CHAPTER 8

Moth-Flame Optimization-Based Deep Feature Selection for Cardiovascular Disease Detection Using ECG Signal

Arindam Majee and Shreya Biswas

Jadavpur University

Somnath Chatterjee

Future Institute of Engineering and Management

Shibaprasad Sen

University of Engineering and Management

Seyedali Mirjalili

Torrens University Australia
Yonsei University

Ram Sarkar

Jadavpur University

DOI: 10.1201/9781003205326-10

CONTENTS

8.1 INTRODUCTION

Cardiovascular diseases (CVDs), such as Coronary, Cerebrovascular, Peripheral, Ischemic, Hypertensive, Congenital, Rheumatic, and Non-rheumatic valvular, are considered as the main reason of fatality with almost 17.9 million deaths each year. It accounts for almost 31% of all deaths worldwide [1]. In the United States, affecting most ethnic groups, it is one of the main diseases present [2]. The electrocardiogram (ECG) is used as a major tool in diagnosing CVDs. Automated analysis of standard ECG gained paramount importance in order to save time and effort, as the world transitioned from analog to digital. ECG is usually performed when a patient experiences acute chest pain following which the treatment can be immediately determined [3].

However, the problem is that many physicians solely rely on elemental diagnostic analysis of ECG results whereas they always require the understanding and confirmation by a trained technician [4].

On the other hand, Deep Learning (DL) has achieved remarkable success in medical diagnosis tasks [5,6] and has the potential to improve health care and clinical practice on a large scale [7]. Though an expert's confirmation is probably required in many clinical settings, DL can help an expert in the clinical environment. Studies show that Supervised Learning can perform better than a human specialist in medical testing and diagnosis [8,9]. However, efficient training of Deep Neural Networks (DNNs) requires a big dataset which, for medical applications, is very scarce—mainly due to confidentiality issues [10].

The ECG is the most common evaluation technique of the heart providing a proper evaluation of the patient's heart activity – including the cardiac rhythm, repolarization, arrhythmias, coronary syndromes, and effects of drugs. Thus, an automatic DL approach interpreting ECGs can be useful for better diagnosis of the patients.

ECG analysis has been done using DL in recent works [11], in which the authors trained DNNs on a fairly large-sized dataset consisting of 91,232 single-lead ECGs achieving a ROC area of 0.97. In Ref. [12], the authors used DNN to train the large publicly available PhysioNet Challenge dataset and achieved better performance when compared to that of cardiologists. Out of a total of 75 teams that entered the challenge, 4 teams won with an F1 score of 0.83.

The standard short-duration 12-lead ECGs are often performed in healthcare units, often with no specialists to analyze the ECG signals. Also, doctors during their training may lack a complete understanding of these tracings [13]. The automatic yet precise interpretation of ECG is the need of the hour with CVDs increasing at an alarming rate—owing to the unhealthy lifestyles that most people are incorporating nowadays [14]. In countries where maximum deaths are related to CVDs and people often do not have access to trained cardiologists, this can be a successful alternative.

The benefits of DNN usage for ECG evaluation are still largely unexploited due to the shortage of medically accurate digital ECG databases [15]. According to the authors of Ref. [16], there are very few databases with a large number of ECG tracings and their standardized explanation, hence limiting their effectiveness as training datasets for the DL methods which follow the supervised learning approach. Despite this disadvantage, the results that DL models produce are better than the analysis of most physicians.

In light of the above facts, under this procedure, we have initially applied a pre-trained Deep Convolutional Neural Network (DCNN) model—VGG16 for arrhythmia classification. Though this gives decent outcomes, the usage of DL models involves large computation cost and requires huge training data so as to generalize over new samples. However, we have used Transfer Learning (TL) in order to extract main features from the data for further processing. But the sizable features so produced may have some redundant and unessential features. Therefore, these obtained features need to be optimized by some means to reduce the computation time and storage requirement apart from the reduction of the redundancy of the features.

For dimension reduction of the feature vector, we have applied a swarm intelligence-based metaheuristic algorithm, the Moth-Flame Optimization (MFO) algorithm, which was originally proposed in Ref. [17]. This optimal subset of the feature set produced by tuning the parameters of MFO was then fed into a Support Vector Machine (SVM) classifier [18] to detect arrhythmia accurately.

Research highlights of this work are listed below:

1. Combining a DCNN model with a nature-inspired metaheuristic feature selection algorithm to classify ECG signals for identification of the CVDs—hence creating a two-staged approach architecture.
2. To manage the overfitting problem, we have used TL. The initial weights are generalized to get the full benefit of TL. Another use of this is to reduce the training time. Training a large model like VGG16 from scratch would have taken a lot of time.
3. For dimension reduction of the feature set that we got from the original DCNN model, a popular metaheuristic, called the MFO algorithm, has been used.
4. The proposed model gained state-of-the-art results when evaluated on the publicly available MIT-BIH Arrhythmia database. The optimized feature subset results in better classification accuracy than the non-optimized feature set produced by VGG16.

The remaining portion of this chapter has been divided into the following sections: Section 8.2 contains the related work in this field of ECG classification and diagnosis using DL methods. Section 8.3 mentions the motivation behind this work. Section 8.4 provides information about the dataset we have used in this work. Section 8.5 explains the pre-processing we have applied in the MIT-BIH dataset to obtain the image data. Section 8.6 contains the methodology and Section 8.7 explains the results we have obtained and also compares our work with some state-of-the-art works in this area. Section 8.8 concludes this chapter by discussing some future prospects to work on.

8.2 RELATED WORK

Here, we briefly discuss some past works related to the said research topic. In Ref. [19], the authors proposed an algorithm to classify noisy ECG signals from the MIT-BIH Arrhythmia database [20] using a one-dimensional CNN. The authors had applied two approaches. In one approach, a five-class classification from the 1D ECG data is obtained using 1D CNN, achieving about 97.4% accuracy. In another approach, they converted the 1D ECG data into 2D gray-scale image data and then used an eight-class classification on those images using a 2D CNN, achieving a classification accuracy of 99.02%.

In Ref. [21], the authors used Marine Predators Algorithm (MPA) with CNN, naming the method MPA-CNN, to classify four different types of arrhythmia. They worked on three separate datasets—the MIT-BIH, the European ST-T dataset (EDB) [22], and the St Petersburg INCART [23]. They obtained precisions of 99.31%, 99.76%, and 99.47% on the datasets, respectively.

The authors of the work reported in Ref. [24] proposed hybrid feature extraction of T-wave for arrhythmia classification on the MIT-BIH database. They used the windowing technique, feature extraction, followed by classification for the purpose, and obtained an accuracy of 98.3%, specificity of 98.0%, and sensitivity of 98.6%. The authors of Ref. [25] used cross-correlation feature extraction. A Least-Square SVM classifier was used for a three-class classification on the MIT-BIH database. The accuracy achieved was about 96%.

