International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I8P247

Volume 70 Issue 8, 463-474, August 2022

Original Article

Grasshopper Optimization for R-R Interval Selection and CBNN as Classifier

Parminder Kaur[1], Hardeep Singh Saini[2], Bikrampal Kaur[3]

[1]*Electronics Engineering, IKG Punjab Technical University, Punjab, India*
[2]*Electronics and Communication Engineering, IGCET, Abhipur, Punjab, India*
[3]*Computer Science Engineering, Chandigarh Engineering College, Landran, Punjab, India*

[1]p.kaur80@gmail.com

Received: 01 July 2022 Revised: 19 August 2022 Accepted: 24 August 2022 Published: 31 August 2022

Abstract - *Electrocardiogram has always been an area of motivation for researchers due to their significance in heart disease identification and classification. As different types of heart diseases result in spike changes in the ECG signal, R-R interval selection becomes crucial if the prediction must be done. This research article proposes a Swarm Intelligence-based Improved Grasshopper algorithm as RR-GHO. The grasshopper food selection policy introduces a novel grouping behaviour, and both the exploration and exploitation phases have been designed and implemented. In addition, a novel fitness function has been designed to improve the overall co-relation between the R-R intervals based on the intervals available in other groups. The optimized set has been trained using a Conjugate Based Neural Network, and the validation ratio has been kept at 70-30. The simulation has been done using MATLAB on an open-source MIT-BIH Arrhythmia dataset. The proposed algorithm architecture has also been compared with other research works in the same context based on the quantitative parameters, namely Precision, Recall, F-measure, and Accuracy. The accuracy of the proposed algorithm was improved by 11% compared to existing techniques.*

Keywords - *ECG Monitoring, Grasshopper Optimization, Neural Network, Swarm Intelligence.*

1. Introduction

Electrocardiogram (ECG) signals play a pivotal role in extracting information about Cardiac Arrhythmias [1]. Heart Arrhythmia is characterized by an irregular heartbeat caused by fluctuating electrical signals in the heart that can result in the heart beating too fast or too slow. Arrhythmias can further increase the risk of stroke and heart failure. Sudden deaths due to heart disease and Arrhythmias contribute to 15-20% of deaths globally, and the number is ever increasing, with major risk factors being High Blood pressure, diabetes, and other lifestyle factors like smoking, stress, and drug abuse [2]. Early diagnosis of Arrhythmias can lead to many lives being saved in the long run. Specifically, two types of irregularity patterns formed during the heartbeat classification: non-life-threatening and life-threatening. To diagnose non-life-threatening arrhythmias, recording the long pattern of ECG signals is vital to classify heartbeats. Cardiologists analyse the heart activity and rhythm using ECG displays. Though an ECG signal contains 5 peaks, namely P-Q-R-S and T but the identification of the disease in any classification method is based on R-R peak interval (distance between two adjacent R peaks) analysis. The PR segment is the base regarding which any deflection of the ECG signal is studied, and it represents atrial depolarization. The ventricular depolarization is represented by the QRS complex and should ideally be of a smaller duration. The T wave follows the QRS complex and is generally a small curve. The R-R peak interval in healthy individuals ranges from 0.6 to 1.2 seconds. Any variations in this help identify heart disease conditions [3].

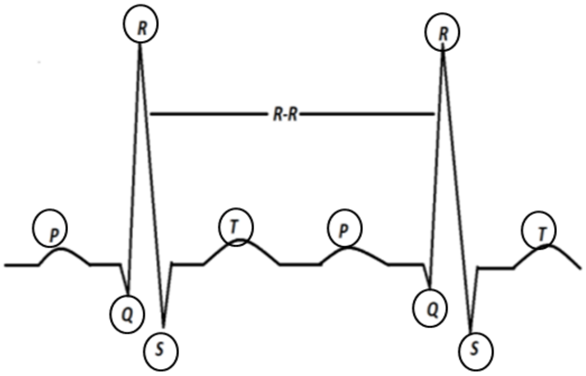


Fig. 1(a) ECG Signal with all peak

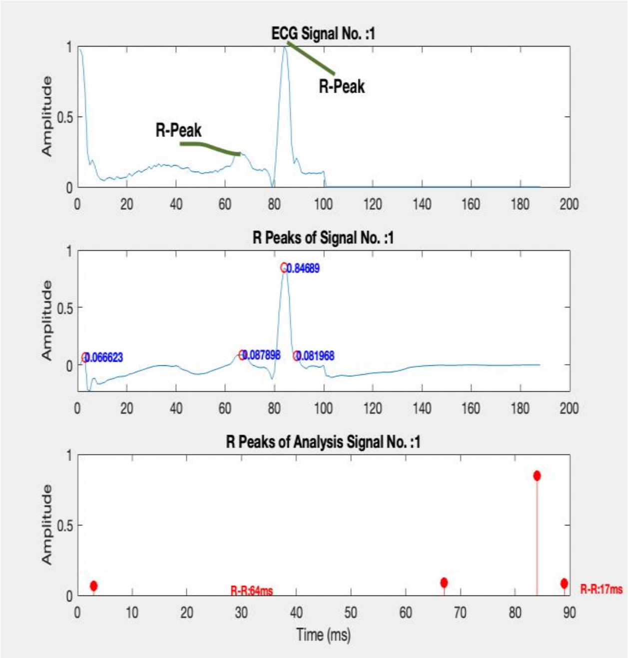


Fig. 1(b) ECG R Peak and R-R interval

Fig. 1(a) shows the position of all peaks, and Fig. 1(b) represents the real-time analysis of an ECG signal borrowed from the Kaggle dataset [4].

It is hard to analyze these peaks and identify the disease based on them manually when the amount of data is large; therefore, computational intelligence techniques help determine the disease using the ECG samples with less computation complexity and high accuracy if the data has high co-relation [5,6].

The classification of heart disease is based on the training data and its associated features. As discussed earlier, the R-R peaks are the most precise feature value for an ECG sample to predict disease. The general workflow of training and classification has been illustrated in Fig. 2.

Several processes are associated with the training and classification [7,8] based on any prediction mechanism for ECG, as shown in Fig. 2. The raw data can be extracted from various dataset repositories like Kaggle and UCI Machine Learning. The raw data is an ECG signal that comes labeled with a disease set. In the case of the proposed work, there are 5 classes of identification, as illustrated in the proposed methodology section. The ECG signal is passed for feature extraction, and in the case of the proposed work, it would be the R-R peaks, as shown in Figure 1(b). In the proposed work, R-R peak selection, feature selection plays a vital role in the training architecture. Many researchers have shown their interest in the optimisation process [9,10,11]. The optimization or selection algorithms work on the base of the co-relation that is formed in the features. The aim is to select optimal features from the given context and process them accordingly. Swarm Intelligence (SI) has been known to explore and extract the best suitable features related to the problem. In the case of ECG signals, many SI-based algorithms have been explored, as illustrated in Fig. 2. Grasshopper is the recent SI-based algorithm. It works on both exploitation and exploration architecture [12,13]. The proposed work has utilized Grasshopper as the base optimization algorithm and has incorporated new design behavior for the feature selection.

