Digital Innovations & Contemporary Research in SCIENCE, ENGINEERING & TECHNOLOGY
Vol. 12. No. 2, 2024 Series

Article Citation Format
Muhammad-Thani, S., Olabiyisi , S. O., Omidiora, E. O. & Ganiyu, R. A (2024): A Comparative Analysis of Metaheuristic Algorithms on Hybrid Features for Feature Selection on ECG-Based Arrhythmia Detection. Journal of Digital Innovations & Contemporary Research in Science, Engineering & Technology. Vol. 12, No. 2. Pp 1-16. www.isteams.net/digitaljournal. dx.doi.org/10.22624/AIMS/DIGITAL/V11N2P1

Article Progress Time Stamps

Article Type: Research Article
Manuscript Received: 16th February, 2024
Review Type: Blind Peer
Final Acceptance: 22nd March, 2024

# A Comparative Analysis of Metaheuristic Algorithms on Hybrid Features for Feature Selection on ECG-Based Arrhythmia Detection

[1]Muhammad-Thani, S., [2]Olabiyisi , S. O., [3]Omidiora, E. O. & [3]Ganiyu, R. A
[1]Department of Computer Science, Kwara State College of Education, Ilorin, Nigeria
[2]Department of Computer Science, Ladoke Akintola University of Technology, Ogbomoso, Nigeria
[3]Dept of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria,
**E-mails:** s.mthani@kwcoeilorin.edu.ng; soolabiyisi@lautech.edu.ng; eoomidiora@lautech.edu.ng, raganiyu@lautech.edu.ng

## ABSTRACT

Electrocardiogram (ECG)-based arrhythmia detection is crucial for diagnosing cardiovascular diseases promptly. However, the high-dimensional nature of ECG signals poses challenges for accurate detection. Feature selection plays a vital role in enhancing classification performance by identifying relevant features. This paper presents a comparative analysis of metaheuristic algorithms for feature selection in ECG-based arrhythmia detection using a hybrid feature set based on morphological and heart variability features. We evaluate the performance of Grey Wolf Optimization (GWO), and Firefly Algorithm (FFA) in terms of classification accuracy, sensitivity, Positive precision, and specificity. The experimental results shows that GWO performs FFA better in selecting relevant features for arrhythmia classification by 2.8%. However, both algorithms demonstrate the efficacy of metaheuristic algorithms and hybrid features in improving arrhythmia detection and provide insights into their comparative performance.

**Keywords:** ECG Signal, Arrhythmia, Feature Selection, Metaheuristic Algorithms.

## 1. INTRODUCTION

Cardiovascular diseases remain a leading cause of mortality worldwide, highlighting the importance of early detection and intervention. (2) Heart arrhythmia, or irregular heartbeat, is the most obvious sign of cardiovascular disease. The most widely used method for evaluating heart function is the electrocardiogram (ECG), which captures the electrical impulses the heart produces from electrodes placed on the body and graphically displays the patterns of activity (Prystowsky and Padanilam, 2021). The main variables generated by ECG signals are the P, Q, and QRS complex waves. The form, connections, and interval ECG signals are widely used for diagnosing cardiac arrhythmias due to their ability to capture the heart's electrical activity. Due to high volume of ECG patient data, computer-aided diagnostics systems have been a core research area to assist physicians in the prompt detection of arrhythmias.

However, the high dimensionality of ECG data presents challenges for accurate and efficient arrhythmia detection. An effective heartbeat classification system typically comprises three important steps: feature extraction, feature selection, and classification (Kaya et al., 2017). Feature extraction and selection are two important modules as they affect classifier performance. Consequently, any ECG classification system aims to achieve optimal accuracy through adequate feature extraction and reduction of feature sizes. While feature extraction centers on expunging representation of important delineation characterized in ECG signals, feature selection specifically seeks to select relevant features that represent ECG signals by reduction in feature sizes while retaining the original representation of the signals.

Therefore, feature selection techniques aim to identify a subset of relevant features, towards achieving an improved classification performance. This study evaluates the performance of FFA (Firefly Algorithm) and Grey Wolf Optimization (GWO) metaheuristics as feature selection techniques on MultiLayer Perceptron Neural Network classifier in the detection of five arrhythmia classifications according to Association of the Advancement of Medical Instrumentation standards (AAMI). It further evaluates the performance of morphological and heart variability features in arrhythmia classification.

## 2. RELATED WORKS:

Many research efforts have been driven towards adopting metaheuristic algorithms as feature selection paradigms in achieving efficient and accurate diagnosis of arrhythmias. Metaheuristic algorithms have gained attention due to their ability to efficiently search large solution spaces and identify optimal feature subsets. Metaheuristics are also known to have an impactful effect on classifier performance (Tsai et al., 2016) . Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Grey Wolf Optimization, Bacteria Foraging Optimization (BFO) and Firefly algorithm (FFA) are among the widely used metaheuristic algorithms for feature selection in various arrhythmia classification problems. (Bouaziz et al (2019); Kora (2017); Lu et.al (2016) However, a comprehensive comparative analysis of these algorithms for ECG-based arrhythmia detection is lacking.

The work of Llamedo and Mart´ınez (2010) marked the introduction of feature selection paradigm and database generalization criteria in ECG analysis and classification. Time domain and morphological features were extracted using DWT. Sequential floating forward selection (SFFS) was applied to extracted features to select optimal features subset of which Linear Discriminant classifier was used for classification. A multiclass approach in ECG classification was adopted classifying ECG signal into five classes as recommended by Association for the Advancement of Medical Instrumentation (AAMI).

