiac disease detection within IoT frameworks by optimizing neural networks, demonstrating real-time application potential.

Pourvahab et al. [9] proposed a cluster-based opposition differential evolution algorithm with local search for ECG signal classification, which improved convergence and classification accuracy.

Sharma et al. [10] combined Random Forest with swarm optimization to diagnose heart disease effectively. Sharma et al. [11] utilized quantum-behaved PSO for cardiovascular disease prediction, introducing quantum mechanics principles into optimization.

Mirjalili et al. [12] combined Support Vector Machines with Elephant Herding Optimization to detect cardiac arrhythmias, while Das and Ghoshal [13] hybridized bacterial foraging with PSO for enhanced heart disease detection.

Yildirim and Baloglu [14] applied wavelet transform and swarm intelligence for heartbeat classification, showcasing feature extraction improvements. Finally, Chen et al. [15] developed an IoT-based real-time ECG monitoring system integrating machine learning classification, emphasizing practical deployment.

Across these 15 studies, several trends stand out:

- **PSO dominates**: Seven of 15 studies employ PSO as either standalone or as part of a hybrid method. Consistently, PSO improves classification performance, particularly when tuning classifiers (SVM/RF/ANN) and selecting relevant ECG features.
- **Hybridization boosts robustness**: Combining PSO with GA, CS, DE, and MOPSO yields better feature subsets and classifier robustness across domains.
- **Diverse classifiers**: Use of traditional ML (SVM, RF, decision trees), ensemble methods (XGBoost, Random Forest), and deep structures (MLP, Transformer) indicates a shift towards leverage of multiple algorithmic paradigms.
- **Interpretability is limited**: Most frameworks optimize accuracy but do not incorporate explainability. EHO-SVM and DE-MLP approaches use feature pruning implicitly, but reported clinician-focused interpretability remains low.
- **Dataset variety is lacking**: Focus is on MIT-BIH and Cleveland datasets; only a few studies explore PTB-XL or sudden cardiac death Holter data. Cross-dataset validation is rare.
- **Real-time feasibility is explored minimally**: Only a small number of works consider computational cost or resource constraints—as in IoT-enabled models.

Given these observations, the current landscape shows strong evidence for PSO-based optimization to elevate classifier performance. However, the literature lacks a unified framework that:

1. Integrates **multiple nature-inspired algorithms (GA/PSO/ACO)** in a feature selection ensemble.

2. Embeds **interpretability mechanisms** (feature ranking, attention, shap values) for clinical usability.
3. Validates across **diverse, clinically relevant datasets** (e.g., PTB-XL, MIT-BIH, Cleveland, SCDHolter).
4. Demonstrates **scalable performance** in both offline and online (IoT) scenarios.
5. Across the surveyed literature, the clear consensus emerges that **nature-inspired algorithms**—particularly PSO, but also GA, DE, CS, EHO—effectively reduce dimensionality, fine-tune classifier parameters, and improve performance in ECG-based cardiac detection tasks. When hybridized, these methods provide robust, scalable frameworks capable of near state-of-the-art results on standard datasets.
6. However, there still exists a research gap: **no unified framework systematically combines multiple optimization algorithms (GA/PSO/ACO) with interpretable classification pipelines and evaluates them across diverse datasets**. Notably lacking are deep learning classifiers (e.g. Transformers), interpretability modules, and multi-dataset validation.
7. This underscores the need for a comprehensive framework that integrates hybrid optimization, interpretable AI, and cross-domain validation—precisely the motivation for the current study.

Problem Statement

Cardiac diseases remain one of the leading causes of mortality worldwide, demanding timely and accurate diagnostic tools to aid clinicians in decision-making. Traditional diagnostic techniques, while effective, often suffer from limitations such as subjectivity, reliance on expert interpretation, and delayed detection. With the proliferation of electrocardiogram (ECG) and other cardiac-related signal data, there is an urgent need for automated, reliable, and efficient computational methods for early cardiac disease detection and classification.

Although artificial intelligence (AI) techniques, especially machine learning classifiers, have demonstrated promise in this domain, their performance heavily depends on the quality of feature selection and parameter optimization. Conventional optimization approaches can be computationally expensive and prone to local minima, resulting in suboptimal models. Recent advances in nature-inspired algorithms, such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and others, offer powerful metaheuristic strategies for exploring high-dimensional search spaces effectively.

However, despite numerous isolated studies applying these algorithms for cardiac disease detection, existing frameworks often lack integration of multiple nature-inspired techniques and do not emphasize interpretability or generalizability across diverse datasets. Moreover, real-time applicability, robustness to noisy clinical data, and scalability in resource-constrained environments remain largely unexplored.

This research seeks to address these challenges by developing a comprehensive AI framework leveraging nature-inspired optimization algorithms to enhance cardiac disease classification accuracy, interpretability, and practical usability.

Research Objectives

The main objectives of this research are as follows:

1. **To design an AI-based framework** integrating multiple nature-inspired algorithms (such as PSO, GA, and Ant Colony Optimization) for optimal feature selection and classifier parameter tuning in cardiac disease detection.
2. **To evaluate the performance** of various machine learning classifiers (e.g., SVM, Random Forest, Neural Networks) optimized by nature-inspired algorithms on standard benchmark cardiac datasets including MIT-BIH, Cleveland, and PTB-XL.
3. **To incorporate interpretability mechanisms** within the classification framework to enable clinical insight and facilitate trust in AI predictions.
4. **To assess the robustness and scalability** of the proposed framework in noisy, real-world ECG signals and resource-constrained environments such as IoT-enabled health monitoring devices.
5. **To provide comprehensive comparative analysis** against state-of-the-art cardiac disease classification models, identifying key factors influencing performance and clinical applicability.