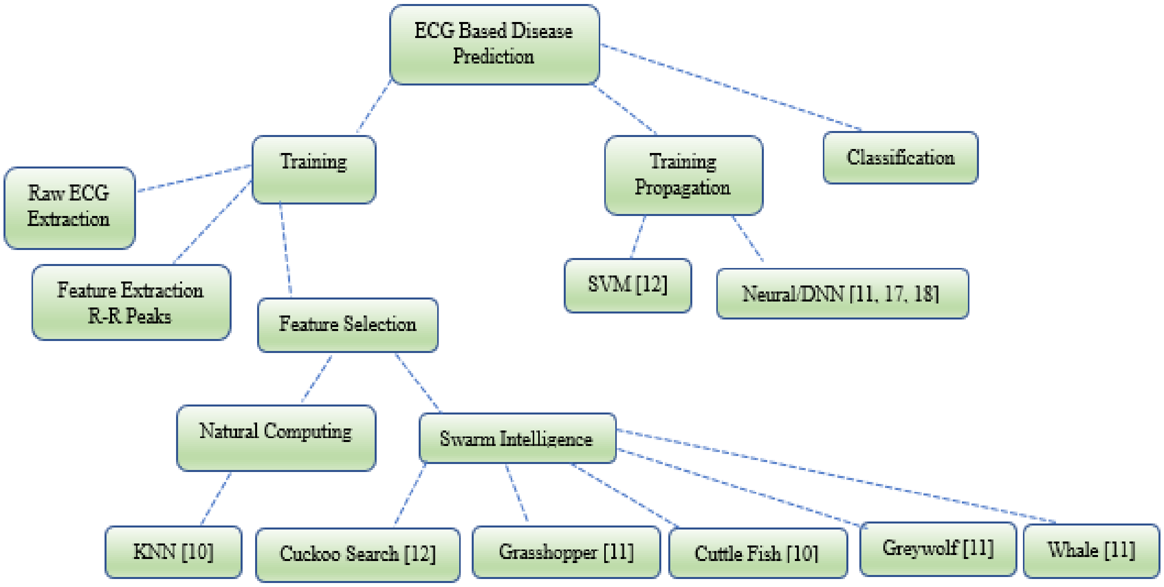


Fig. 2 General Processing Architecture

The optimal feature set is passed to the training algorithm in which several training mechanisms belong to Neural Network (NN) computing and Deep Neural Network (DNN) based computing, as shown in Figure 2. The classification mechanism takes the trained architecture against the supplied test data and classifies it with the same method that has been used for training [14,15]. The paper contribution is given as follows.

- Introduction of new novel food selection mechanism for Grasshopper Optimization Algorithm (GHOA)
- Introduction of a reward mechanism in GHOA
- Training and classification of more than 50000 data record with improved accuracy.

1.1. Problem Formulation

In hospitals, patients requiring even emergency care have to wait for long hours. The problem considered in this research is classifying the arrhythmia diseases in the healthcare application architecture. The modus of operation should be to quantify the patients based on different diseases. The motivation for the work comes from the humongous health issues of people worldwide. The main problem is organising the queue using the optimization technique and the CBNN classifier. The limitation of the CBNN classifier was avoided using the optimizer.

The organization of this research article is given as follows. Related work in which other researchers carried out the research was presented in section 2. The following section presents the proposed work incorporating the proposed GHOA optimization algorithm's algorithmic architecture. Section 4 illustrates the simulation results and conclusion presented in section 5.

2. Related Work

In the past, different techniques had been employed by the researchers considering the different classification techniques such as Long-short term memory (LSTM), Machine Learning (ML) models, Scaled Conjugate Gradient (SCG), etc.

Ahmad et al. [10] classify beats using ECG signals, and features were selected using an optimization algorithm such as Cuttlefish. The authors classified each beat's statistical and interval features, and five classes were used. The classifiers used in this paper were Scaled Conjugate Gradient and SVM, along with ANN and K-Nearest Neighborhood (KNN). The output results show that ANN accuracy is around 96%, 95% using KNN, and about 93% using SVM. The experimental analysis was tested using the MIT-BIH dataset, which recorded 1400 beats and 14 ECG signal records. The study is limited to providing the desired classification results with SVM.

Rajeshwari and Kavitha [11], the authors, used the Deep Neural Network approach for the classification of the ECG sample. The authors used different optimisation algorithms to extract the features, such as Grasshopper, whale optimization, and the grey wolf, which had been levied for the best solution. The proposed outcomes show that sensitivity and accuracy of about 96.95% and 98.11% using ensemble technique and whale optimization DNN provide 96.47% sensitivity. Outcomes show the betterment of ensemble classification with the DNN model.

Sharma et al. [12] considered the DWT technique for ECG signal classification. The authors used the SVM classifier to train and classify the signals into five classes. The selected features using the proposed method had been optimized using the Cuckoo search technique, and outcomes show that 98.93% accuracy had been obtained using the SVM with Neural Network and 97.95% with optimization technique, when tested on 3600 samples and outcomes, were improved using the NN in contrast to existing methods.

Kaya et al. [14] used the classifiers to detect the heartbeat, namely Premature Ventricular Contractions (PVC), from the ECG signals, and abnormality was detected using the time series approach. Further, the authors used dimension reduction techniques, and experiments analysis was presented using the Machine Learning (ML) models. Findings revealed that classification accuracy obtained using the classifiers is around 99.6%, sensitivity is about 99%, and specificity rates were about 99.8%. The main drawback of the presented study is the slow diagnosis process that makes the proposed system less reliable and inefficient.

Savalia et al. [15] considered the TensorFlow diagrams to classify heart disease. The authors used the four hidden layers, and then Multilayer Perceptron was used for training and testing the extracted data. The dataset in this paper was extracted from Kaggle, and accuracy was better compared to current techniques. But, the main problem is that training takes a lot of time, thus making the classification process long.

Rohmantri, and Surantha [17], the authors presented a method to classify arrhythmia disease using a Convolution Neural Network 2D Model using a variety of inputs having multiple sizes. The authors used the MIT-BIH dataset to classify the disease and transform the ECG peaks for the best accuracy. The accuracy for 2 classes is about 99%, around 98% for 7 classes, and for 8 classes, accuracy is around 98%. The limitation of the paper is that the CNN model is complex, and it is difficult to differentiate the classes for smaller sizes.

Ahmad et al. [18] used the scaled conjugate gradient (SCG) algorithm –ANN for the signal classification into five different classes. The work includes denoising signals for classification, statistical high order feature extraction, and the SCG algorithm was suitable for the recognition of each beat. The experimental results show that about 96% accuracy was obtained using 14 records considering the MIT-BIH arrhythmia dataset for about 1400 beats. But the main limitation of using the SCG algorithm in this paper is that there is no line search and needs a minimum *n* number of simulations to provide the required results, making the proposed system inefficient.