Performance evaluation results indicated 93% global accuracy, 95% sensitivity; 98% positive predictive value. Supraventricular beats had 77% and 39% sensitivity and positive predictive values respectively, while for ventricular beats, 81% sensitivity and positive predictitviy 87% were achieved. The limitation of the proposed approach is the Gaussian assumption of the data imposed by the classifier implemented. However, the effect of the lack of Gaussianity can be mitigated using more complex classifiers, like ANN.

Kora and Kalva (2015) developed a feature selection technique for ECG classification system based on the hybridization of Bacteria Foraging Optimization (BFO) and Particle Swarm Optimization (PSO) techniques. It was observed that BFO has the problem of achieving global optimum because the chemotaxis process of BFO depends on random search solutions; thus there is need for hybrid BFPSO for optimal feature selection. PSO was used for global search while local search was carried out using BFO. Twenty optimal features out of 200 ECG features were selected using BFPSO. Four classifiers viz KNN, SVM, NN and LMNN were implemented to evaluate the performance of BFPSO. BFPSO outperformed GA, PSO and BFO in conjunction with LMNN with an accuracy of 98.1%, sensitivity of 98.97% and specificity of 98.7%.

Kora and Krishna (2016) observed that traditional feature extraction techniques such as autoregressive modelling and discrete wavelet transform used in ECG analysis yielded large number of features. Additionally, Firefly algorithm had the problem of getting trapped into local minimum. Hence, hybrid firefly and Particle Swarm Optimization (PSO) was proposed for the detection of bundle branch block. PSO operator was used to accomplish global search and alleviate the problem of firefly observed. Hybrid firefly and Particle Swarm Optimization was used to select relevant feature while reducing feature dimensionality. It showed a promising result of 99.1%. The algorithm was used to select features for the detection bundle branch block only.

A suggested methodology for increasing the accuracy of ECG pattern detection using feature dimensionality reduction was offered by Lu et al. (2016). Their work focused on employing enhanced Genetic Algorithm and Empirical Mode Decomposition for feature selection in ECG Signal Processing. It was found that GA presented computational challenges related to population size selection. Therefore, an inductive learning approach was introduced into GA to address issues found. ECG signals were broken down into intrinsic mode functions (IMFs) using the Empirical Mode Decomposition (EMD) technique for feature extraction. From IMFs, 71 statistical features were taken out. The best traits were then chosen using an enhanced genetic algorithm. A 96.13% accuracy rate overall was achieved.

Atal and Singh (2020) used a bat-rider optimization to optimize CNN on the MIT- BH database to detect arrhythmia. Gabor, wave and heart variability features were fed into CNN based classification model to identify arrhythmia or no arrhythmia in a patient. An accuracy of 93.19%, specificity of 95% and sensitivity of 93.9% was achieved. This result is due to the limitation of CNN models, which are computationally expensive and not effective on small datasets.

Houseein et al (2021) proposed an arrhythmia detection system based on Mantra foraging optimization for feature selection in ECG-based arrhythmia classification. ECG signal feature descriptors based on one-dimensional local binary pattern (LBP), wavelet, higher-order statistics, and morphological information were introduced for feature extraction. MRFO was used to optimize SVM parameters and select the relevant feature subsets that provide the best classification performance. The proposed MRFO-SVM approach was trained on the MIT-BIH Arrhythmia database containing four abnormal and one normal heartbeat. An overall accuracy of 98.26% accuracy was recorded.

Hassaballah et al. (2023) proposed an ECG-based arrhythmia scheme based on the Marine Predator Algorithm. (MPA). MPA was utilized specifically to overcome the choice of parameter issue that most machine learning classifiers have. For the classification problem, the MPA approach was used to optimize the learning parameters of four supervised machine learning classifiers: random forest (RF), k-nearest neighbors (kNNs), gradient boosting decision tree (GBDT), and support vector machine

(SVM). Several experiments were carried out on three common databases, namely the Massachusetts Institute of Technology (MIT-BIH), the European Society of Cardiology ST-T (EDB), and the St. Petersburg Institute of Cardiological Techniques 12-lead Arrhythmia (INCART), to validate the benefit of the proposed approach. The findings collected demonstrated that the integration of the MPA algorithm greatly enhanced the performance of all tested classifiers, with an average accuracy of 99.92% and a sensitivity of 99.81% for the classification of ECG arrhythmias.

Deepa et.al (2023) implemented a black widow optimization metaheuristic algorithm to optimize hybrid wavelet features for arrhythmia classification. Kernel-based Support Vector machine was used to classify 15,000 ECG signals obtained from MIT-BH ECG database into five arrhythmia classes. The result showed that the best accuracy of 99.91% was achieved with 2$^{nd}$ wavelet coefficients using an RBF kernel SVM.

Khafaga et. al (2023) implemented a Dipper Throated optimized K Nearest neighbor classifier to categorize arrhythmia into 16 classes. The Dipper Throated Algorithm was used to select important features out of 279 attributes present in the arrhythmia UCI database. An overall accuracy of 99.8% was achieved.

Tunç, & Cangöz, (2024) worked on a comparative study of the effects of minimum redundancy maximum relevance (MRMR), Chi-square and Matched score feature selection techniques on three classifiers via; Support Vector Machine, K-Nearest Neighbor and Decision Tree. The Arrhythmia Dataset obtained from the UCI repository was used in this work. The matched score method outperformed MRMR and chi-square feature selection methods on SVM classifier with an accuracy of 81.27%.

