Analysis Of ECG Arrhythmia For Heart Disease Using SVM And Island-Based Cuckoo Search With Highly Disruptive Polynomial Mutation (Icspm)

Dr. Dinesh D. Patil [1], Manish Sunil Wadile[2]
Department of Computer Science and Engineering
Shree Sant Gadge Baba Engineering College, Bhusawal

Abstract

Electrocardiogram arrhythmias are abnormal heart rhythms that can lead to serious medical conditions if undetected. Early and accurate detection of arrhythmia will play an important role in patient care. In this paper, we proposed a more optimized approach for detecting and analyzing arrhythmia using the MIT database. The classification is done for the Detection of abnormal heart rhythms using a Support vector machine (SVM) and Island-based Cuckoo Search with a highly disruptive polynomial mutation (iCSPM) Optimized Neural network.

Keywords: Island-based cuckoo search; support vector machine; Neural Networks; Machine Learning; Electrocardiograms; QRS; Cardiovascular disease;

I. Introduction

Cardiovascular disease (CVD) is a major public health concern, particularly in India, where it ranks as a leading cause of mortality. Heart attacks and strokes, responsible for over 80% of cardiovascular deaths, are linked to risk factors such as tobacco use, sedentary lifestyles, obesity, high blood pressure, diabetes, and genetic predispositions.

The heart, central to the cardiovascular system, is susceptible to various diseases, from benign conditions like tachycardias to life-threatening ones like myocardial infarctions. Early detection of cardiac arrhythmias is essential, as minor arrhythmias can indicate potential progression to severe forms of the condition.

Electro Cardio Grams (ECGs) are invaluable for this purpose, offering a quick, cost-effective, and non-invasive means of assessing cardiac health. ECGs represent the heart's electrical activity as waveforms, with deviations indicating cardiac dysfunctions or arrhythmias.

Recognizing arrhythmias involves preprocessing noisy Holter recordings, extracting relevant features, and using classifiers. The effectiveness of this process hinges on meticulous data preparation, feature extraction, and classification—a multidisciplinary endeavour aimed at improving cardiac healthcare.

II. Proposed Method

A. *Data processing and feature extraction*

Data Preprocessing:
The MIT-BIH Arrhythmia database contains ECG recordings with annotated beat labels. We will preprocess the raw ECG signals by applying filtering techniques and noise reduction to ensure the data's quality. ECG signals will be segmented into individual beats, and the beat labels will be extracted for supervised learning.

Feature Extraction:
The key to ECG classification is the selection and extraction of informative features. We will compute a range of descriptors, including:
Time Intervals: Features like RR intervals (time between successive R-peaks) are crucial for characterizing heart rhythms.
Wavelet-Based Features: Wavelet transformations are used to capture signal details across different scales.
Local Binary Patterns (LBP): LBP analysis can describe the texture of the ECG waveform.
Higher-Order Statistics (HOS): Features like kurtosis and skewness provide insights into waveform shapes.
Amplitude Values: Mean and standard deviation of signal amplitudes are included.
Formulas and algorithms are applied to compute these features, and mathematical expressions, such as the discrete wavelet transform and LBP calculations, are essential components of the feature extraction process.

Ensemble SVM Modeling:
To perform ECG beat classification, we'll implement multiple SVM models, each designed for a specific feature category. SVMs are powerful classifiers that find the hyperplane which maximizes the margin between classes. We'll optimize SVM hyperparameters, such as the kernel function and regularization parameter (C), by using techniques like binary SVM.

Ensemble Learning:
Our approach involves combining the decisions of different SVM models to produce the final prediction. We will explore ensemble methods such as product, sum, and majority rules to create a consensus prediction. This ensemble strategy leverages the strengths of individual SVM models and helps in achieving improved classification performance.

Evaluation and Benchmarking:
We will evaluate the performance of our ECG classification model using metrics such as accuracy, precision, recall, and F1-score. A separate validation dataset will be used to fine-tune the model and prevent overfitting. We will compare our ensemble SVM approach with a single SVM model using the same features and also assess the results against previous machine learning approaches and state-of-the-art methods for ECG classification.