Yildirim et al. [19] used the LSTM classifier to recognise arrhythmia. The signal had been coded with the CNN autoencoder for automatic recognition. The authors used the MIT-BIH arrhythmia dataset, and outcomes revealed that the average root means square approaches 0.70% while accuracy was over 99%. The study significantly reduced the computational time and thus obtained high performance at a lower cost.

Ihsanto et al. [20] proposed an ECG authentication system considering the two stages: classification and beat detection. The authors used the Residual depth-wise CNN algorithm to classify the disease. The outcomes show that 48 patients had been tested, and 90 people were recorded for the ECG database. Thus, 100% accuracy has been obtained, and the proposed method outperforms the state of art techniques. Still, the study is limited to providing the desired results as the system is complex, making automatic recognition a cumbersome task.

Mathunjwa et al. [21] integrate the Recurrence plot (RP) and the other deep learning approach in two different stages. The authors used the 2D images and RP tool to explore the arrhythmia features for classification. The spatial information had been leveraged, and image classification was done to acquire better outcomes. R-peaks detection provides an accuracy of about 98.41% compared to other state-of-the-art techniques.

Huang et al. [22], the authors used the 2D CNN model to classify the disease, and time domain signals were used further to realize the time-frequency spectrograms. The classification method provides an accuracy of about 99%, and the learning rate is 0.01%; with the lowest loss, the 2D model was compared with the 1D model, which showed better accuracy of about 90.93% without additional pre-processing of the extracted ECG signal. The study drawback is that authors are limited to pre-processing the signal, such as filtering the noise and reducing features.

Xu et al. [24] developed an IoT-assisted monitoring system for automatically classifying ECG signals. The proposed system was lightweight and used the MIT-BIH dataset for experimentation. The authors also used the Physio Net database to recognize the physical activities in a real-time environment. The mathematical architecture was developed, and sensitivity using the MIT-BIH dataset was 99.78% and 96.45% in a real-time environment. The main drawback of the study was that there is still a need an improvement using the advanced algorithms.

The authors in [25] developed an IoT-based health monitoring system using sensors to avoid the current limitations of the health monitoring system. The authors monitor the real-time vital signs of the patient and monitor the body temperature, heart rate, and other parameters remotely. The drawback of the present study is that only two sensors were used to access the information. Moreover, there is a need for secure log-ins to access the patients' information and continuous monitoring.

The study in [26] implemented an ECG monitoring system for classifying abnormal and normal signals. The proposed system includes the noise removal system using filters and the Spartan 7 FPGA system for processing information. The simulation results showed that the accuracy of the proposed system was 99.92%, with a minimum power consumption of about 0.009W. Although the disease specification ratio was improved by 0.3%, the study's main limitation was processing time due to a large number of iterations.

3. Research Methodology

The proposed work is divided into the Optimal Training Section (OTS) and the Classification Section (CS). OTS is concerned and illustrative about the feature extraction, feature selection by the proposed GHOA algorithm, and running of the training engine to create a trained repository. In contrast, the CS is regarding the classification architecture of the proposed work. The dataset is illustrated in OTS to perform the proposed work, and data distribution is briefed in CS.

3.1. Optimal Training Section

The OTS section extracts the ECG features as the R-R intervals supplied in the dataset. Furthermore, it performs the R-R peak selection through the RR-GHOA algorithm that uses the R-R interval and modifies the selection procedure of the Grasshopper based on the co-relation of the intervals presented in the dataset.

3.1.1. Dataset

The dataset used for disease classification is the MIT-BIH Arrhythmia Dataset using five different classes: N, S, V, F, and Q [23]. The annotations of these classes are defined in [11]. The sampling frequency used is 125Hz, and the number of samples used is 109446.

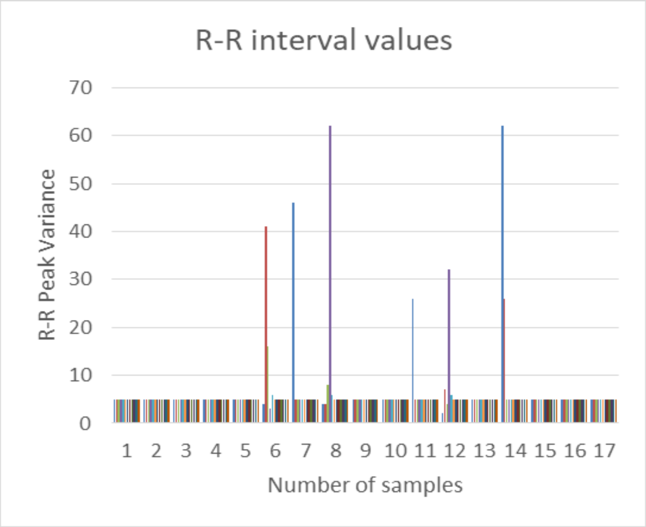


Fig. 3(a) Consistent R-R peaks for N and S disease sample

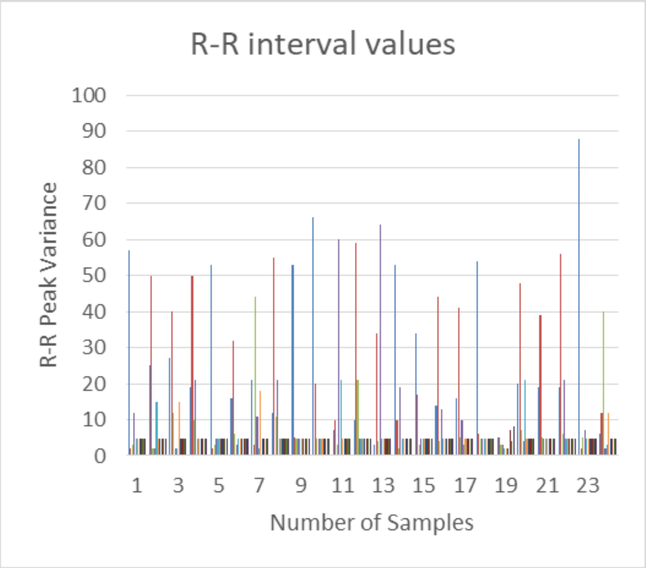


Fig. 3(b) Inconsistent R-R peaks for V, F, and Q

The data is quite co-related; hence, the R-R interval values are consistent in most cases with disease types N and S, but for V, F and Q, the values are not consistent, as shown in Figures 3(a) and (b).

Hence an R-R interval selection mechanism shows its relevance. Algorithm RR-GHOA describes the working of feature selection, whereas Figure 4 presents the overall workflow of the proposed model.

As shown in Figure 4, the raw dataset from Kaggle contains the ECG data for five categories. It is further supplied to feature extraction, where R-R intervals have been extracted. Post feature extraction, the data is aggregated as per their ground truth. GT. The proposed RR-GHOA plays its role in selecting the most relevant features that utilize the behaviour of the Grasshopper and introduces a new and novel fitness function that helps select the features based on the co-relation illustrated in the algorithm RR-GHOA. The selected features are then passed for data distribution, where 70% of the data is supplied for the training, whereas the rest, 30% data, will be used as the test sample. The training engine supports a Conjugate based Neural Network (CBNN), and the same engine is utilized to classify the test data. RR-GHOA inputs the RR interval and performs both the exploitation and exploration phase over the supplied data.