In Ref. [26], the authors proposed a 1D DNN model to correctly classify three different arrhythmias on the MIT-BIH database, achieving a 97.44% accuracy. In Ref. [27], the authors presented a classification system designed for detecting Ventricular Ectopic Beats. Accuracy–sensitivity performances of the system were 98.3%–84.6% and 97.4%–63.5%, respectively.

The authors of the work [28] used a DCNN model to classify five different arrhythmias and Myocardial Infarction (MI) from the MIT-BIH and PTB Diagnostics datasets. They obtained 93.4% and 95.9% accuracy on each dataset. The authors of Ref. [29] used Particle Swarm Optimization (PSO) to enhance the classification results of the SVM classifier using data from the MIT-BIH dataset. PSO gave an 89.72% accuracy.

The authors of Ref. [30] used wavelet-transformation on waveforms from 18 files of the MIT-BIH database. They generated the feature set for classification and obtained an accuracy of 95.16%–96.82%. In the work [31], the authors used a deep 2D CNN model for arrhythmia classification, achieving 99.05% accuracy and 97.85% sensitivity.

Recently, many Swarm intelligence optimization techniques (inspired by animal groups and insect colonies, mimicking their behavior) have been proposed and tested for general and medical diagnosis [32]. Robustness, flexibility, and the ability to quickly find the optimal solution to a particular problem are some of the usefulness of these metaheuristic algorithms [33]. Genetic Algorithm [34], Ant Colony Optimization [35], PSO [36], Artificial Bee colony optimization [37], and Cuckoo Search [38] are a few of them. Some of the applications of swarm intelligence-based optimization algorithms in the medical field deal with cancer screening [39,32], the fusion of MRI and CT scans [40], endocrinology [41], tumor classification [42], and so on.

8.3 MOTIVATION

MIT-BIH dataset has been used in this study because of its large amount of data as a DL model needs to properly train in order to correctly predict ECG classes. We have chosen a DL-based model for initial feature extraction as it has been seen that DL models perform better than machine learning-based models as the former can learn complex patterns automatically from the raw inputs. The choice of the MFO algorithm as the chosen metaheuristic algorithm is because it is a population-based algorithm with a local search strategy. This helps to find out the best possible solution using global and local exploitation [43]. MFO has few parameters, is flexible and easy to implement, and has faster convergence. Hence, it is being used to solve many problems such as parameter estimation [44], classifications [18], medical diagnoses [45], and image processing [46]. Other applications of MFO have been elucidated in Ref. [47].

8.4 DATASET USED

We have used the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia Database [20] that has 48 properly annotated records (each with a duration of 30 minutes and sampled at 360 samples per second with 11-bit resolution over a 10 mV voltage range). This dataset gives a proper view of all important waves present in an ECG signal, including P-waves, Q-waves, R-waves, S-waves, and T-waves. There are almost 15 labels, including 1 normal class. Among these labels, some are unclassified and some are normally not used in Arrhythmia detection. We have chosen to work with the MIT-BIH database because it provides a properly annotated collection of normal and several different types of arrhythmia to train our classifier.

For this work, we have collected the required arrhythmia recordings from the database which has more than 110,000 ECG beats with almost 16 different types of arrhythmia and 1 normal. As we have already discussed in Section 8.2, past authors have done arrhythmia classification tasks with different classes. After surveying previous works, we have decided to move with an eight-class classification as this contains normal and the most common Arrhythmia classes. From the MIT-BIH database, we have considered eight different ECG beats. These are normal beats (NOR) and seven different types of ECG arrhythmias. These seven different types of Arrhythmia are as follows: Premature Ventricular Contraction (PVC), Paced Beat (PAB), Right and Left Bundle Branch Block Beat (RBB and LBB), Atrial Premature Contraction (APC), Ventricular Flutter Wave (VFW), and Ventricular Escape Beat (VEB). We have ignored some less important beats in ECG arrhythmia classification studies like non-conducted P-wave and unclassifiable beats.

8.4.1 Pre-Processing

ECG signal is a time-based signal. To use these ECG 1D data in a 2D CNN model, we need to convert these into image data. We have followed a pre-processing strategy for this conversion. We plotted every ECG signal as an individual 128×128 gray-scale image and transformed them into corresponding ECG images. The MIT-BIH dataset is sampled with a frequency of 360 Hz, and the label is specified during the heartbeat's Q peak. So, each signal is segmented based on the Q-wave peak time and converted into gray-scale images after plotting. Thus, each ECG beat's Q-wave peak is kept at the center and the first and last 20 sampled values from the previous and afterward Q-wave peak signals are excluded as the duration of the Q-wave peak is 0.03 s or less. So, in the time domain, we can define the range of a single ECG beat as follows;

$$T\left(\left(Q_{\text{peak}}\right)(n-1)+20\right) \leq T(n) \leq \left(\left(Q_{\text{peak}}\right)(n+1)-20\right)$$

Following the above formula, we have excluded the first and last ECG beats. Figure 8.1 shows some images obtained after pre-processing. As a result, we have obtained a total of 1,07,620 images belonging to eight different classes. Then, we have divided this dataset into training, validation, and testing set in a ratio of 70%, 15%, and 15%. The number of images in every class is presented in Table 8.1.

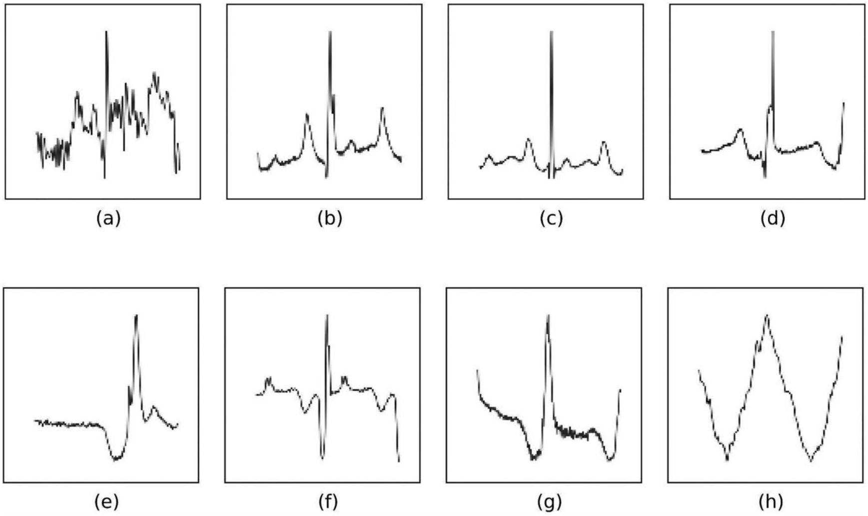

FIGURE 8.1 ECG signal Images formed after pre-processing. A sample from all eight classes is shown here: (a) APC, (b) LBB, (c) normal beat (NOR), (d) RBB, (e) PAB, (f) PVC, (g) VFW, and (h) VEB.