Ensemble iCSPM Optimized Neural network :
the Island-Based Cuckoo Search with Polynomial Mutation (iCSPM) algorithm offers a valuable solution. This approach integrates iCSPM, a population-based optimization algorithm, into the training process. It begins with data preprocessing and defining the neural network architecture. Key hyperparameters, such as learning rates, batch sizes, and architecture-specific parameters, are identified for optimization.

iCSPM operates using an island-based parallel model, where a population of candidate solutions represents various hyperparameter configurations for the neural network. These candidates are spread across islands, each optimizing the neural network with a specific set of hyperparameters. Periodic migration among islands facilitates the sharing of the best-performing solutions, enhancing the collective exploration of the hyperparameter space. Fitness evaluation is conducted to assess each neural network's performance on its respective island, with a best-worst migration policy to select the most promising candidates for exchange.

The final result is an optimized neural network configuration with fine-tuned hyperparameters. This configuration is determined by the best-performing solutions identified during the iCSPM optimization process. The optimized neural network is then trained on the entire dataset for ECG arrhythmia classification, allowing for improved classification accuracy and robust generalization capabilities on unseen data.

B. Support Vector Machine for Classification

The One-vs-One (OvO) multiclass Support Vector Machine (SVM) approach extends binary SVM classification to handle problems with multiple classes. In the OvO strategy, a binary classifier is trained for every possible pair of classes, resulting in $\frac{N(N-1)}{2}$ classifiers for N classes. Each binary classifier is designed to distinguish between one class and another. The formulation for the OvO SVM can be expressed as follows:

For each pair of classes (*i*, *j*) where $i \neq j$ and *i,j* are in the range from 1 to N, a binary classifier is created. The binary classifier's objective is to find a hyperplane that separates class *i* from class (*j*) by maximizing the margin:

$$min_{w_{ij},b_{ij}} \frac{1}{2} \left\| w_{ij} \right\|^2$$

Subject to the constraints:
$y_k(w_{ij}.x_k + b_{ij}) \geq 1$ for all k where $k \in \{i,j\}$

The above equations ensure that data points from class *i* are correctly classified with a positive margin, while data points from class *j* are correctly classified with a negative margin. The decision boundary between classes *i* and *j* is determined by the weight vector w_{ij} and bias term b_{ij}.
During the prediction phase, for a given test sample x, each binary classifier (i,j) computes a score $f_{ij}(x)$
$f_{ij}(x) = w_{ij}.x + b_{ij}$

The class with the highest sum of scores across all binary classifiers is selected as the predicted class:

$$\hat{y} = argmax_k \sum_{(i,j)} vote_{(i,j)}(k)$$

Here, $vote_{(i,j)}(k)$ represents the "votes" for class k from binary classifier (i,j). The class with the most votes is chosen as the predicted class.
The OvO strategy provides a versatile approach for multiclass classification, particularly when SVMs are used as binary classifiers. However, the number of classifiers grows quadratically with the number of classes, making it computationally intensive for problems with many classes. Nonetheless, it remains a powerful method for handling multiclass classification tasks using SVMs.

C. Island-based Cuckoo Search with a highly disruptive polynomial mutation (iCSPM) Optimized Neural network.

The Island-Based Cuckoo Search with Highly Disruptive Polynomial Mutation (iCSPM) algorithm is a significant advancement in the domain of optimization algorithms. It specifically addresses the challenge of premature convergence; a recurring issue that traditional optimization techniques often encounter. Premature convergence refers to the situation where optimization algorithms converge to suboptimal solutions early in their execution, limiting their effectiveness in solving complex problems.
iCSPM introduces a novel parallel variation of the Cuckoo Search (CS) algorithm, which is renowned for its simplicity and efficiency. What sets iCSPM apart is its unique approach to population management. The algorithm organizes the candidate solutions into isolated islands, creating a structured and controlled environment for optimization. To enhance exploration, iCSPM replaces the conventional Lévy flight operator with a highly disruptive polynomial mutation (HDP) method. This modification introduces a level of diversity that was previously lacking in CS, mitigating the problem of premature convergence.
A key feature of iCSPM is its island model, which employs a periodic migration process to facilitate the exchange of information among islands. The migration is governed by parameters such as migration frequency (Mf) and migration rate (Mr). These parameters determine how often and to what extent candidate solutions move between islands. Moreover, iCSPM uses a best-worst migration policy, aligning with the principle of survival of the fittest, to select the most promising solutions in an island for exchange with the weakest solutions in a neighboring island. This strategy of promoting the best candidates over random ones enhances the optimization process and is widely favored in the island model framework.
In summary, iCSPM is a cutting-edge optimization algorithm designed to overcome the challenge of premature convergence by introducing parallelism, structured population management, and a disruptive mutation operator. This combination of features makes iCSPM a powerful tool for tackling complex optimization problems that demand both diversity and efficiency in the search for optimal solutions.