To perform the exploitation and exploration phases, the entire hopper set, viz., all R-R values has been divided into hopper fields equal to the total number of diseases in the dataset. With each field exists the centroid, which will be termed the global food value for the hoppers to survive in the list. The hoppers of the same class will be extracted, and a random sample of group values will be considered for the exploitation process. All the hopper's food viz. R-R intervals will be grouped with other hoppers of the same hypothetical group to perform the exploitation phase. A total of 10 levy flights will be lifted, out of which 5 levy flights will be for the exploitation phase, and 5 other levy flights will be for the exploration phase. The purpose of implementing both the exploitation and the exploration phase is that the peak values should dignify the original class and contribute to other classes' decision-making. In every levy flight, the hopper RR value will be grouped with other hoppers. It will be passed to the fitness function for exploitation, illustrated in fitness function exploitation. If the hopper satisfies the fitness function, the hopper will get a reward of 10 units, and if the hopper does not satisfy the fitness function, it will receive 0 rewards.

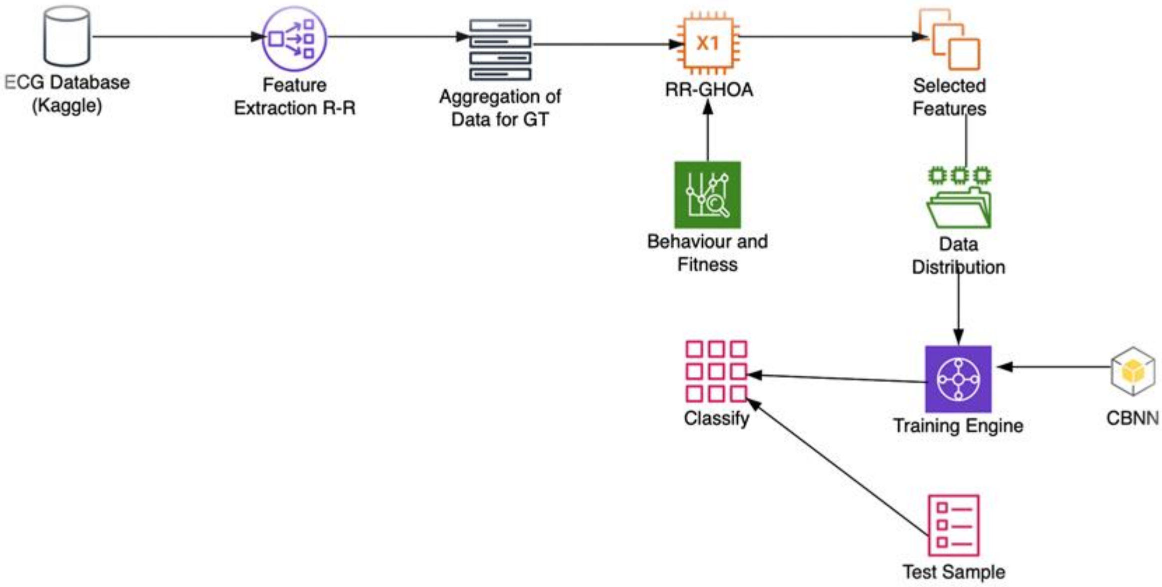


Fig. 4 Proposed Work Flow Diagram

In the exploration phase, as the hopper gets wings and can fly to other regions, the current hopper will be grouped with the hoppers of other indexes. The aim is to check the co-relation and the similarity values of the existing RR peak value with other RR peak values; hence, the exploration phase has been implemented. The fitness function is different for both the exploitation and the exploration phase.

Algorithm 1 Algorithm **RR-GHO**

```
Input: R-R intervals as $RR_A$ , Ground Truth of $RR_A$ as $GT_A$
Output: Selected R-R interval as $RRs$
$hf=GT_A.Classes;$ Hopper field are 5 as per dataset
$[h_{id},h_c]$=k-means($RR_A$,/$hf$ ) Divide the data into hypothetical fields where $h_{id}$ contains the centroid id of the hoppers and $h$, is the hypothetical centroid
$gt$= 1 Starting from first ground truth class
while $gt \le hf$ do
    hp= Find ($GT_A$.Classes ==$gt$) Hopper population
    $fs=RR_A.h_p$. Find hopper food as R-R intervals
    for i in $fs$ do
        $Fscore$=[]; Initiate Flight Score to Null
        $l_f$=5; 5 levy flights for exploitation
        $p_l$=1; Propagated flights
        while $p_l \le l_f$ do Exploitation Phase
            $n_p=h_p$.randomPop() Generate a random Population
            $n_p$.merge(i) Merge the current hopper to the population
            $Fs=RRs.n_p$ Collect the hopper food
            $Gf=h_c.n_p$. Extract the hypothetical centroid value as global food
            [$f_s$]= Hopper-Fitness-Exploitation ($F_s$, $G_f$); Score attained
            $p_l$++
            $Fscore[p_l]=f_s$;
        end while

        Exploration Phase Begins
        $l_f$=5; 5 levy flights for exploration
        $p_l$ = 1; Propagated flights
                $h_p$=find ($h_{id} \neq gt$) Find hoppers from other fields
        while $p_l \le l_f$ do Exploration Phase
```

```
$n_p$=$h_p$.randomPop() Generate a random Population
np.merge(i) Merge the current hopper to the population
$F_s$=$RRs.n_p$ Collect the hopper food
$G_f$=$h_c.n_p$. Extract the hypothetical centroid value as global food
 [$f_s$]= Hopper-Fitness-Exploration ($F_s$, $G_f$): Score attained
$p_l$++
Fscore[$p_l$+5] = fs: 5 flights have been covered in Exploitation
end while
th=50;
if( (ΣFscore)/$l_f$ )*2 $\geq$th
RRs .Merge(i);
End if
end for
        Return:  RRs
```

As the hopper viz., the interval value should be significant to its ground truth value. Still, it should also be insignificant for the other category to be most precise on classification architecture. The fitness function for exploitation is defined as follows. In contrast, the fitness function for exploration is just the reverse of exploitation, as the hopper has to avoid other group values to be most significant in the current group.