A novel deep-learning method employing recurrent CNN with Grey Wolf Optimization (GWO) for ECG classification-based arrhythmia prediction was proposed by Singh and Mahaptra (2024). Support Vector Machine, Decision tree, K-NN and Recurrent Convolutional Neural Network (RCNN) Classifiers were employed to examine the impact of GWO metaheuristic algorithm. RCNN-GWO performed best with an overall accuracy of 98%. This study seeks to evaluate GWO and FFA metaheuristics as feature selection methods while considering their performance on hybrid features in ECG-based arrhythmia classification.

Digital Innovations
& Contemporary Research in SCIENCE, ENGINEERING & TECHNOLOGY
Vol. 12. No. 2, 2024 Series

## 3. MATERIALS AND METHODS

The figure below is the block diagram of the workflow of the proposed system

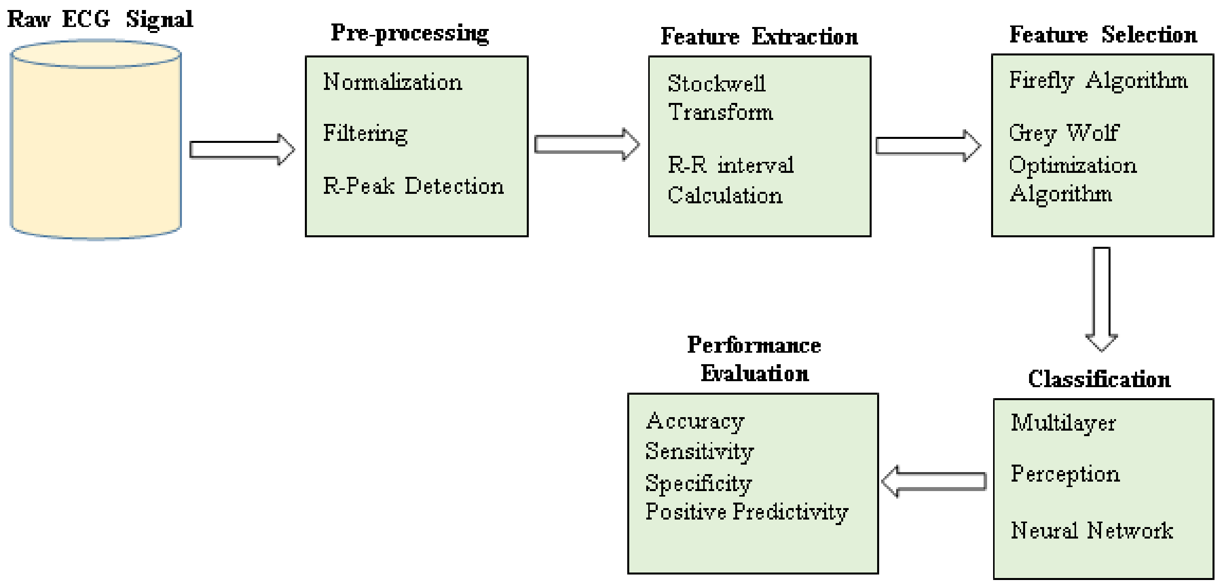


Figure 1 Block Diagram Of The Workflow Of The Proposed ECG-Based Arrhythmia Detection Scheme.

### 3.1 Database

This study employed MIT-BIH arrhythmia database consisting of 48 ECG signal records. The digitized (11-bit resolution) ECG signals in MIT-BIH have been sampled at 360 Hz. For this study, and according to the Association for the Advancement of Instrumentation (AAMI), standards only 44 of the recordings are considered, while 4 recordings containing paced beats are excluded due to poor signal quality. The 44 recordings consist of 100716 heartbeat samples distributed on five arrhythmia classes. Ninety thousand, one hundred and eight (90,108) samples are classified as normal heartbeats (N), and 2781 samples are classified as supraventricular heartbeats (S). While 7009, 803 and 15 samples are classified as ventricular (V), fusion (F) and unknown (Q) heartbeats respectively.

### 3.2 Preprocessing

It is common for noise and artifacts to contaminate ECG data obtained from a patient's body. It is challenging to discern between arrhythmia and normal ECG signals due to noise in the data. Therefore, to further improve the feature extraction phase and improve classification accuracy, signals for the ECG classification task must be preprocessed to eliminate noise and artifacts (Rangayyan, 2002). Additionally, preprocessing is required to increase signal-to-noise ratio (SNR), which is helpful for the later detection of the fiducial point (R peak and QRS complex) and the classification of heartbeats. ECG signals were normalized ECG signals are normalized to zero mean and standard deviation of unity. The purpose is to reduce DC offset and remove amplitude variance for each ECG signal.

Thereafter, Pan Tompkins algorithm was implemented to remove noise using high and low bandpass filters. R peak detection was achieved by adjusting thresholds on QRS complexes of each signal. Figure 2 shows filtered ECG signals.