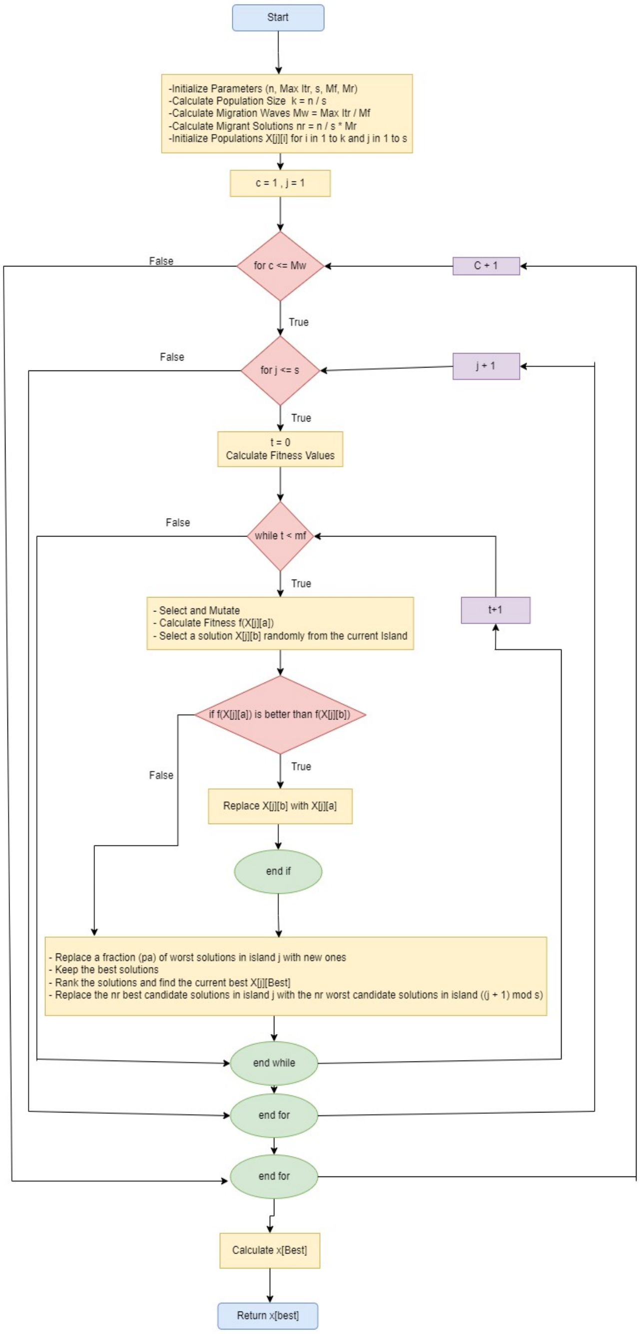

Fig.1. Flow chart of Island-based cuckoo search

We have following terms used in flowchart:

1. ***n:*** *Total number of candidate solutions.*
2. ***Max Itr:*** *Maximum number of iterations.*
3. ***Mf:*** *Migration Frequency*
4. ***Mr:*** *Migration rate*
5. ***s:*** *number of Islands*
6. ***k****: The population size for each island. It's calculated by dividing the total number of candidate solutions (n) by the number of islands (s).*
7. ***Mw****: The number of migration waves. It's determined by dividing the maximum number of iterations (Max Itr) by the migration frequency (Mf).*
8. ***nr:*** *The number of migrant solutions exchanged between islands. It's calculated by multiplying the total number of candidate solutions (n) by the migration rate (Mr).*
9. ***X[j][i]:*** *A candidate solution (often represented as a vector) in island j and position i within that island's population.*

III. Limitations

Computational Complexity: iCSPM's computational demands can be a significant limitation, especially for large datasets or complex neural network architectures. The algorithm's population-based nature and exploration of a wide hyperparameter space can result in high computational costs. This limitation might require extensive computational resources, which could be a challenge for users with limited computing power.

Hyperparameter Space Exploration: The effectiveness of iCSPM depends on its ability to efficiently explore the hyperparameter space. When the search space is vast and complex, finding the optimal configuration can be challenging. Users must carefully define and restrict the hyperparameter search space to ensure the algorithm's efficiency and effectiveness.