TABLE 8.1 Division of the Dataset—A Total of Eight Classes—into Train Set, Test Set, and Validation Set

	No. of Samples in		
Class	**Training Data**	**Validation Data**	**Testing Data**
APC	1,760	382	402
LBB	5,630	1,204	1,238
Normal	52,509	11,250	11,257
PAB	4,937	1,036	1,051
PVC	4,983	1,097	1,050
RBB	5,095	1,088	1,073
VEB	78	16	12
VFW	342	70	60
Total	75,334	16,143	16,143

8.5 METHODOLOGY

Before using the MFO algorithm, we have first extracted the features from the images. We have employed the VGG16 model pre-trained on the ImageNet dataset as our feature extractor model. After freezing the top layers up to block 5, we have fine-tuned the model on our ECG dataset. The pre-trained VGG16 model accepts three channel input images. So, we

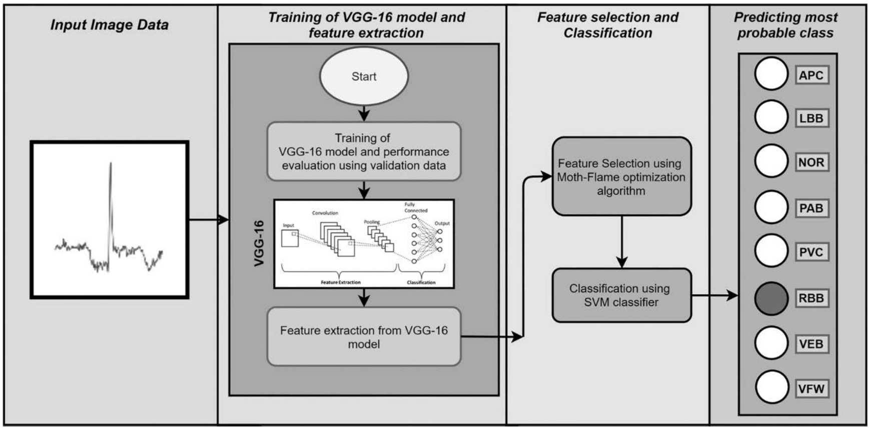

FIGURE 8.2 Block diagram representing the proposed methodology.

have converted the gray-scale images obtained after pre-processing into three-channel (RBG) images. The block diagram depicting the flow of our proposed methodology can be seen in Figure 8.2.

8.5.1 VGG16: A Brief Overview

In the first stage of the procedure, we have applied VGG16 for feature extraction on our dataset. The initial layers of this model are initialized with ImageNet weights [48], whereas the final layers are made trainable to modify their weights according to the images in our dataset. VGG [49], proposed in 2014, has its main aspect in its small kernel size of (3×3). This feature enables VGG to apprehend provincial features, thus improving the performance. Figure 8.3 shows the VGG16 model we have used in this work after inserting a few supplementary layers for achieving better results fine-tuned especially for our purpose.

8.5.2 Moth-Flame Optimization

MFO is a nature-inspired metaheuristic algorithm simulating the movement of moths, and this movement is commonly known as the Transverse Orientation. While traveling at night, a moth maintains a fixed angle with the Moon (a bright source of light) as a result of which it stays in a straight line. But in most cases, the moths keep on spiraling to the source until they get exhausted. This is because these moths are fooled by the presence of artificial lights and they try to follow a similar movement method around them. Being extremely close to the Moon, these lights resulted in a spiral path for the moths. Figure 8.4 shows diagrammatically the spiral motion.

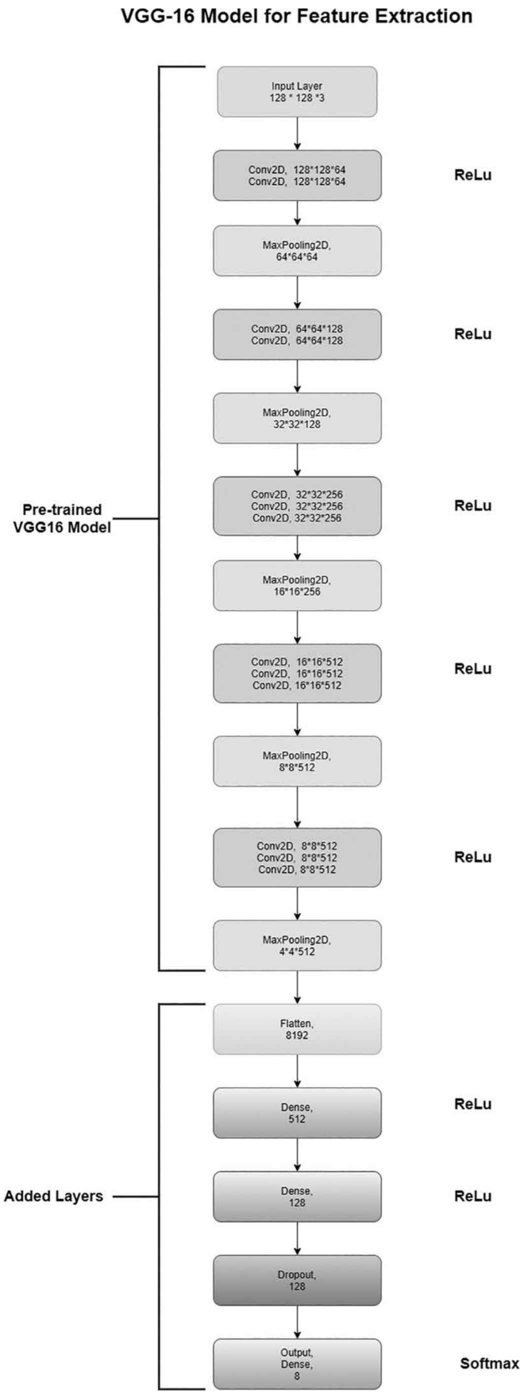

FIGURE 8.3 VGG16 model used for feature extraction from ECG records.

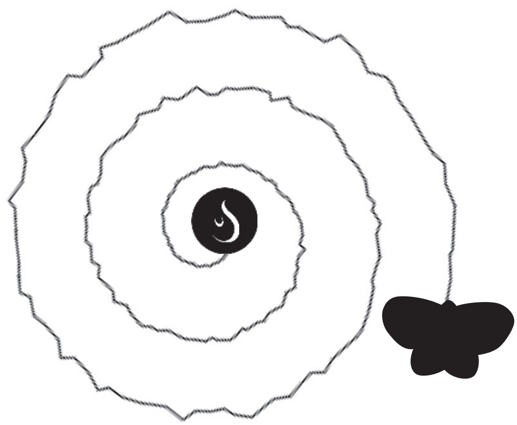

FIGURE 8.4 Movement of moths around an artificial light following the transverse orientation ultimately leading to a spiraling motion.