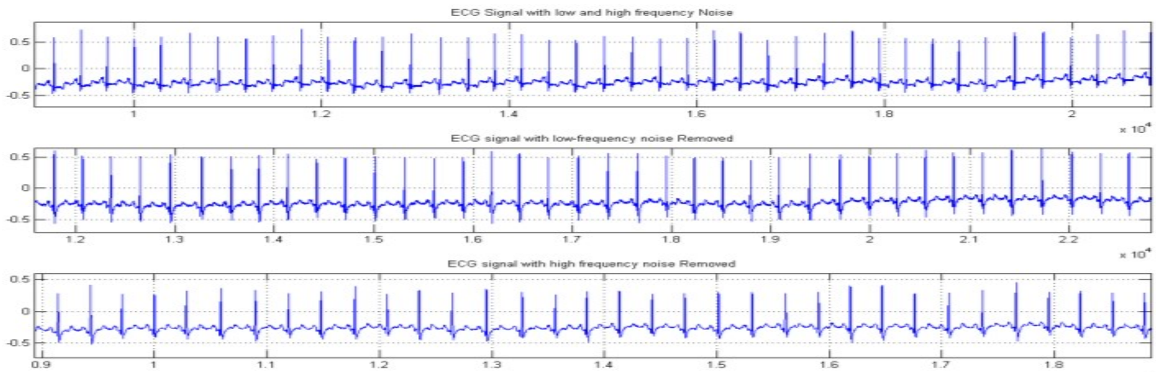


Fig. 2 Denoised and Filtered ECG signal

### 3.3 Beat Segmentation

Prior feature extraction, each beat in the ECG needs to be isolated and assembled into a single beat format. An ECG signal with a window length of 200 points per beat is extracted based on the location of the QRS complex, which is identified using Pan Tompkin's technique (Pan and Tompkin, 1985). The final point is completed on the R point itself. These points are established with 99 and 100 ones on the left and right sides of the R point, respectively. The location of the R point for the associated beat was determined with the aid of the MIT-BIH database's annotation files.

### 3.4 Feature Extraction

Morphological and heart variability features are considered in this work and extracted from the preprocessed and segmented ECG signals. Both features are important in medical practice in detecting arrhythmias because morphological features capture changes in the waveform in amplitude and time domains. In addition, heart variability features also known as R-R interval are dynamic revealing changes in heartbeat around the R-peak of the ECG signals. (Can Ye, 2012). Similarly, The R-R interval is an important feature in ECG signal classification, as it indicates the time taken between each ventricular blood ejection or two consecutive heartbeats. In consonance with the choice of morphological and heart variability features is the work of Lin and Yang (2014) which exhibit an increase in detection accuracy of arrhythmia. Stockwell Transform is used to extract morphological features. It is utilized in determining signal properties in the time and frequency domain. An essential characteristic of the S-transform is its ability to integrate, uniquely, the Fourier Transform's globally referenced phase measurement feature and the Wavelet Transform's frequency-dependent resolution (Agrawal and Vijay, 2013).

It has been applied in ECG analysis in literature (Das and Ari, 2014; Jevoka *et al*., 2020). S-transform is defined as follows:

$$S(\tau, f) = \int_{-\infty}^{\infty} ECG_T(t) . \frac{|f|}{\sqrt{2\pi}} . e^{\frac{-(\tau-t)^2 . e^2}{2}} . e^{-i2\pi f t} dt \qquad (1)$$

where $ECG_T(t)$ is the ECG signal in the time domain, $\tau$ denotes the time of spectral localization and $f$ is the frequency in the Fourier transform.

The Fourier transform $ECG_T(t)$ is

$$ECG_T(f) = \int_{-\infty}^{\infty} ECG_T(t) \,.\, e^{-i2\pi f t} dt \qquad (2)$$

The S-transform could be represented as operations on the Fourier spectrum $ECG_T(f)$ of the time signal $ECG_T(t)$:

$$S(\tau, f) = \int_{-\infty}^{\infty} ECG_F(f+\alpha) \,.\, e^{\frac{-2\pi^2 . \alpha^2}{f^2}} \,.\, e^{-i2\pi f t} d\alpha \qquad (3)$$

The output of the S-transform is a matrix of complex values with rows and columns indicating the frequency and time of the analyzed ECG signals. R-R interval features are considered in four ways described as follows:

i. pre-RR-interval: describes the R-R interval between a given heartbeat and the previous heartbeat;
ii. post-RR intervals: the R-R interval between a given heartbeat and the following heartbeat;
iii. average RR-intervals: the mean of the RR-intervals for a recording and is considered as the same value for all heartbeats in a recording;
iv. local average RR-interval: taking as the average of ten RR-intervals surrounding a heartbeat (Anwar, 2018). These features were calculated using the following equations 4-7

$$RR_{pre}(i) = R(i) - R(i-1) \qquad (4)$$

$$RR_{post}(i) = R(i+1) - R(i) \qquad (5)$$

$$RR_{local} = \frac{1}{10} \sum_{i=5}^{5} RR(i) \qquad (6)$$

$$RR_{ave} = \frac{1}{N_{RR}} \sum_{i=1}^{N_{RR}} RR(i) \qquad (7)$$

where $i$ is the location of the current R peak. $RR_{pre}$, $RR_{post}$, $RR_{local}$ and $RR_{aver}$ depict the previous, post, local and average RR interval respectively. R ($i$ – 1) and R ( $i$ + 1) represent the previous and the post R peaks while $N_{RR}$ shows the total number of RR in an ECG segment (Anwar, 2018). Morphological features were extracted by obtaining the S-transform coefficients from the preprocessed signals using equations 1 to 3. One hundred and eighty (180) features were extracted from one ECG cardiac cycle by selecting a window of −250ms to +250ms around the R-peak equivalent to 200 beat samples (Das, 2015).