Sensitivity to Initializations: iCSPM's performance can be sensitive to the choice of initializations, as is the case with many optimization algorithms. The selection of initial populations and islands can impact the convergence to a global optimum. To mitigate this limitation, multiple runs with different initializations may be required to increase the likelihood of discovering the best hyperparameter configuration.

Acknowledgment

I would like to express my sincere gratitude to my research guide Dr. Dinest D. Patil for their invaluable guidance, unwavering support, and expert mentorship throughout this research project. Their dedication and insightful feedback have been instrumental in shaping the outcome of this work.

I am also thankful to my college Shri Sant Gadge Baba College, Bhusawal, Maharashtra,India for their contributions and support during this research endeavor. Their assistance has been greatly appreciated.

References

[1] Dinesh D. Patil1, R. P. Singh, Vilas M. Thakare "Analysis of ECG Arrhythmia for Heart Disease Detection using SVM and Cuckoo Search Optimized Neural Network" International Journal of Engineering & Technology,2018.

[2] Bilal H. Abed-alguni "Island-based Cuckoo Search with Highly Disruptive Polynomial Mutation." ResearchGate, Feb 2019

[3] Marriwala, H., Iyer, R. "Arrhythmia Detection in ECG Using Ensemble of Machine Learning Algorithms." Procedia Computer Science, 2019, Vol. 132.

[4] Pourmohammadi, A., Marateb, H.R., Eslahchi, C. "Deep Learning-Based Arrhythmia Detection Using Long Short-Term Memory (LSTM) Network." Journal of Medical Signals & Sensors, 2020, Vol. 10, No. 1.

[5] Abiyev, R.H., Cosar, A. "Detection of Arrhythmia in ECG Signals Using Deep Learning." Studies in Health Technology and Informatics, 2019, Vol. 264.

[6] Zhang, J., Kang, X., Liu, L., et al. "ECG Arrhythmia Detection Using Feature-Based Data Mining and Fuzzy Neural Network." Expert Systems with Applications, 2013, Vol. 40, No. 16.

[7] Maheswari, S., Thanushkodi, K., Dash, S. "Arrhythmia Classification Using Modified Grey Wolf Optimization and Random Forest." Biomedical Signal Processing and Control, 2018, Vol. 39.

[8] Bang, S.Y., Lee, T.H. "ECG Beat Classification Using Particle Swarm Optimization." Expert Systems with Applications, 2014, Vol. 41, No.4.

[9] Rajpurkar, F., Hannun, A.Y., Haghpanahi, M., et al. "ECG Arrhythmia Classification Using a Convolutional Neural Network." arXiv preprint, 2017.

[10] Haque, B.J., Razib, S.N.M., AlZoubi, R.Y. "Automatic Detection of Arrhythmias in ECG Signals with Decomposition and Ensemble Learning." Biomedical Signal Processing and Control, 2013, Vol. 8, No. 6.

[11] Usama, M., Zhang, R., Yi, S., et al. "A Hybrid Deep Learning-Based ECG Classification Algorithm." Computers in Biology and Medicine, 2020, Vol. 120.

[12] Huang, X., Yu, Z., Huang, S., et al. "ECG Arrhythmia Classification Using Feature Selection and Random Forest." Proceedings of the International Joint Conference on Neural Networks (IJCNN), 2017.

[13] Feng, W., Yan, Y., Wang, Y., et al. "ECG Arrhythmia Detection Using a Novel Deep Learning Network." Biomedical Signal Processing and Control, 2020, Vol. 62.

[14] Tang, H., Yu, Z., Huang, S., et al. "Multi-Channel ECG Classification Using Extreme Learning Machine." Frontiers in Physiology, 2017, Vol. 8.

[15] Samadi, H., Alhussein, N. "ECG Arrhythmia Classification Using a Convolutional Neural Network with Causal Convolution." Proceedings of the International Joint Conference on Neural Networks (IJCNN), 2021.

[16] Amaral, I.F., do Amaral, L.M.F., Araújo, R.M.R. "A Hybrid Machine Learning Approach to ECG Arrhythmia Classification." Computers in Biology and Medicine, 2016, Vol. 74.

[17] Hussain, A., Dogra, D.P., Suganthan, P.N. "ECG Arrhythmia Classification Using a Deep Convolutional Neural Network." Future Generation Computer Systems, 2020, Vol. 107.