The MFO algorithm can solve many complex optimization problems. For this, the moths and the artificial light sources (flames) are assumed to be the solutions to the problem, and the position of the moth in space becomes the variable in any number of dimensions. We represent the moth population as a matrix M:

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & \cdots & m_{1,d} \\ m_{2,1} & m_{2,2} & \cdots & \cdots & m_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & \cdots & m_{n,d} \end{bmatrix} \tag{8.1}$$

where n is the total number of moths and d the number of dimensions in which the moth travels.

A problem-specific fitness function is defined and an array OM stores the corresponding fitness values. As there are n number of moths, there are n fitness values in the array, as shown below:

$$\mathrm{OM} = \begin{bmatrix} \mathrm{OM}_1 \\ \mathrm{OM}_2 \\ \vdots \\ \mathrm{OM}_n \end{bmatrix} \tag{8.2}$$

Another matrix *F*, with dimensions same as that of *F*, stores the best-obtained solutions so far as flames, and a matrix OF stores the fitness values of the individual flames as shown in the below two equations:

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & \cdots & F_{1,d} \\ F_{2,1} & F_{2,2} & \cdots & \cdots & F_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ F_{n,1} & m_{n,2} & \cdots & \cdots & F_{n,d} \end{bmatrix} \quad (8.3)$$

$$\text{OF} = \begin{bmatrix} \text{OF}_1 \\ \text{OF}_2 \\ \vdots \\ \text{OF}_n \end{bmatrix} \quad (8.4)$$

The MFO algorithm is defined as follows:

$$\text{MFO} = (I, P, T) \quad (8.5)$$

Here, the function *I* generates a random population of moths and their corresponding fitness values:

$$I:\ \varnothing \rightarrow \{M, \text{OM}\} \quad (8.6)$$

The function *P* returns the updated positions of the moths:

$$P: M \rightarrow M \quad (8.7)$$

The function *T* returns true if the stopping condition is satisfied:

$$T: M \rightarrow \{\text{True}, \text{False}\} \quad (8.8)$$

S is the spiral function (usually a logarithmic spiral) using which the position of a moth with respect to a flame is updated and stored in *M*.

$$M_i = S(M_i, F_j) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j \quad (8.9)$$

where D_i indicates the distance of the i^{th} moth for the j^{th} flame, *b* is a constant, and *t* is a random number in [–1,1].

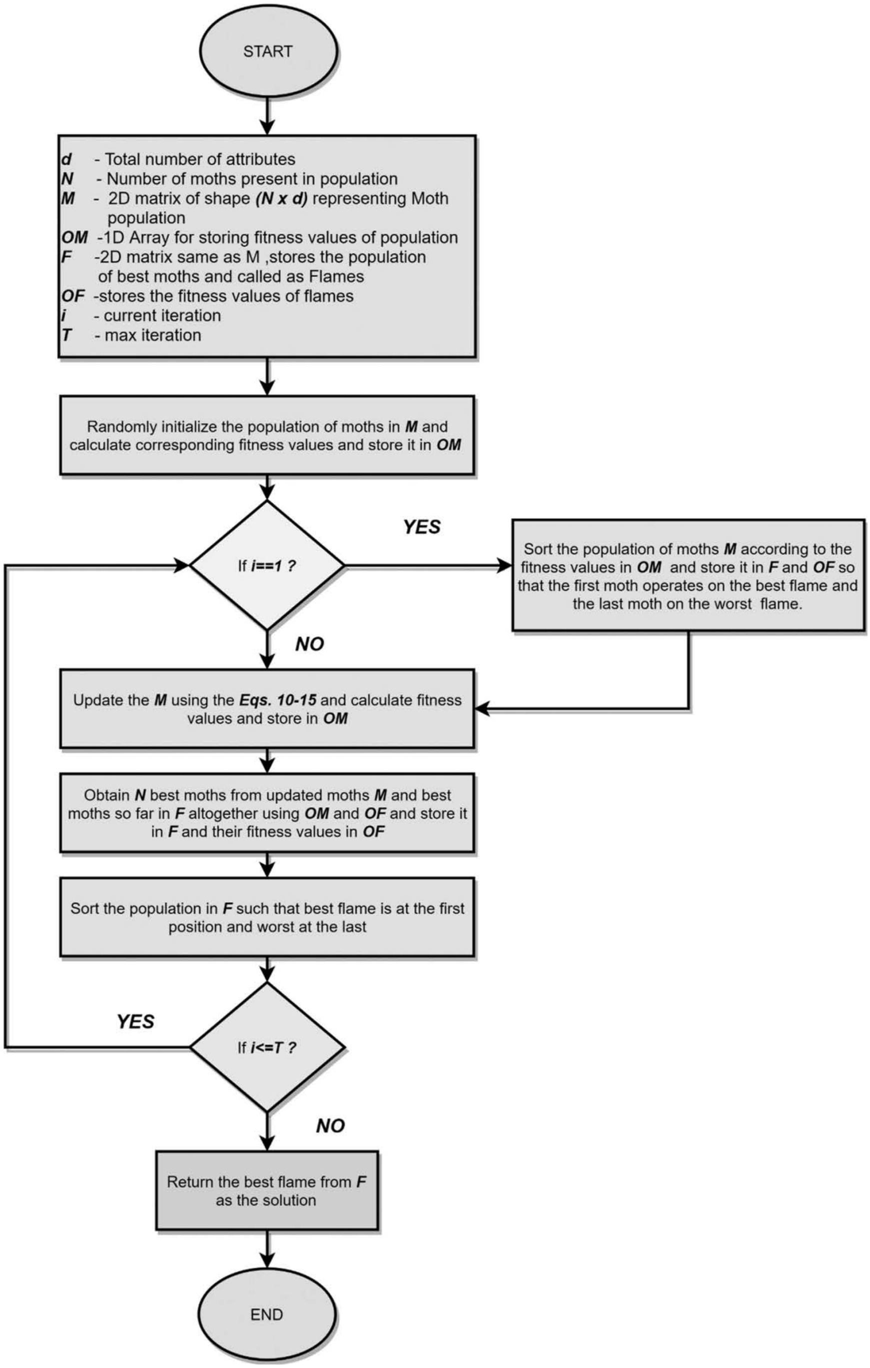

FIGURE 8.5 Flowchart representing the feature optimization using MFO.

The number of flames adaptively decreases throughout iterations as shown:

$$\text{Flame number} = \text{Round}\left(N - l * \frac{N-1}{T} \right) \tag{8.10}$$

where l, N, and T denote current iteration, maximum number of flames, and maximum number of iterations, respectively.

This continues till T returns true, and in the end, the best moth is returned as the best-obtained approximation of the optimum.

Figure 8.5 shows the flowchart depicting the MFO algorithm.