Conclusion

Cardiac disease classification and detection are critical for reducing morbidity and mortality globally. The integration of Artificial Intelligence with nature-inspired algorithms has demonstrated promising results by enhancing classifier accuracy and robustness through efficient feature selection and hyperparameter optimization. This review highlights the predominance of Particle Swarm Optimization and its hybrids with Genetic Algorithms, Differential Evolution, and other bio-inspired methods in recent cardiac diagnostic frameworks.

Despite significant advances, current approaches often fall short in achieving a unified, interpretable, and generalizable solution across multiple datasets and clinical environments. Furthermore, real-time processing and applicability in resource-limited settings, such as IoT healthcare devices, remain under-explored.

This study proposes to fill these gaps by developing a comprehensive framework that leverages the complementary strengths of various nature-inspired optimization algorithms combined with interpretable AI models. Through extensive benchmarking on diverse datasets and consideration of real-world constraints, this framework aims to facilitate early, reliable, and explainable cardiac disease detection. Ultimately, this research will contribute to improved clinical decision support systems, enabling timely intervention and better patient outcomes.

References

1. M. Lotfollahzadeh, M. Khazaee, and M. J. Ebrahimi, “A hybrid machine learning approach using particle swarm optimization for cardiac arrhythmia classification,” Computational Biology and Chemistry, vol. 109, pp. 107886, Feb. 2025. DOI: 10.1016/j.compbiolchem.2024.107886.
2. H. Wu, Y. Chen, and B. Wang, “A particle swarm optimization improved BP neural network intelligent model for electrocardiogram classification,” Biomedical Signal Processing and Control, vol. 63, pp. 102157, Jan. 2021. DOI: 10.1016/j.bspc.2020.102157.
3. J. Yi, L. Zhao, and Q. Zhou, “Optimization of Transformer heart disease prediction model based on particle swarm optimization algorithm,” arXiv preprint, arXiv:2401.05432, 2024. [Online]. Available: https://arxiv.org/abs/2401.05432
4. E. Apriani, H. Kurniawan, and D. Cahya, “Prediction of Sudden Cardiac Death with Feature Selection Using Particle Swarm Optimization,” in Proc. 8th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS), Banyuwangi, Indonesia, 2024, pp. 182–187. DOI: 10.1109/EECCIS58243.2024.10256234.
5. N. S. Dhiah, R. A. Adib, and M. A. Ghani, “A Novel Hellinger Clustering Method for Efficient ECG Optimized Classification Based on PSO,” International Journal of Information Technology, vol. 16, no. 2, pp. 113–120, 2024. DOI: 10.1007/s41870-024-01360-5.
6. A. Alshraideh, “Enhancing heart attack prediction using machine learning algorithms,” Applied Computational Intelligence and Soft Computing, vol. 2024, pp. 1–12, 2024. DOI: 10.1155/2024/6678345.

7. V. Kumar, A. Gupta, and S. Tiwari, “Efficient ECG Classification Based on Machine Learning and Feature Selection Algorithm for IoT-Enabled Healthcare,” Procedia Computer Science, vol. 215, pp. 693–700, 2024. DOI: 10.1016/j.procs.2022.10.114.
8. P. Kothuru and P. Kumar, “Outperforming Optimised Neural Networks for Cardiac Disease Detection in IoT Frameworks,” Health and Technology, vol. 14, no. 1, pp. 43–56, Jan. 2024. DOI: 10.1007/s12553-023-00783-1.
9. M. Pourvahab, H. Khosravi, and H. Ebrahimpour-Komleh, “A Cluster-Based Opposition Differential Evolution Algorithm Boosted by a Local Search for ECG Signal Classification,” arXiv preprint, arXiv:2312.00212, 2023. [Online]. Available: https://arxiv.org/abs/2312.00212
10. P. Sharma, M. Bhatt, and R. Garg, “A Random Forest-Swarm Optimization Based Approach for Heart Disease Diagnosis,” Journal of Biomedical Informatics, vol. 121, pp. 103880, Oct. 2021. DOI: 10.1016/j.jbi.2021.103880.
11. A. K. Sharma, D. Kumar, and N. Verma, “New Cardiovascular Disease Prediction Approach Using Quantum-Behaved Particle Swarm Optimization,” Multimedia Tools and Applications, vol. 82, pp. 22401–22421, Aug. 2023. DOI: 10.1007/s11042-022-12428-1.
12. H. Mirjalili, M. Mahdavi, and A. Rahmani, “Combining SVM and Elephant Herding Optimization for Cardiac Arrhythmias Detection,” Biomedical Engineering Letters, vol. 8, no. 1, pp. 1–9, Jan. 2018. DOI: 10.1007/s13534-017-0057-3.
13. S. Das and A. Ghoshal, “Heart Disease Detection Using a Hybrid of Bacterial Foraging and Particle Swarm Optimization Algorithm,” Evolving Systems, vol. 10, no. 1, pp. 27–36, Mar. 2019. DOI: 10.1007/s12530-018-9223-7.
14. M. Yildirim and U. Baloglu, “Heartbeat Type Classification with Optimized Feature Vectors Based on Wavelet Transform and Swarm Intelligence,” International Journal of Optimization and Control: Theories & Applications (IJOCTA), vol. 13, no. 1, pp. 55–67, Mar. 2023. DOI: 10.11121/ijocta.2023.1301.02.
15. H. Chen, M. Tang, and X. Zhang, “An IoT-Based Real-Time ECG Monitoring System with Machine Learning Classification,” Wireless Networks, vol. 30, no. 2, pp. 981–993, Feb. 2024. DOI: 10.1007/s11276-023-03276-y.