Algorithm 1 Fitness Function RR-GHO

```
Input: Total food of hoppers as Tf  Defines the food value
Field Centroids as Fc. Defines the hypothetical centroid by k-means
Rest Hopper as Rh. Defines hoppers other than evaluating hopper
Output f as fail or pass margin value attained by hopper
H fr=0;
where H f r represents Hopper food reliability
Hopper food reliability is the factor that denotes the situation that if the current hopper does not consume healthy food that is being the global best, the hopper is not going to survive
for i = 1 :Tf.Count
Hfr= Hfr+($Tf_i$-Fc)$^2$;
   End for i
Hfr= $\sqrt{}$Hfr
Gr=0 Global reliability
for j=1:Rh.Count
Gr=Gr+{$Rh_j$—Fc)$^2$;
   End f or j
Gr= $\sqrt{}$Gr;
   fc= ((Gr-Hfr)/Gr) * 1OO; food change
   if fc $\leq$70     f = 1;
      margin=fc;
      else
      margin=fc;
      end If
      Return f, margin
```

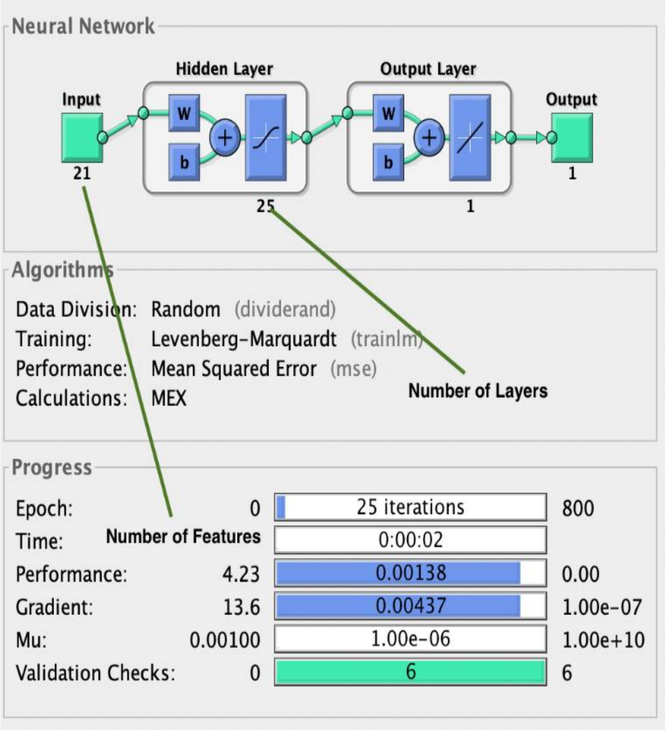
The proposed grasshopper algorithm provides selected R-R interval values supplied to CBNN, and the ordinal measures are defined in table 1 as follows.

Table 1. CBNN Propagation Architecture

Propagation Algorithm	**Conjugate-Based Neural Network**
Decision Criteria	Entropy
Stopping Criteria	Mean Squared Error (MSE)
Propagation Type	Polynomial
Layer Count	10-25

CBNN is a three-layer architecture that contains the Input Layer (IL), Hidden Layer (HL), and Output Layer (OL). The input layer takes the feature vector, R-R interval, in the case of the proposed work, along with the ground truth values that are diseases in the case of the proposed work. The HL is tuned based on the attained regression value supplied against each class. Fig. 5 illustrates the working of CBNN.

As shown in Fig. 5, 21 R-R intervals out of 24 have been depicted to be most relevant after applying the proposed algorithm. A total of 25 layers have been supplied, and one output architecture has been received. This output architecture will be used to classify the test data as illustrated as a part of the CS section.



Data Division: Random (dividerand)
Training: Levenberg-Marquardt (trainlm)
Performance: Mean Squared Error (mse)
Calculations: MEX

Epoch: 0 25 iterations 800
Time: 0:00:02
Performance: 4.23 0.00138 0.00
Gradient: 13.6 0.00437 1.00e-07
Mu: 0.00100 1.00e-06 1.00e+10
Validation Checks: 0 6 6

Fig. 5 Propagation architecture of CBNN

To be precise on the classification architecture, 70-30 data distribution, viz. 70% of data is being used for the training purpose, whereas the rest of the 30% of data is being used for the classification purpose. The data selection of training and classification has been kept as Levy distribution. The results have been evaluated based on the quantitative parameters and are illustrated in section 4.

4. Results and Illustrations

The current study uses the MIT-BIH dataset to classify the Arrhythmia diseases, considering the five categories, such as N, S, V, F, and Q, labelled 0,1,2,3 and 4, respectively, in the GT disease. N-Normal, S – premature atrial contraction, V-premature ventricular contraction, F- a fusion of ventricular and normal, Q – Fusion of paced and normal [23]. The dataset is free for the researchers. In the present study, MATLAB software was used to implement the results.

4.1. Performance Parameters

The performance metrics used in this study are accuracy, F-measure, recall, and precision, indicating the classification of diseases. The use of four performance metrics evaluates the performance of the classifier. True Positive (T_P) indicates the correctly classified signals, and false Positive (F_P) signifies the erroneous recordings. The proposed methodology ensures that positive value is predicted to reflect greater quality. The performance metrics are explained in the given equations.

$$\text{Accuracy} = \frac{T_P + T_N}{T_N + F_P + F_N + T_P} \qquad (1)$$

$$\text{Recall} = \frac{T_P}{T_P + F_N} \qquad (2)$$

$$\text{Precision} = \frac{T_P}{T_P + F_P} \qquad (3)$$

$$\text{F measure} = 2.\frac{P \times R}{P + R} \qquad (4)$$

Table 2 shows the computation of performance metrics such as Precision, Accuracy, Recall, and F-measure using the proposed methodology. The simulation results show that the precision for 1200 samples is around 85%. An increase in ECG samples also results in the variation in precision, recall, and F-measure value.

For 12000 test samples, there is about 90% precision, around 86% recall, and about 88% F-measure. For 15600 samples, the value for precision approaches 90.4%, recall approaches 86.7%, and F-measure reaches 88.3%. Consequently, for 18000 samples, the value for precision approaches 90.4%, recall approaches 86.6%, and F-measure reaches 88.5%. Thus, the average precision obtained using the proposed RR-GHOA is 87.6%, while that of recall is about 86.3%. The F-measure average value obtained using the proposed RR-GHOA is 86.98%, and the accuracy is 91%. The improved precision, recall, and accuracy values are due to the use of optimization techniques in the proposed methodology. Although, the shortcomings of the CBNN have been avoided using the grasshopper technique.

