IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 22, 2020, accepted October 28, 2020, date of publication November 3, 2020, date of current version November 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035531

A Hybrid Meta-Heuristic Feature Selection Method Using Golden Ratio and Equilibrium Optimization Algorithms for Speech Emotion Recognition

ARIJIT DEY[1], SOHAM CHATTOPADHYAY[2], PAWAN KUMAR SINGH[3], (Member, IEEE), ALI AHMADIAN[4,5], (Member, IEEE), MASSIMILIANO FERRARA[5], AND RAM SARKAR[6], (Senior Member, IEEE)

[1]Department of Computer Science and Engineering, Maulana Abul Kalam Azad University of Technology, Kolkata 700064, India
[2]Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India
[3]Department of Information Technology, Jadavpur University, Kolkata 700106, India
[4]Institute of IR 4.0, The National University of Malaysia (UKM), Bangi 43600, Malaysia
[5]DiGiES & Decisions Laboratory, Mediterranea University of Reggio Calabria, 89124 Reggio Calabria, Italy
[6]Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, India

Corresponding author: Ali Ahmadian (ahmadian.hosseini@gmail.com)

This work was financially supported by Department of Law Economics and Human Science, Mediterranea University of Reggio Calabria Italy, under Grant: Piano di Eccellenza L.232/2016.

ABSTRACT Speech is the most important media of expressing emotions for human beings. Thus, it has often been an area of interest to understand the emotion of a person out of his/her speech by using the intelligence of the computing devices. Traditional machine learning techniques are very much popular in accomplishing such tasks. To provide a less expensive computational model for emotion classification through speech analysis, we propose a meta-heuristic feature selection (FS) method using a hybrid of Golden Ratio Optimization (GRO) and Equilibrium Optimization (EO) algorithms, which we have named as Golden Ratio based Equilibrium Optimization (GREO) algorithm. The optimally selected features by the model are fed to the XGBoost classifier. Linear Predictive Coding (LPC) and Linear Prediction Cepstral Coefficients (LPCC) based features are considered as the input here, and these are optimized by using the proposed GREO algorithm. We have achieved impressive recognition accuracies of 97.31% and 98.46% on two standard datasets namely, SAVEE and EmoDB respectively. The proposed FS model is also found to perform better than their constituent algorithms as well as many well-known optimization algorithms used for FS in the past. Source code of the present work is made available at: `https://github.com/arijitdey1/Hybrid-GREO`.

INDEX TERMS Speech emotion recognition, feature selection, golden ratio based equilibrium optimization, speech analysis, LPC and LPCC features, equilibrium optimization, golden ratio optimization, meta-heuristic.

I. INTRODUCTION

Speech signal happens to be the most common medium of communication among human beings. So, the automatic recognition of speech signals through computing devices is considered as an interesting problem among the research fraternity. Most of the time, along with information, speech express the emotion of a person. Speech emotion recognition (SER) plays an important role in modern Artificial Intelligence (AI) based systems, such as autonomous vehicle [1], voice assistance software, human physiology analysis, and medical services [1]–[3]. For an example, by using a SER system, one can predict the driver's emotion and can judge whether the driver is capable of driving or not. This prevents road accidents by telling the driver about his/her fatigue state. In medical science, a doctor can easily use a SER system

The associate editor coordinating the review of this manuscript and approving it for publication was Shiqing Zhang.

as a disease prediction tool for a physiologically depressed person or an autistic child. Recognition of human emotion from audio signals is one of the most challenging tasks in the domain of speech processing [4]. Humans communicate verbally through speech. The physiological studies explain that both facial expressions [5] and speech signals are informative for recognizing human emotions [6], which need to identify the adjustments of facial muscle and changing tone. In a nonverbal communication, the facial expression is responsible for 55% and voice intonation 38% and words 7% in the message perception [7], [8].