8.6 EXPERIMENTAL RESULTS AND DISCUSSION

In this study, a bi-stage model for the classification of ECG beats has been designed. For its evaluation, we have considered the MIT-BIH database. In the current experiment, 70% of the total database is used as the training set. Remaining 30% is equally divided to represent the validation and testing set, respectively. The gray-scale sample images are firstly transformed into RGB images (since VGG accepts only RGB images) of a dimension of $128 \times 128 \times 3$ before feature extraction by VGG16. Training is done over 20 epochs with a batch size of 64. A 0.2 Dropout is employed to prevent overfitting. Also, TL helps reduce the training time substantially.

In the present experiment, the feature set produced from VGG16 is used in the next stage of the framework. From Figure 8.6, it can be observed that

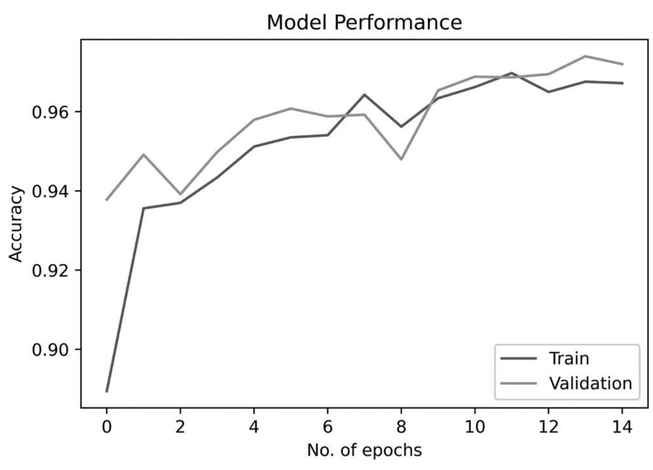

FIGURE 8.6 The train and validation accuracy curves with respect to the number of epochs for the VGG-16 classifier.

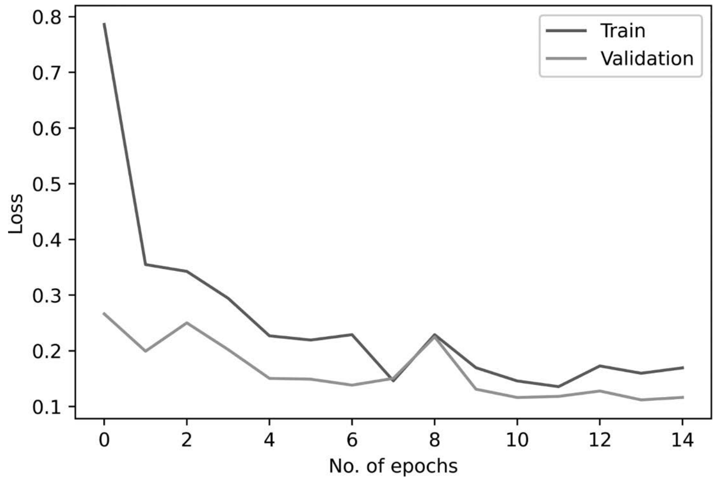

FIGURE 8.7 The train and validation loss curves with respect to the number of epochs for the VGG-16 classifier.

the VGG16 model has a good performance and achieves a 97.36% accuracy on the test set. Figures 8.7, and 8.8 represent the loss curves and confusion matrix, respectively. In the next stage of our framework, the MFO algorithm has been used to optimize the feature vector achieved from the VGG16 model. The mechanism eliminates the irrelevant features and thus helps in reducing the dimension of the feature set without compromising the classification performance. The results of the VGG16 model before applying the MFO algorithm are given in Table 8.2.

From the VGG16 model, we have extracted 128 features from the penultimate layer of the model. Convolutional and Pooling blocks extract the generic features, while the Dense layer in the model distinguishes those special features of the images which makes the classification tasks easier. So, we have used the features of the second last layer for optimization as they contain most of the information related to the dataset.

It has been observed that the MFO algorithm can reduce the feature set's dimension from 128 to 58, i.e., almost a 55% reduction without any significant loss in performance. Instead, the performance of the overall classification task improves marginally by 0.5%. The results are detailed in Table 8.3.

After applying the MFO-based feature selection approach, the confusion matrix is shown in Figure 8.9. It is clear from this figure that after optimizing the feature set, the performance has improved.

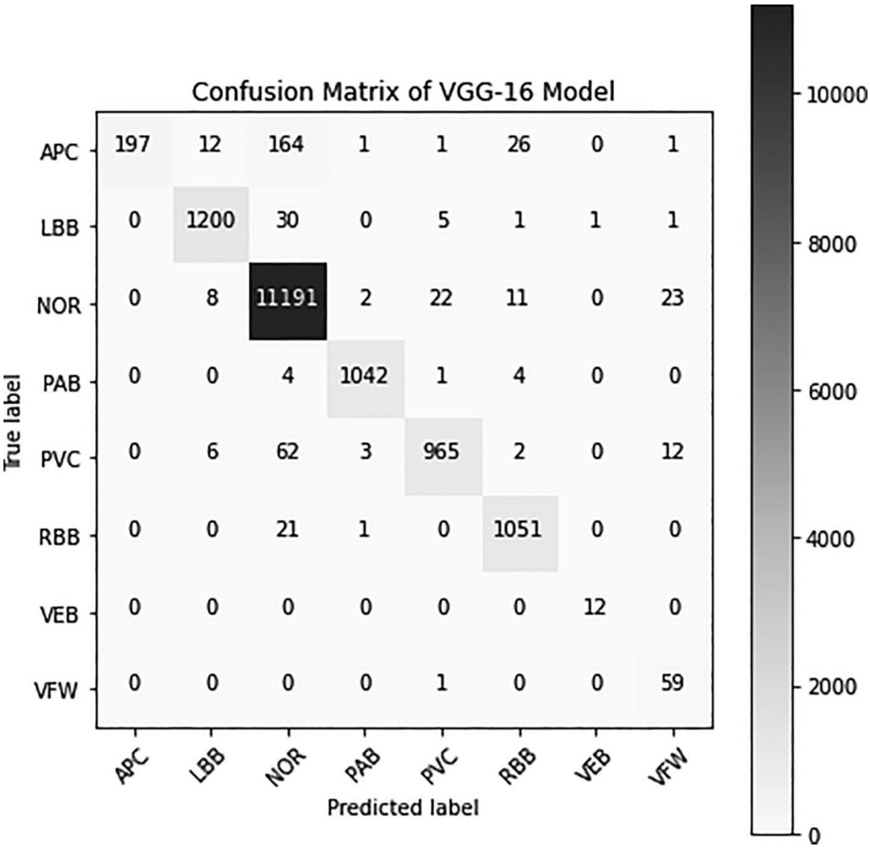

FIGURE 8.8 Confusion matrix of the VGG-16 classifier model.