Table 2. Performance Metrics using the proposed technique

ECG Test Samples	**Precision**	**Recall**	**F-measure**	**Accuracy**
1200	0.855745041	0.845743002	0.850714623	90.47633735
2400	0.858144822	0.847865735	0.852974311	86.76681764
3600	0.857897398	0.843363707	0.850568473	87.25989852
4800	0.873033855	0.858030409	0.865467113	90.6090431
6000	0.874428545	0.845897858	0.859926618	90.48303807
7200	0.86369605	0.859108712	0.861396274	93.75473036
8400	0.856957247	0.868947009	0.862910482	87.68328312
9600	0.855119333	0.879230555	0.867007345	94.2847022
10800	0.896941779	0.88606368	0.891469546	91.23003371
12000	0.904162967	0.860263165	0.881666941	89.66288173
14400	0.899447831	0.896176375	0.897809123	88.79473172
15600	0.900256814	0.867356714	0.883500584	95.60159572
16800	0.86897743	0.866132592	0.867552679	93.01438947
18000	0.904255878	0.866700829	0.885080156	87.60232277

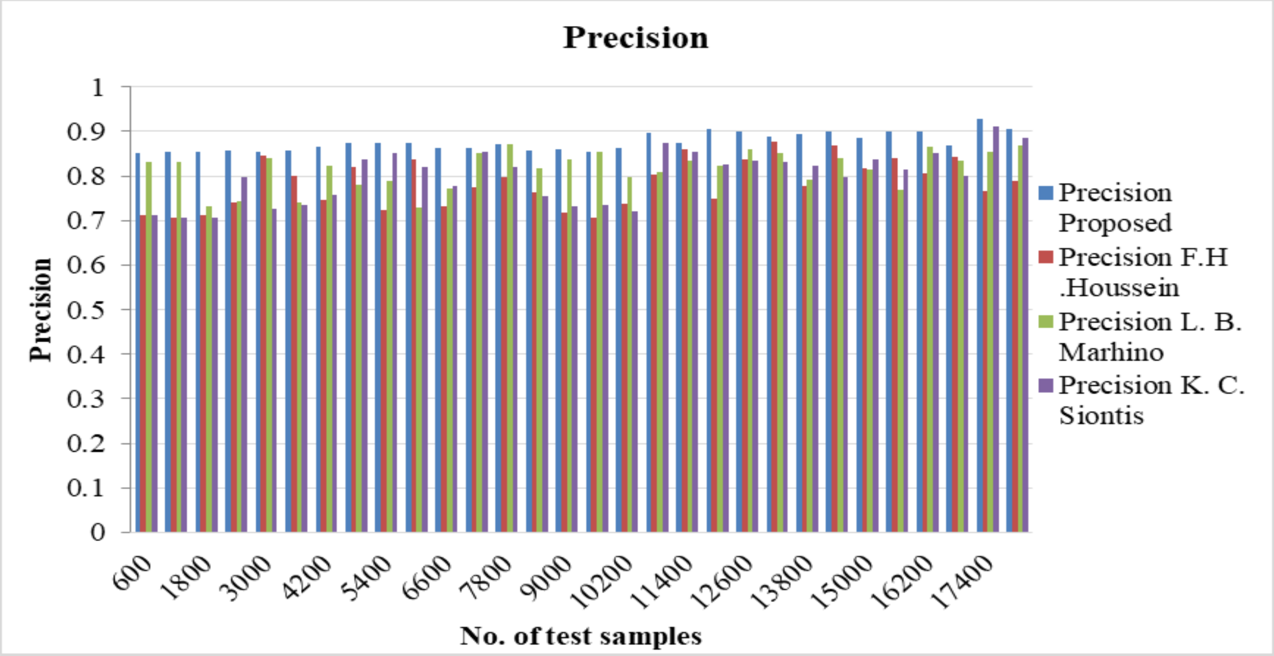


Fig. 6. Precision comparison with existing techniques

Fig. 6 shows the comparative analysis of the proposed approach for the computation of precision. The analysis results showed that the proposed technique's average value is 87.65% while existing approaches showed 78.38% [8] and 81.5% using [9]. The precision obtained using the technique proposed by [16] is about 80%. Thus, outcomes show that the proposed technique has been improved by 11.8%, 7.5%, and 9.6% in comparison to existing techniques [8], [9], and [16], respectively.

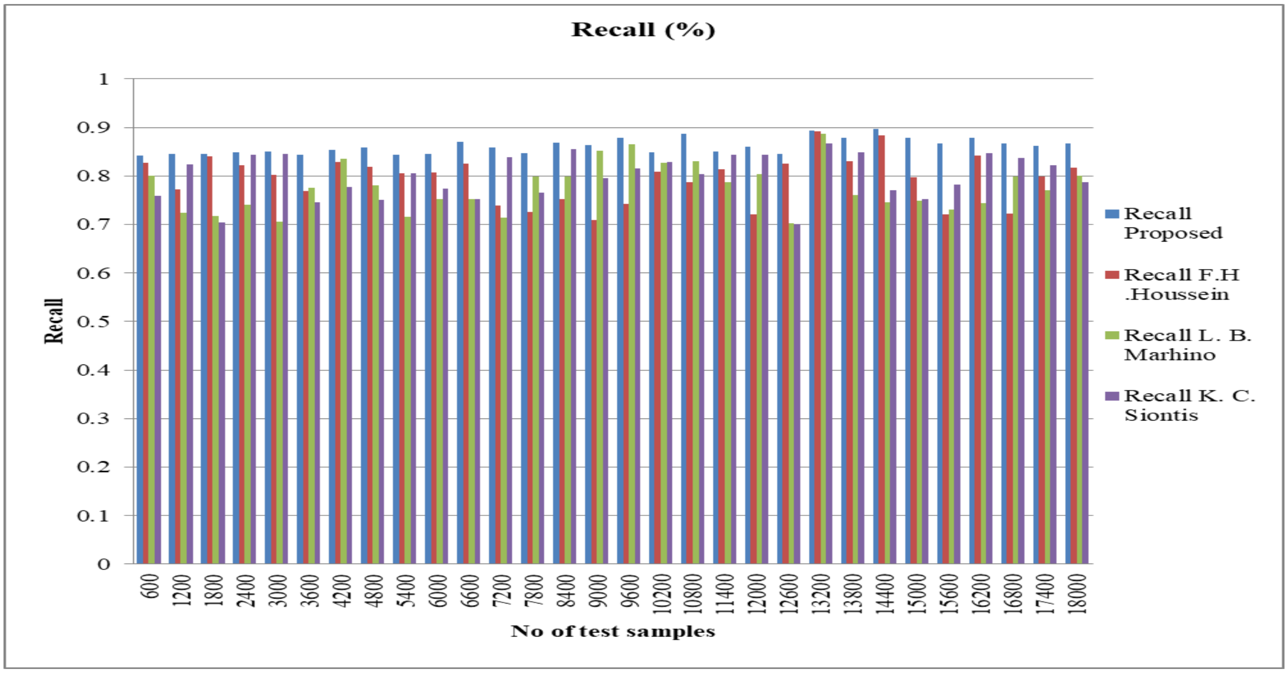


Fig. 7 Recall Comparison with existing techniques

Fig. 7 shows the comparative analysis of the proposed approach for recall computation. The analysis results showed that the average recall value using the proposed technique is 86.14%, while existing approaches showed 79.48% [8] and 77.5% using [9]. Similarly, recall obtained using the existing technique proposed by Siontis et al. is about 80%. Thus, there is an improvement of about 8.3%, 11%, and 7.7% in comparison to existing techniques [8], [9], and [16], respectively.

Fig. 8 shows the computation of the F-measure compared with the existing techniques. The analysis results showed that the average F-measure value using the proposed technique is 86.88%, while existing approaches showed 78.8% [8] and 79.3% using [9]. Similarly, the F-measure value obtained using the existing technique proposed by Siontis et al. is about 79.81%. Thus, the proposed approach has been improved by 10% from [8], 9.4% from [9], and 8.8% from the existing technique [16]. Thus, outcomes showed that the proposed technique is better than existing techniques.