Let X be the S-transform-based morphological feature vector. That is,

$$X = [M_1, M_2, M_3, \ldots M_{180}] \qquad (8)$$

Let Y represent the heart variability feature vector

$$Y = [T_1, T_2, T_3, T_4]. \qquad (9)$$

The morphological and temporal features are combined via serial concatenation to form a feature vector containing 184 features. Therefore, the feature vector resulting in (184x 200) dimension was used in this work. The combined feature vector is represented thus:

$$Z = [M_1, M_2, M_3, \ldots M_{180}, T_1, T_2, T_3, T_4] \qquad (10)$$

## 4. FEATURE SELECTION

The excessively large feature vectors degrade the system's performance. As a result, feature selection is required. The goal of feature selection is to obtain the best features that maximize the prediction accuracy of the classified ECG signals. In this work, the hybrid features are used to optimize the MLPNN classifier using Firefly and Grey Wolf Optimization Algorithm.

### 4.1 Firefly algorithm

Firefly Algorithm (FFA) is a nature-inspired algorithm used in optimization problems proposed by Yang (2007). FFA adopts a search mechanism simulated by the social behavior of fireflies and the principle of bioluminescent communication characterized by the function of the flashing light, possessing the ability to attract partners or swarm against potential prey. In FFA, there are two principal variables; they are light intensity and attractiveness. Thus, it beholds that the attractiveness of a firefly is proportional to the light intensity (brightness). Every member $x_i$ of the firefly swarm is determined by its brightness $I_i$ which can be directly measured as a corresponding fitness function (Yang, 2009).

The light intensity I and attraction of the firefly which is inversely proportional to a particular distance r from the light source is calculated using the equation in 11

$$I = \frac{I_0}{e^{\gamma r_{ij}}} \tag{11}$$

Where $I_0$ is the initial fluorescence strength of a firefly, γ represents the light absorption coefficient and r is the spatial distance between two fireflies *i* and *j*.

As a firefly's attractiveness is proportional to the light intensity observed by adjacent flies, the variation of the attractiveness β with the distance r is defined by equation 12

$$\beta = \beta_0 e^{-\gamma r^2} \tag{12}$$

where $\beta_0$ refers to the attractiveness at r = 0.
The distance between two fireflies is defined using Cartesian distance as expressed in Equation 13

$$r_{i,j} = x_i - x_j = \sqrt{\sum_{k=1}^{d}\left(x_{i,k} - x_{j,k}\right)^2} \tag{13}$$

Invariably, Firefly i can travel towards firefly j using this equation. The movement of firefly i being attracted towards the brighter firefly j is defined as

$$x_i = x_i + \beta_0 e^{-\gamma r^2{ij}} \left(x_i - x_j\right) + \alpha \epsilon_i \tag{14}$$

The first term of the equation describes the present location of a firefly, while the second term represents the attractiveness β of a firefly and the last indicates the random part of a firefly. In the above equation, γ characterizes the variation of the attractiveness of fireflies and its value is critically paramount in determining the convergence speed of firefly algorithm. Therefore, if γ approaches zero, i.e. (γ → 0), the attractiveness and brightness becomes constant while β = $\beta_0$. The third term is randomization.

The vector of random variables $\epsilon_i$ is drawn from different distributions such as the Uniform distribution, Gaussian distribution and Lévy flight (Zhou *et al.,* 2019). The term, α is a scaling parameter that controls the step size and it should be linked with the interest of the problems. By implication, a firefly can be seen in any position within a complete global search, as γ approaches infinity (γ ⟶ ∞), the attractiveness and brightness decrease.

$$x_i = x_i + \alpha\,(\text{rand} - 0.5) \tag{15}$$

Where α is a random variable, and rand is a random number consistently spread over the space (0, 1).

### 4.2 Grey Wolf Optimization Algorithm

Grey Wolf Optimization algorithm (GWO) was modeled by Mirjalili et al. based on the social hierarchy and the hunting behavior of grey wolves. The wolves prefer to live in packs operating a very strict social dominance hierarchy, consisting of alpha, beta, delta, and omega.

To mathematically model the GWO based on the social hierarchy characteristics of wolves, the alpha ($\alpha$) is considered to be the fittest solution. The beta (β) and the delta ($\delta$) are consequently utilized as the second- and third-best solutions, respectively. The omega ($\omega$) corresponds to the remainder of the candidate solutions. The hunting (optimization) in the GWO algorithm is guided by the three top wolves, and the rest of the wolves move according to the positions of these three leaders. The GWO is based on three major steps: encircling, hunting and attacking. The encircling behavior of grey wolves is indicated in equation 16

$$\vec{D} = |\vec{C}\,.\vec{X}_p(t) - \vec{X}(t)| \tag{16}$$

where $t$ is the iteration, $Xp$ is the prey position, and $X$ indicates position of the grey wolf. A is a coefficient matrix and D is a vector that depends on the location of the prey $X_p$ which is calculated as follows:

Where A and C are coefficient vectors calculated as follows:

$$\vec{A} = 2a\,.\vec{r_1} - a \tag{17}$$

$$\vec{C} = \overrightarrow{2r_2} \tag{18}$$

where $a$ is linearly diminished throughout iterations controlling *exploration* and *exploitation*, and $r_1$ and $r_2$ are random vectors in the range of [0, 1]. The value of $a$ is the same for all wolves in the search space.