TABLE 8.2 Performance of the Pre-Trained VGG16 Classifier for Extraction of the Feature Set

Model	**Accuracy (%)**	**Precision**	**Recall**	**F_1 Score**
VGG16	97.36	0.97	0.97	0.97

TABLE 8.3 Classification Performance of ECG Samples Before and After Application of the MFO Algorithm

Method	**Accuracy (%)**	**Dimension**	**Precision**	**Recall**	**F1 Score**
Before feature selection	97.36	128	0.97	0.97	0.97
After applying MFO-based feature selection approach	97.86	58	0.98	0.98	0.98

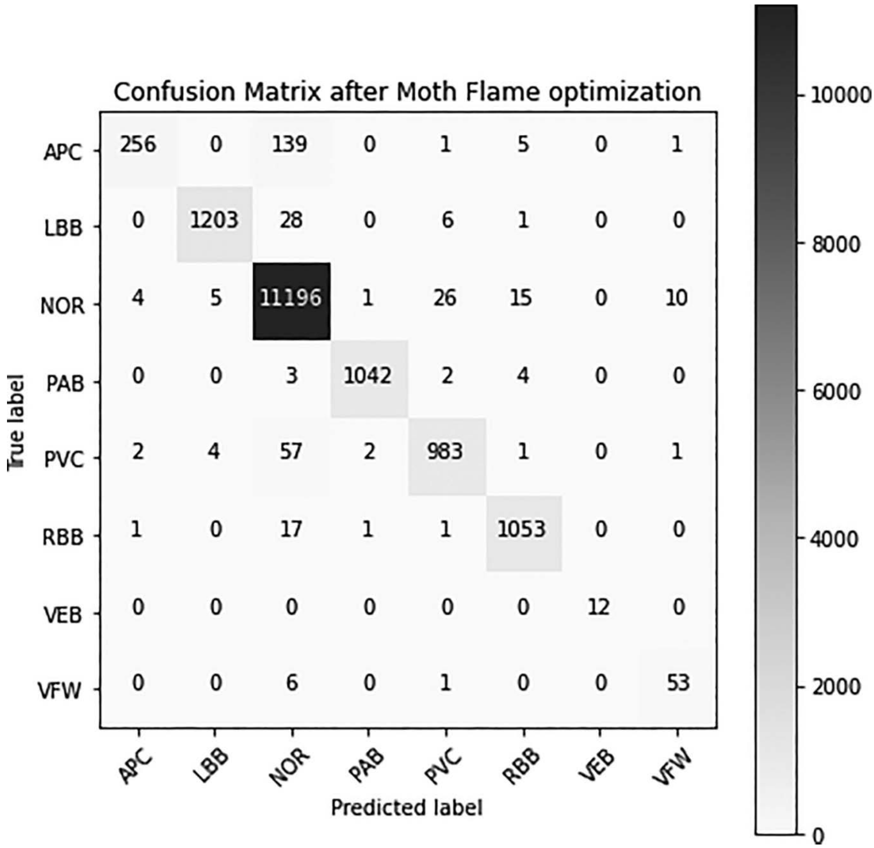

FIGURE 8.9 Confusion matrix after applying MFO algorithm and SVM classifier.

The ROC curves before and after using the MFO algorithm are given in Figures 8.10 and 8.11, respectively.

We have used 20 moths for optimization, the fitness and dimension of each moth are given in Table 8.4. This represents the dimensions of each moth in the population after the final iteration along with the classification accuracy.

Table 8.3 justifies that the proposed method is assuring with respect to performance and efficiency. We have tried to provide a comparative study with the other existing approaches evaluated on the same dataset in Table 8.5. In Ref. [50], the authors claim a classification accuracy of 98.71% on eight classes with a test set of size 4,900. First, they have used Independent Component Analysis (ICA) to separate independent sources from mixed ECG signals components. After that, they used neural

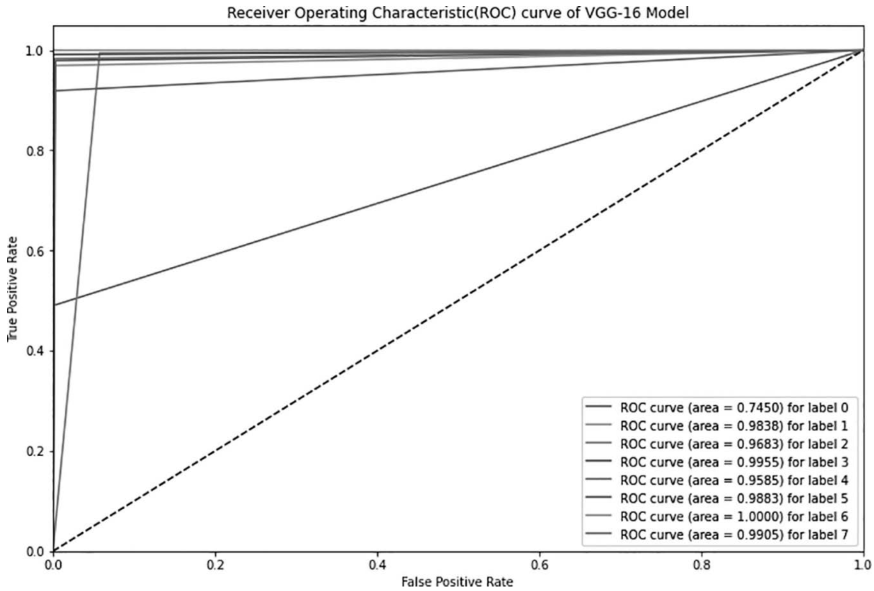

FIGURE 8.10 ROC curve of VGG-16 model.

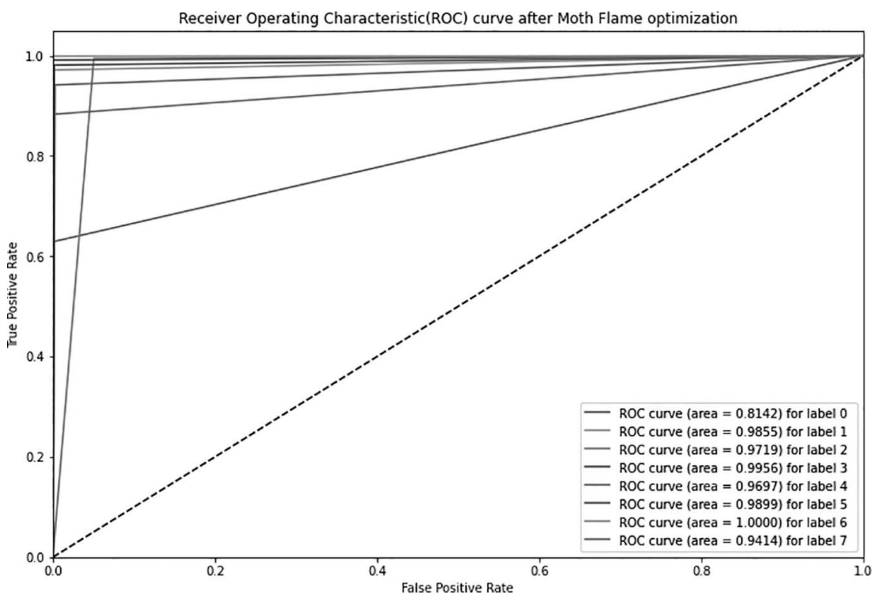

FIGURE 8.11 ROC curve after Moth-Flame Optimization.