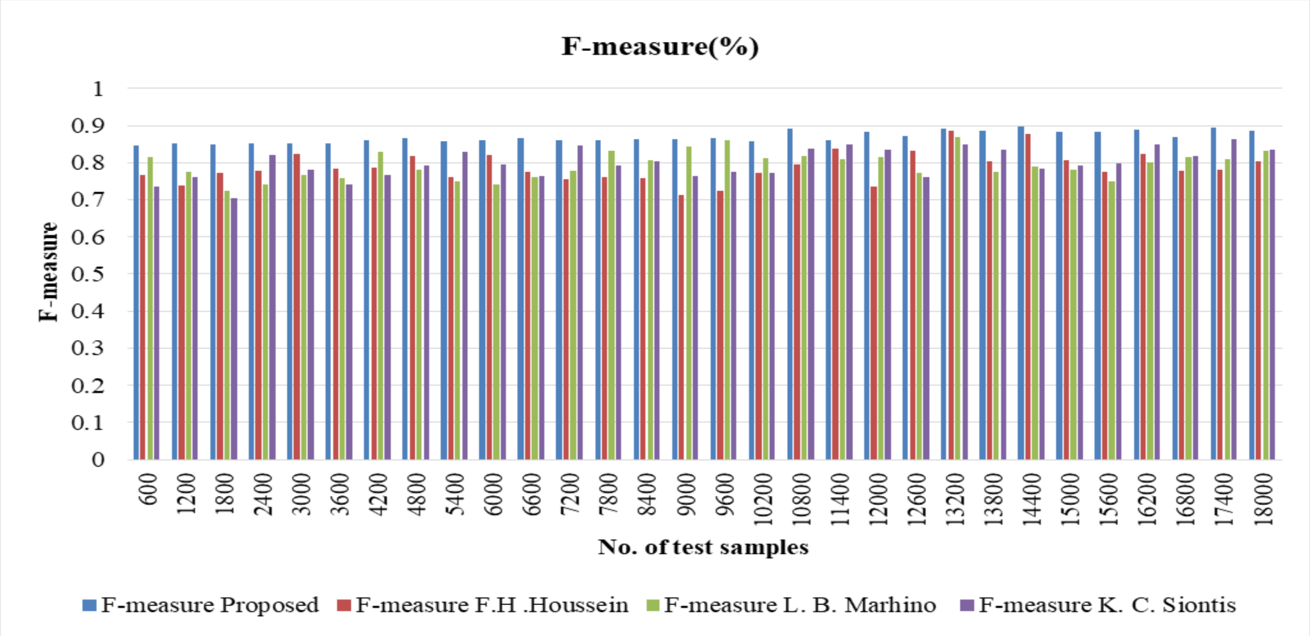


Fig. 8 F-measure comparison with the existing techniques

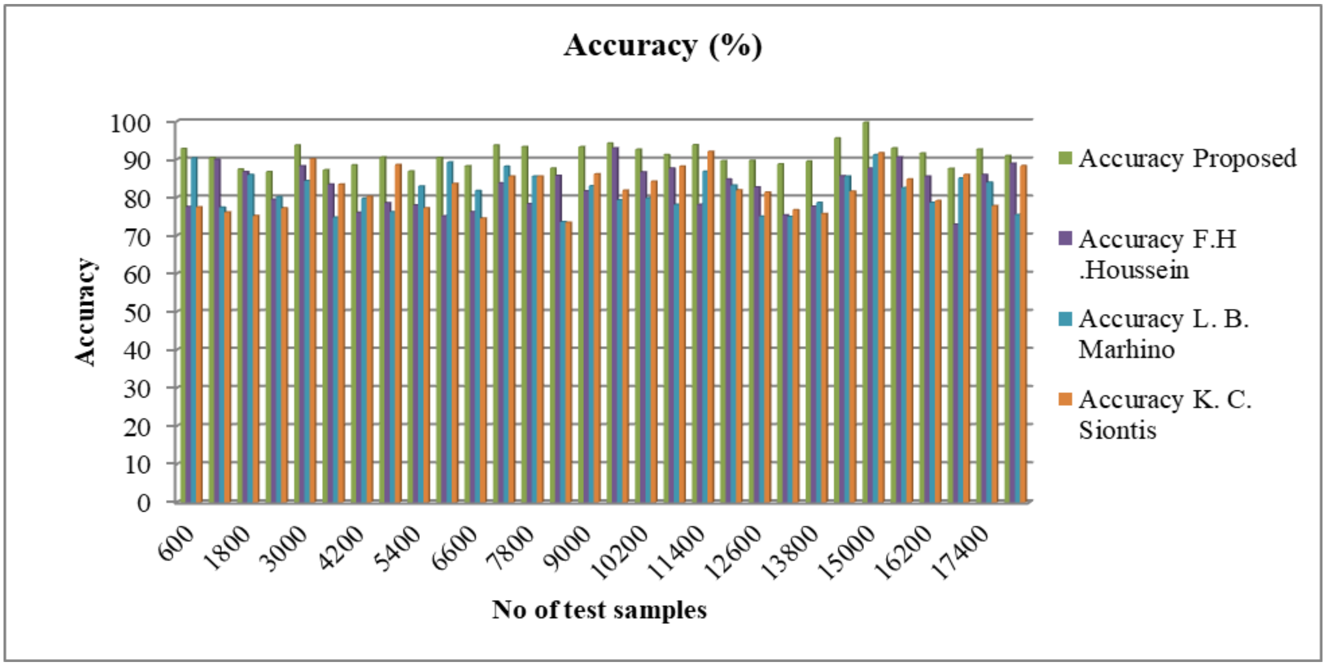


Fig. 9 Accuracy comparison with the existing techniques

Fig. 9 shows the computation of accuracy compared with the existing techniques. The analysis results showed that the average Accuracy value using the proposed technique is 91.07%, while existing approaches showed 82.76% [8] and 81.74% using [9]. Similarly, the accuracy value obtained using the existing technique proposed by Siontis et al. is about 82.21%. Thus, the proposed approach has been improved by 10% from [8] and [16] and 11% from [9]. Thus, outcomes showed that the proposed technique is better than existing techniques.

5. Conclusion

The paper illustrates an improved feature selection mechanism with one of the recently developed architectures of Swarm Intelligence, the Grasshopper Algorithm. The proposed algorithm takes the R-R interval of the ECG signals selected from the open-source Kaggle repository. As the intervals were not static across all the defined ground truth classes, an optimized RR-GHOA is being designed and explained in the algorithm architecture. The proposed algorithm presents a novel behaviour of RR interval selection and reduces the overall classification error compared to other state-of-the-art techniques discussed in the result section. CBNN has been applied to classify the data, which outcasts the other classification algorithms used by comparative researchers. The evaluation has been done on the base of quantitative parameters. The comparative analysis shows that the proposed approach has improved by 11.8% for precision, 10% for accuracy, 8.3% for recall, and 10% for F-measure. Thus, outcomes show that the proposed technique has improved compared to the existing state-of-the-art techniques.