2) Hunting: It is performed by the whole pack based on the information coming from the *alpha*, *beta*, and *delta wolves*, which are expected to know the prey location, as given in the following:

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{19}$$

Where $\overrightarrow{X_1}$, $\overrightarrow{X_2}$, and $\overrightarrow{X_3}$ are defined as follows:

$$\overrightarrow{X_1} = |\overrightarrow{X_\alpha} - \overrightarrow{A_1}\,.\,\overrightarrow{D_\alpha}| \tag{20}$$

$$\overrightarrow{X_2} = |\overrightarrow{X_\beta} - \overrightarrow{A_2}\,.\,\overrightarrow{D_\beta}| \tag{21}$$

$\overrightarrow{X_3} = |\overrightarrow{X_\delta} - \overrightarrow{A_3} . \overrightarrow{D_\delta}|$ (22)

Where $\overrightarrow{X_\alpha}$, $\overrightarrow{X_\beta}$, and $\overrightarrow{X_\delta}$ are the first three best solutions at a given iteration *t*, $\overrightarrow{A_1}$, $\overrightarrow{A_2}$, *and* $\overrightarrow{A_3}$ are defined in equation 2.25, and $\overrightarrow{D_\alpha}$, $\overrightarrow{D_\beta}$, and $\overrightarrow{D_\delta}$ are defined using the following:

$\overrightarrow{D_\alpha} = |\overrightarrow{C_1} . \overrightarrow{X_\alpha} - \overrightarrow{X}|$ (24)
$\overrightarrow{D_\beta} = |\overrightarrow{C_2} . \overrightarrow{X_\beta} - \overrightarrow{X}|$ (25)
$\overrightarrow{D_\delta} = |\overrightarrow{C_3} . \overrightarrow{X_\delta} - \overrightarrow{X}|$ (26)

Where $\overrightarrow{C_1}$, $\overrightarrow{C_2}$, and $\overrightarrow{C_3}$ are defined as in (2.26). This is interpreted by the fact that *alpha, beta, and delta* wolves know the best position of the prey and all the other wolves update their positions based on the position of these wolves using equation.

## 4.2 Classification

Multilayer Perceptron Neural Network (MLPNN), otherwise referred to as Feed Forward Neural Network (FFNN) is a biologically inspired classification algorithm. It consists of a large number of simple neuron-like processing units arranged in layers such that each unit in a layer is connected with all the units in the previous layer (Asteris, 2017). Each connection has different weights of which encodes the knowledge of the network. Such units are often referred to as nodes. The nodes in an FFNN are arranged in layers starting with the input layer and ending with the output layer. In between them are some internal layers called the hidden layers, which provide most of the computational power of the network. Multilayer Perceptron Neural Network was used because it has the advantages of being easily implemented on a hardware platform because of its simple structure, and it can be used to model non-linear problems due to its ability to learn system behavior. (Hamidi, 2012).

A supervised learning approach called backpropagation (BP) was be used to train the Multilayer perceptron neural network. Figure 2 shows a multilayer feed-forward neural network. The algorithm for MLPNN with a backpropagation training algorithm is described as follows. At first, weights are initialized to small random numbers and each unit is associated with a bias. After that, the input data is fed to the network's input layer. For an input unit *j,* its corresponding output is $O_j$. Each input is connected to a hidden unit and each connection has a weight. The total input $I_j$ is computed by multiplying each input by its corresponding weight and the sum is calculated. Hence, for a unit *j* the net input $I_j$ is

$$I_j = \Sigma_i w_{ij} O_i + \theta_j \quad (27)$$

Where $w_{ij}$ is the connection weight from node *i* to node *j* and $O_i$ is the output of node *i*. $\theta_j$ is the bias of node *j*. A sigmoid or logistic function that maps a large range input onto a smaller range of normally 0 to 1 is used.

The output of node *j*, denoted as $O_j$ is calculated as

$$O_j = \frac{1}{1+e^{-Ij}} \quad (28)$$

Finally, the weights are adapted using the backpropagation algorithm. The error $Err_j$ of node $j$, is determined as

$$Err_j = O_j(1 - O_j)(T_j - O_j) \tag{29}$$

Where $O_j$ = Actual output of node $j$
$T_j$ = Target output of node $j$

Then the hidden layer error is calculated as

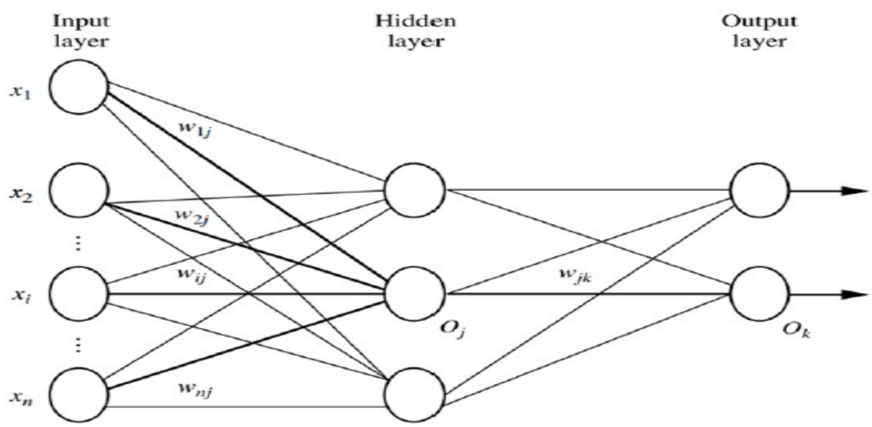


Figure 3. Multilayer Feed-Forward Neural Network (Nahiduzzaman, 2019)

$$Err_j = O_j(1 - O_j)\Sigma_k Err_k w_{jk}$$

Where $w_{jk}$ = connection weight from node $j$ to a node $k$. The weights are adapted as follows

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{30}$$

Where $l$ = Learning rate and the bias is adapted as follows

$$\Delta\theta_j = (l)\mathrm{Err}_j \tag{31}$$

$$\theta_j = \theta_j + \Delta\theta_j \tag{32}$$

The number of input nodes is equal to the number of optimal features selected by the GWO and FFA algorithms. The network has five outputs in conformity with AAMI standards of heartbeat arrhythmia classification.