TABLE 8.4 Fitness and Feature Dimensions of the Population After the Final Iteration

Moth No.	Accuracy (%)	Dimension	Moth No.	Accuracy (%)	Dimension
1	97.86	58	11	97.83	76
2	97.86	65	12	97.83	67
3	97.85	71	13	97.83	79
4	97.85	71	14	97.82	67
5	97.84	87	15	97.82	80
6	97.84	72	16	97.82	74
7	97.84	65	17	97.82	85
8	97.84	63	18	97.82	57
9	97.83	70	19	97.81	72
10	97.83	72	20	97.81	67

TABLE 8.5 A Comparison between the Proposed Approach and Other Existing Approaches

Work Ref.	Method Used	No. of Data Samples in the Test Set	Classification Accuracy (in %)
Yu et al. [51]	Integration of ICA and classification using neural networks	4,900	98.71
Yu et al.[50]	Novel independent components arrangement approach for feature extraction and classification using SVM	4,900	98.70
Wang et al. [52]	They used PCA and LDA for ECG feature set minimization. Then, a probabilistic neural network was used to classify the reduced feature set	4,900	99.71
Proposed Method	A two-staged network involves optimizing the features of ECG scans using MFO and classifying it using SVM	16,143	97.86

networks to classify the ECG beats. The dataset used in Ref. [51] is a small subset of the original dataset MIT-BIH. They had selected two samples at random from each record corresponding to the eight classes which produced a dataset of size 9,800. A similar kind of random sub-sampling is used by Yu et al. [50] and Wang et al. [52]. The size of the mentioned datasets in these said papers [50,52] is <10% of what we have considered in the present study. Therefore the classification accuracies provided in these papers [50,52] may not act as benchmarking results rather motivate us to achieve comparable results.

Table 8.5 shows that the proposed method in this work is comparable with past works in ECG classification and produces satisfactory results.

8.7 CONCLUSION

This work proposes a bi-stage framework for arrhythmia classification using ECG signals. In the first level, pixel-level features are extracted from the 2D images (constructed from the 1D ECG signals) using the VGG16 model. After this, to optimize this feature set, we have applied a meta-heuristic MFO algorithm on it. The obtained performance shows that this method successfully reduces the size of the feature set and also improves the classification results. This method can be hence used as an aid in arrhythmia diagnosis in humans. As future work, this framework can be evaluated on other domains to measure its robustness. Again, in the future, an ensemble of different DCNN models or the stacking of different feature sets can be tried out to have better recognition accuracy. Besides, the MFO algorithm can be modified and/or blended with other optimization algorithms for better optimization of the feature set produced by some classification models at the initial stage.

REFERENCES

1. https://www.who.int/health-topics/cardiovascular-diseases/.
2. Heron, M., Deaths: Leading causes for 2017. *Nat. Vital Stat. Rep.*, 68(6), 77 (2019).
3. American Heart Association, Electrocardiogram (ECG or EKG) (2015). https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/electrocardiogram-ecg-or-ekg.
4. Kligfield, P., The centennial of the Einthoven electrocardiogram. *J. Electrocardiol.*, 35(Suppl), 123–129 (2002).
5. Bakator, M., & Radosav, D., Deep learning, and medical diagnosis: A review of literature. *Multimodal Technol. Interact.*, 2, 47 (2018). doi: 10.3390/mti2030047.
6. Stead, W.W., Clinical implications and challenges of artificial intelligence and deep learning. *JAMA,* 320, 1107–1108 (2018).
7. Naylor, C., On the prospects for a (deep) learning health care system. *JAMA,* 320, 1099–1100 (2018).
8. Bejnordi, B. E. et al., Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA,* 318, 2199 (2017).
9. De Fauw, J. et al., Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.,* 24, 1342–1350 (2018).
10. Beck, E. J., Gill, W. & De Lay, P. R., Protecting the confidentiality and security of personal health information in low- and middle-income countries in the era of SDGs and Big Data. *Glob. Health Action,* 9, 32089 (2016).
11. Hannun, A. Y. et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. Med.,* 25, 65–69 (2019).

12. Clifford, G. D. et al., AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. *Comput. Cardiol.*, 44, 1–4 (2017).
13. Cook, D. A., Oh, S., & Pusic, M. V., Accuracy of physicians' electrocardiogram interpretations: A systematic review and meta-analysis. *JAMA Internet Med.,* 180(11), 1461–1471 (2020). doi: 10.1001/jamainternmed.2020.3989.
14. Robertson, S., Unhealthy lifestyle raises heart disease risk more than genetics. *News-Medical* (2019) viewed 12 July 2021, https://www.news-medical.net/news/20190903/Unhealthy-lifestyle-raises-heart-disease-risk-more-than-genetics.aspx.
15. Sassi, R. et al., PDF-ECG in clinical practice: a model for long-term preservation of digital 12-lead ECG data. *J. Electrocardiol.,* 50, 776–780 (2017).
16. Lyon, A., Mincholé, A., Martínez, J. P., Laguna, P. & Rodriguez, B., Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. *J. R. Soc. Interface,* 15, 20170821 (2018).
17. Mirjalili, S., Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Syst.*, 89, 228–249 (2015). doi: 10.1016/j.knosys.2015.07.006.
18. Zawbaa, H.M., Emary, E., Parv, B., & Sharawi, M., Feature selection approach based on the moth-flame optimization algorithm. *In 2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, Canada, pp. 4612–4617 (2016).
19. Ullah, A., Rehman, S. U., Tu, S., Mehmood, R.M., Fawad, & Ehatisham-ul-Haq, M., A hybrid deep CNN model for abnormal arrhythmia detection based on cardiac ECG signal. *Sensors*, 21, 951 (2021). doi: 10.3390/s21030951.
20. Moody, G. B. & Mark, R. G., The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol.*, 20(3), 45–50 (2001). doi: 10.13026/C2F305.
21. Houssein, E. H., Abdelminaam, D. S., Ibrahim, I. E., Hassaballah, M., & Wazery, Y. M., A hybrid heartbeats classification approach based on marine predators algorithm and convolution neural networks. *IEEE Access*, 9, 86194–86206 (2021). doi: 10.1109/ACCESS.2021.3088783.
22. Taddei, A., Distante, G., Emdin, M., Pisani, P., Moody, G. B., Zeelenberg, C., & Marchesi, C., The European ST-T database: Standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *Eur. Heart J.*, 13, 1164–1172 (1992). doi: 10.13026/C2D59Z.
23. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., & Stanley, H. E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220 (2000). doi: 10.13026/C2V88N.
24. Raghu, N., Arrhythmia detection based on hybrid features of T-wave in electrocardiogram. In: J. Joshua Thomas, et al. (eds.), *Deep Learning Techniques and Optimization Strategies in Big Data Analytics*. IGI Global, pp. 1–20 (2020). doi: 10.4018/978-1-7998-1192-3.ch001.
25. Dutta, S., Chatterjee, A., & Munshi, S., Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification. *Med. Eng. Phys.*, 32(10), 1161–1169 (2010). doi: 10.1016/j.medengphy.2010.08.007.