References

[1] A. Tyagi, and R. Mehra, “Intellectual Heartbeats Classification Model for Diagnosis of Heart Disease from ECG Signal using Hybrid Convolutional Neural Network with GOA,” *SN App. Sci*., vol. 3, no. 2, pp. 1-14, 2021.

[2] T. N. Srinivasan and J. R. Schilling, “Sudden Cardiac Death & Arrhythmias,” *Arrhythmia and Electrophysiological Review*, pp. 101-117, 2018.

[3] A. Mahalakshmi, N.Nithya and P. Nandhini, “Diagnosis of Cardiovascular Diseases based on R-R interval using ECG Signals,” *International Journal of Engineering Research & Technology (IJERT)*,vol. 3, no. 1, pp. 2722-2728, 2014.

[4] R. He, Y. Liu, K.Wang, N. Zhao, Y.Yuan, H. Li, and Q. Zhang, “Automatic Cardiac Arrhythmia Classification using Combination of Deep Residual Network and Bidirectional LSTM,” *IEEE Acc*., vol. 7, pp. 102119-102135, 2019.

[5] D. Hooda, and R. Rani, “An Ontology Driven Model for Detection and Classification of Cardiac Arrhythmias using ECG Data,” *J. of Int. Inf. Sys*., vol. 58, no. 2, pp. 405-431, 2022.

[6] T. N. Nguyen, and T.H. Nguyen, “Deep Learning Framework with ECG Feature-Based Kernels for Heart Disease Classification,” *Elektronika Ir Elektrotechnika*, vol. 27, no. 1, pp. 48-59, 2021.

[7] J. Rahul, M. Sora, L. D. Sharma, and V.K. Bohat, “An Improved Cardiac Arrhythmia Classification using an RR Interval-Based Approach,” *Biocybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 656-666, 2021.

[8] E. H. Houssein, I. E. Ibrahim, N. Neggaz, M. Hassaballah, and Y.M. Wazery, “An Efficient ECG Arrhythmia Classification Method Based on Manta Ray Foraging Optimization,” *Exp. Sys. with App.*, vol. 181, pp. 115-131, 2021.

[9] L.B. Marinho, De M.M. Nascimento, N. Souza, J. W. M. Gurgel, M.V. Rebouc¸ P. P. Filho, de Albuquerque, V.H.C., “A Novel Electrocardiogram Feature Extraction Approach for Cardiac Arrhythmia Classification,” *Future Gener. Comput. Syst.*, vol. 97, pp. 564-577, 2019.

[10] A.S.S. Ahmad, M.S. Matti, A.S. Essa, O.A. ALhabib, and S. Shaikhow, S., “Features Optimization for ECG Signals Classification,” *Inter. J. of Adv. Comp. Sci. and Apps*., vol. 9, no. 11, pp. 383-389, 2018.

[11] M.R. Rajeshwari, and K.S. Kavitha, “Arrhythmia Ventricular Fibrillation Classification on ECG Signal using Ensemble Feature Selection and Deep Neural Network,” *Cluster Comp*., pp. 1-18, 2022.

[12] P. Sharma, S.K. Dinkar, and D.V. Gupta, “A Novel Hybrid Deep Learning Method with Cuckoo Search Algorithm for Classification of Arrhythmia Disease using ECG Signals,” *Neural Comp. and App.*, vol. 33, no. 19, pp. 13123-13143, 2021.

[13] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, “ECG Heartbeat Classification: A Deep Transferable Representation,” In *IEEE Inter. Conference on Healthcare Informatics (ICHI), IEEE*, pp. 443-444, 2018.

[14] Y. Kaya, and H. Pehlivan, H, “Classification of Premature Ventricular Contraction in ECG,” *Inter. J. of Adv. Comp. Sci. and App.*, vol. 6, no. 7, 2015.

[15] S. Savalia, V. Emamian, “Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks,” *Bioengineering*, vol. 5, no. 2, pp. 35-42, 2018.

[16] K. C. Siontis, P.A. Noseworthy, Z.I. Attia, and P. A. Friedman, “Artificial Intelligence-Enhanced Electrocardiography in Cardiovascular Disease Management,” *Nature Rev. Card.*, vol. 18, no. 7, pp. 465-478, 2021.

[17] R. Rohmantri, and N. Surantha, N., “Arrhythmia Classification using 2D Convolutional Neural Network,” *International Journal of Adv. Comput. Sci. Appl*., vol. 11, pp. 201-208, 2020.

[18] A.S.S. Ahmad, A. S. S., Matti, M. S., Al-Habib, O. A., and Shaikhow, S, “ECG Signal Classification using Scaled Conjugate Gradient Learner Algorithm,” *Inter. Journal of Med. Res. & H. Sci*., vol. 7, no. 5, pp. 11-17, 2018.

[19] O. Yildirim, U.B. Baloglu, R.S. Tan, E.J. Ciaccio, U. R. Acharya, “A New Approach for Arrhythmia Classification using Deep Coded Features and LSTM Networks,” *Comput. Methods P. Biomed*., vol. 176, pp. 121-133, 2019.

[20] E. Ihsanto, K. Ramli, D. Sudiana and T.S. Gunawan, “An Efficient Algorithm for Cardiac Arrhythmia Classification using Ensemble of Depthwise Separable Convolutional Neural Networks,” *Appl. Sci.*, vol. 10, no. 2, pp. 483-501, 2020.

[21] B.M. Mathunjwa, Y.T. Lin, C.H. Lin, M.F. Abbod and J.S. Shieh, “ECG Arrhythmia Classification by using a Recurrence Plot and Convolutional Neural Network,” *Biomed. Signal P. Con.*, vol. 64, pp. 102-122, 2021.

[22] J. Huang, B. Chen, B. Yao, and W. He, “ECG Arrhythmia Classification using STFT-Based Spectrogram and Convolutional Neural Network,” *IEEE Access*, vol. 7, pp. 92871-92880, 2019.

[23] [Online]. Available: https://www.kaggle.com/datasets/shayanfazeli/heartbeat.

[24] G. Xu, “IoT-Assisted ECG Monitoring Framework with Secure Data Transmission for Health Care Applications,” *IEEE Access*, vol. 8, pp. 74586-74594, 2020.

[25] S. Ahammed, N. Hassan, S.H. Cheragee, and A.Z.M.T. Islam, “An Iot-Based Real-Time Remote Health Monitoring System,” *International Journal of Research Engineering Sci*., vol. 8, no. 3, pp. 23-29, 2021.

[26] D.D.Salam, and D.L.S. Singh, “Cardiovascular Disease Classification using ECG Signal,” *International Journal of Engineering Trends and Tech,* vol. 69, no. 11, pp. 134-139, 2021.