## 5. RESULTS AND DISCUSSIONS

### 5.1 Experimental Set Up and Data Division

Experiments were conducted using a benchmark dataset containing a diverse set of ECG recordings from MIT-BH ECG databased. Training and testing data were according to Chazal et *al.* (2004) proposed a division of the heartbeats from the MIT-BH database by dividing the 44 recordings considered according to AAMI standards into two namely datasets A and B (DS A, DS B). This allow coherency and conformity with real-life scenarios as heartbeats from DS A and DS B come from different individuals. DS A was used for training while DS B was used for validation (testing). This division protocol is known as inter-patient scheme in literature.

### 5.2 Feature Selection

GWO and FFA were run in 100 iterations. GWO and FFA selected optimal features at the $50^{th}$ and $60^{th}$ iteration respectively. GWO selected 50 optimal features while FFA had an optimal feature size of 65. Both selected features were fed into MLPNN classifier to categorize ECG signals into five arrhythmia classes.

### 5.2 Evaluation of Classification Performance

Performance metrics used to evaluate the effectiveness of each metaheuristic algorithm on MLPNN classifier are based on four statistical parameters; accuracy sensitivity, specificity, and precision. Comparative analyses were performed to identify the strengths and weaknesses of the metaheuristic feature selection approach while considering the impacts of morphological and heart variability features on classifier performance. Furthermore, all 184 features were also input into MLPNN for without applying GWO and FFA as feature selection techniques.

Table 1 and 2 shows the performances of both GWO and FFA metaheuristic algorithms on MLPNN classifier. It is observed that GWO outperformed FFA in all classes of arrhythmia, though both algorithms obtained remarkable results. However, the performances of GWO and FFA dwindle for classes F and Q that is due to a smaller number of beats present in both classes. The highest accuracy of 99.42% was achieved by GWO for class N considering both morphological and heart variability features. An improvement in overall classification accuracy of 2.8% was achieved by GWO over FFA considering both morphological and heart variability features. In comparison, GWO outperformed FFA with an increased overall accuracy of 2.84% considering morphological features only.

An improvement of 0.6% and 0.63% in overall accuracy for morphological feature and heart variability features over morphological features only were recorded in respect GWO and FFA. These results revealed that the choice of combined features from different domains of ECG signals presents an increase in the classification accuracy of arrhythmias. It is important to note that both algorithms had good results for all the metrics presented for the Superventricular (S) and Ventricular (V) classes as these classes are responsible for most life-threatening arrhythmias.

Table 1: Performances of GWO and FFA feature selection methods on the MLPNN classifier using morphological and heart variability features

| CLASS | TECHNIQUE | SEN (%) | SPEC (%) | PREC (%) | ACC (%) |
|---|---|---|---|---|---|
| N | GWO | 98.20 | 98.10 | 98.55 | 99.42 |
| | FFA | 97.23 | 97.11 | 97.03 | 98.12 |
| S | GWO | 98.95 | 99.02 | 99.09 | 99.17 |
| | FFA | 97.31 | 97.23 | 96.11 | 97.56 |
| V | GWO | 98.88 | 98.37 | 99.42 | 99.07 |
| | FFA | 96.97 | 97.76 | 97.63 | 97.47 |
| F | GWO | 93.55 | 92.22 | 93.27 | 94.07 |
| | FFA | 90.84 | 91.42 | 85.21 | 92.15 |
| Q | GWO | 90.00 | 91.71 | 93.82 | 93.78 |
| | FFA | 86.62 | 85.43 | 88.33 | 86.19 |
| AVERAGE | GWO | 95.91 | 95.88 | 96.83 | 97.10 |
| | FFA | 93.79 | 93.79 | 92.86 | 94.30 |

Table 2: Performances of GWO and FFA feature selection methods on the MLPNN classifier using morphological features only.

| CLASS | TECHNIQUE | SEN (%) | SPEC (%) | PREC (%) | ACC (%) |
|---|---|---|---|---|---|
| N | GWO | 97.72 | 98.00 | 98.15 | 99.14 |
| | FFA | 97.03 | 97.00 | 96.61 | 97.42 |
| S | GWO | 98.74 | 98.02 | 98.18 | 99.01 |
| | FFA | 97.01 | 97.12 | 96.03 | 97.56 |
| V | GWO | 98.88 | 98.37 | 99.42 | 99.07 |
| | FFA | 96.97 | 97.76 | 97.63 | 97.50 |
| F | GWO | 91.37 | 90.34 | 93.23 | 93.12 |
| | FFA | 89.48 | 88.82 | 84.17 | 90.71 |
| Q | GWO | 89.05 | 90.11 | 91.28 | 92.16 |
| | FFA | 86.13 | 84.36 | 87.03 | 85.11 |
| AVERAGE | GWO | 95.15 | 94.97 | 96.05 | 96.50 |
| | FFA | 93.32 | 93.01 | 92.29 | 93.66 |

Table 3; Performance of MLPNN classifier without metaheuristic feature selection method

| CLASS | SEN (%) | SPEC (%) | PREC (%) | ACC (%) |
|---|---|---|---|---|
| N | 94.13 | 94.08 | 94.51 | 95.24 |
| S | 94.14 | 93.82 | 94.18 | 94.91 |
| V | 94.80 | 95.03 | 94.12 | 94.64 |
| F | 85.04 | 83.82 | 84.17 | 88.71 |
| Q | 84.38 | 82.80 | 83.86 | 83.51 |
| AVERAGE | 90.49 | 89.91 | 90.17 | 91.40 |

From Table 3, overall average accuracy of 91.40% was achieved for all classes of arrhythmia without implementing either GWO or FFA as selection paradigms. Results shows that GWO had 5.7% and FFA recorded 5.11% increase in classification accuracy which indicates the capability of both algorithms in selecting relevant features for arrhythmia classification.