26. Parveen, A., Vani, R. M, Hunagund, P.V., & Masroor, F., Deep learning: 1-D convolution neural network for ECG signal. *Int. J. Ind. Electron. Electr. Eng. (IJIEEE)*, 8(6), 1–17 (2020).
27. Ince, T., Kiranyaz, S., & Gabbouj, M., A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Trans. Biomed. Eng.*, 56(5), 1415–1426 (2009). doi: 10.1109/TBME.2009.2013934.
28. Kachuee, M., Fazeli, S., & Sarrafzadeh, M., ECG heartbeat classification: A deep transferable representation, *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 443–444 (2018). doi: 10.1109/ICHI.2018.00092.
29. Melgani, F. & Bazi, Y., Classification of electrocardiogram signals with support vector machines and particle swarm optimization. *IEEE Trans. Inf. Technol. Biomed.*, 12(5), 667–677 (2008). doi: 10.1109/TITB.2008.923147.
30. Inan, O. T., Giovangrandi, L., & Kovacs, G. T. A., Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Trans. Biomed. Eng.*, 53(12), 2507–2515 (2006). DOI: 10.1109/TBME.2006.880879.
31. Jun, T., Nguyen, H.M., Kang, D., Kim, D., Kim, D., & Kim, Y., ECG arrhythmia classification using a 2-D convolutional neural network. ArXiv, abs/1804.06812 (2018).
32. Pereira, D. C., Ramos, R. P., & do Nascimento, M.Z., Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. *Comput. Methods Programs Biomed.*, 114(1), 88–101 (2014). doi: 10.1016/j.cmpb.2014.01.014.
33. Blum, C. & Li, X., Swarm intelligence in optimization. In: *Swarm Intelligence*. Springer: Berlin, Heidelberg, pp. 43–85. (2008).
34. Man, K.F., Tang, K.S., & Kwong, S., Genetic algorithms: Concepts and applications in engineering design. *IEEE Trans. Ind. Electr.*, 43(5), 519–534 (1996). doi: 10.1109/41.538609.
35. Dorigo, M., Birattari, M., & Stutzle, T., Ant colony optimization. *IEEE Comput. Intell. Mag.*, 1(4), 28–39 (2006). doi: 10.1109/MCI.2006.329691.
36. Kennedy, J. & Eberhart, R., Particle swarm optimization. *Proceedings of ICNN'95- International Conference on Neural Networks*, pp. 1942–1948, vol. 4 (1995). DOI: 10.1109/ICNN.1995.488968.
37. Karaboga, D., & Basturk, B., Artificial Bee Colony (ABC) optimization algorithm for solving constrained optimization problems. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., & Pedrycz, W. (eds) *Foundations of Fuzzy Logic and Soft Computing.* IFSA 2007. Lecture Notes in Computer Science, vol. 4529. Springer: Berlin, Heidelberg (2007). doi: 10.1007/978-3-540-72950-1_77.
38. Mohamad, A., Zain, A., Bazin, N.E.N., & Udin, A., Cuckoo search algorithm for optimization problems: A literature review. *Appl. Mech. Mater.* (2013). doi: 10.4028/www.scientific.net/AMM.421.502.
39. Ghaheri, A., Shoar, S., Naderan, M., & Hoseini, S.S. The applications of genetic algorithms in medicine. *Oman. Med. J.*, 30(6), 406–416 (2015). doi: 10.5001/omj.2015.82.

40. Valsecchi, A., Damas, S., & Santamaria, J. (eds.), An image registration approach using genetic algorithms. *2012 IEEE Congress on Evolutionary Computation (CEC).* IEEE, Australia (2012).
41. Ling, S.S., & Nguyen, H.T., Genetic-algorithm-based multiple regression with fuzzy inference system for detection of nocturnal hypoglycemic episodes. *IEEE Trans. Inf. Technol. Biomed.*, 15(2), 308–15 (2011).
42. Kumar, A., Ashok, A., & Ansari, M. A., Brain tumor classification using hybrid model of PSO and SVM classifier, *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 1022–1026 (2018). DOI: 10.1109/ICACCCN.2018.8748787.
43. Jangir, N., Pandya, M.H., Trivedi, I.N., Bhesdadiya, R., Jangir, P., & Kumar, A., Moth-flame optimization algorithm for solving real challenging constrained engineering optimization problems. *In 2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS).* IEEE, pp. 1–5 (2016).
44. Hazir, E., Erdinler, E.S., & Koc, K.H., Optimization of CNC cutting parameters using design of experiment (doe) and desirability function. *J. For. Res.*, 29(5), 1423–1434 (2018).
45. Wang, M., Chen, H., Yang, B., Zhao, X., Hu, L., Cai, Z., Huang, H., & Tong, C., Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses. *Neurocomputing*, 267, 69–84 (2017).
46. El Aziz, M.A., Ewees, A.A., & Hassanien, A.E., Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation. *Expert. Syst. Appl.*, 83, 242–256 (2017).
47. Muangkote, N., Sunat, K., & Chiewchanwattana, S., Multilevel thresholding for satellite image segmentation with moth-flame based optimization. *In 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE).* IEEE, pp. 1–6 (2016).
48. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., & Berg, A. C., Imagenet large scale visual recognition challenge. *Int. J. Comput. Vision*, 115(3), 211–252 (2015).
49. Liu, S., & Deng, W., Very deep convolutional neural network-based image classification using small training sample size. *In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, IEEE, Malaysia, pp. 730–734 (2015).
50. Yu, S.-N., & Chou, K.-T., Selection of significant independent components for ECG beat classification. *Expert Syst. Appl.*, 36, 2088–2096 (2009). doi:10.1016/j.eswa.2007.12.016.
51. Yu, S.-N., & Chou, K.-T., Integration of independent component analysis and neural networks for ECG beat classification. *Expert Syst. Appl.*, 34, 2841–2846 (2008). doi: 10.1016/j.eswa.2007.05.006.
52. Wang, J.-S., Chiang, W.-C., Hsu, Y.-L., & Yang, Y.-T.C., ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing*, 116, 38–45 (2013). doi: 10.1016/j.neucom.2011.10.045.