## 6. CONCLUSION:

This study presents a comparative analysis of metaheuristic algorithms for feature selection in ECG-based arrhythmia detection. The experimental results demonstrate the efficacy of Grey Wolf Optimization and Firefly Algorithm in improving classification performance. However, the impressive performance of GWO is attributed to having an edge over FFA through its ability to maintain balance between exploration and exploitation throughout the search space. (Mirjalili, 2014b).

Insights gained from this study can guide researchers and practitioners in selecting appropriate feature selection techniques for ECG-based arrhythmia detection applications. Further work will focus on comparative analysis of hybrid metaheuristic algorithm for feature selection in arrhythmia classification using other classification techniques.

## REFERENCES

Association for the Advancement of Medical Instrumentation (1987). Recommended Practice for Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms.

Atal, D. K, Singh.M (2020). Arrhythmia Classification with ECG signals based on the Optimization-Enabled Deep Convolutional Neural Network, Computer Methods and Programs in Biomedicine. 196, 105607.

Bouaziz, F., and Boutana, Daoud. (2019).Automated ECG heartbeat classification by combining a Multilayer Perceptron Neural Network with Enhanced Particle Swarm Optimization algorithm. Research on Biomedical Engineering 35:143 -155.

C. Ye, B. V. K. Vijaya Kumar and M. T. Coimbra, (2012). Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals. IEEE Transactions on Biomedical Engineering. 59 (10), 2930-2941.

Chazal, P. D, M. O'Dwyer, Reilly, R.B., (2004). Automatic Classification of Heartbeats using ECG Morphology and Heartbeat Interval Features. IEEE Trans. Biomed. Eng, 51(7):1196-2106.

Deepa S. R., Subramoniam M., Swarnalatha R., Poornapushpakala S., Barani S. Meta-Heuristic Optimized Hybrid Wavelet Features for Arrhythmia Classification.Intelligent Automation and Soft Computing. 37 (1), 746-761

Elhaj, Fatin A. et al. (2023). Heartbeat Classification of Arrhythmia using Hybrid Features Extraction Techniques. *20th International Multi-Conference on Systems, Signals & Devices (SSD)* 925-933.

G. B. Moody and R. G. Mark (2001). The Impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine. 20, 45–50.

Hamidi, J. (2012). Application of Multilayer Perceptron Neural Network to Combine Economic and Emission Dispatch. International Journal of Soft Computing and Engineering, 1(6): 242-246.

Hassaballah, M., Wazery, Y. M., Ibrahim, I. E., & Farag, A. (2023). ECG Heartbeat Classification Using Machine Learning and Metaheuristic Optimization for Smart Healthcare Systems. Bioengineering. 10(4), 429.

Houssein E.H., Ibrahim E., Neggaz N, M. Hassaballah, Yaser M. Wazery, (2021). An efficient ECG arrhythmia classification method based on Manta ray foraging optimization. Expert Systems with Applications. 181, 115131

Kaya. Y, Pehivan. H, Tenekeci M. E (2017). Effective ECG beat classification using higher-order statistic features and genetic feature selection. Biomedical Research 28(17), 75.

Khafaga D.S, Alhussan A A, Abdelhamid A. A, Ibrahim A , Saber M, El-kenawy E. M. (2023). Dipper Throated Algorithm for Feature Selection and Classification in Electrocardiogram. Computer systems science and Engineering. 45 (2), 14720- 1482.

Kora, P. and Kalva, S.R. (2015). Improved Bat Algorithm for the Detection of Myocardial Infarction. *Springer Plus* 4, 666.

Kora. P, and Krishna, K.S (2016). Hybrid Firefly and Particle swarm Optimization Algorithm for the Detection of Bundle Branch Block. International Journal of Cardiovascular Academy, 2(1): 44 – 48.

Lin, C.C. and Yang, C.M. (2014). Heartbeat classification using normalized RR intervals and morphological features. Math. Probl. Eng. 12, 1–11.

Mahapatra, R.P and Singh, P.N. (2024). A novel deep learning approach for arrhythmia prediction on ECG classification using recurrent CNN with GWO. International journal of Information and Technology. 16, 577–585

Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014). Grey Wolf Optimizer. Advances in Engineering Software 69, 46-61. http://dx.doi.org/10.1016/j.advengsoft.2013.12.007

Nahiduzzaman, M., Nayeem, J., Ahmed, T., and Zaman, S. (2019). Prediction of Heart Disease using Multi-layer Perceptron Neural Network and Support Vector Machine. Proceedings of 4th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh.Pg 1-6.

Pan J., Tompkins W. J. A real-time QRS detection algorithm (1985). *IEEE Transactions on Biomedical Engineering.* 32(3),230–236.

Prystowsky E. and Padanilam, B. (2021). Cardiac Conduction Disorders. Cardiac Electrophysiology Clinics. 13(4), 1.

Rangayyan RM.(2002) Biomedical Signal Analysis: A case-study Approach. IEEE Press 2002.

Tunç, M., & Cangöz, G. B. (2024). Classification of the Cardiac Arrhythmia Using Combined Feature Selection Algorithms. Turkish Journal of Science and Technology. 19(1), 147-159.