However, in real life, it is not an easy task to classify emotions from speech signals. The main difficulty lies in the field of acoustic signal processing is to extract meaningful as well as optimal features from the speech signals. In recent times, many machine learning as well as deep learning models are found to produce significant results in the field of SER. Most of the SER methods available in the literature are developed to extract new features from the speech signals. However, for every audio-clip, all the features are not of same importance. This is the reason that researchers find it difficult to achieve desirable accuracy using traditional feature extraction techniques. A speech signal, generally, has two different feature types such as temporal features (time domain features) and spectral features (frequency-based features). Some standard feature extraction methodologies introduced in the literature are: Mel-frequency cepstral coefficients (MFCC) [9], Linear Predictive Coding (LPC) [10], Linear Prediction Cepstral Coefficients (LPCC) [11], Perceptual Linear Prediction (PLP) [12] etc. In the present work, we mainly focus on the temporal features. This is because for the SER task, the combination of LPC and LPCC features give a promising result as the LPC parameters are more precise [13]. Furthermore, the reliability and robustness of LPCC features are far better than rest of the techniques [13]. These features capture the appropriate nature of speech signals required for human emotion recognition task than rest of the spectral features. It is also found that the classification results using this feature set give state-of-the-art accuracy as compared to other feature vectors.

The basic target of a feature selection (FS) model is to choose optimal set of features which can reduce the computational cost and storage requirement, as well as enhance the classification accuracy of the problem in hand. For this, the use of optimization algorithms play an effective role to discard redundant features from the original feature vector and to increase the classification accuracy. FS models have been successfully applied by the researchers in various fields. For example, numerous optimization algorithms [14] can be found in the literature. However, in the present work, we propose a novel hybrid optimization [16], [17] algorithm which decreases the size of the feature vector and increases the accuracy of the SER task. The most interesting part is that our proposed algorithm gives a better result than the deep learning models, thereby ensuring low resource requirement. It is to be noted that the constituent algorithms of the proposed FS model namely, Equilibrium Optimization (EO) algorithm [18] and Golden Ratio based Optimization (GRO) algorithm [19] are the meta-heuristic optimization algorithms and have not been used to form a hybrid FS model till date. Our proposed FS model, named as Golden Ratio based Equilibrium Optimization (GREO) algorithm, helps to improve both the exploration and the exploitation phases efficiently. The overall architecture of our proposed GREO based FS model is illustrated in Fig. 1.

A. CONTRIBUTIONS

The contributions of the present work used to solve the SER problem are highlighted below:

- We have designed a new hybrid meta-heuristic FS method named as GREO algorithm by combining GRO and EO algorithms, which has been used for SER from the audio signals.
- In this proposed approach of hybrid GREO algorithm, we have used Average Weighted Combination Mean (AWCM) for hybridizing both algorithms and for nearest neighbour searching of final candidate solution vector, we have chosen Sequential One Point Flipping (SOPF) technique. This combination of various statistical optimizing techniques makes our proposed approach very unique.
- A comparative study of proposed GREO and other popularly used FS algorithms is performed. The reported results aid confirm the idea of choosing particularly EO and GRO algorithms for the hybridization.
- We have evaluated our model on two standard SER datasets, namely Surrey Audio-Visual Expressed Emotion (SAVEE) and Berlin Database of Emotional Speech (EmoDB) and reported a comparative study of our proposed approach with recently evolved state-of-the-art techniques in Section IV.

II. LITERATURE REVIEW

Research on SER task has been started since long back. For example in the 20th century, Nakatsu *et al.* [20] proposed a method for SER using machine learning algorithm. After that, successful implementations of traditional machine learning algorithms in the notion of making speech recognition as an effective interface between robot and human are reported by Adam *et al.* [21] and Kim *et al.* [22]. However, an implementation of hidden Markov model (HMM) by Schuller *et al.* [23] in this particular field brought 76.1% and 71.8% classification accuracies for SER problem on EmoDB and VAM datasets respectively. Next, an improved Markov model is proposed and implemented on the German and English speech datasets each having 5250 samples and produced an average accuracy of 86.8% using global prosodic pitch and energy based features with the help of HMM classifier [24]. Later, Rong *et al.* extracted Zero-Crossing Rate (ZCR), spectral and energy based features and used K-nearest neighbor (KNN) classifier on Mandarin dataset [25].

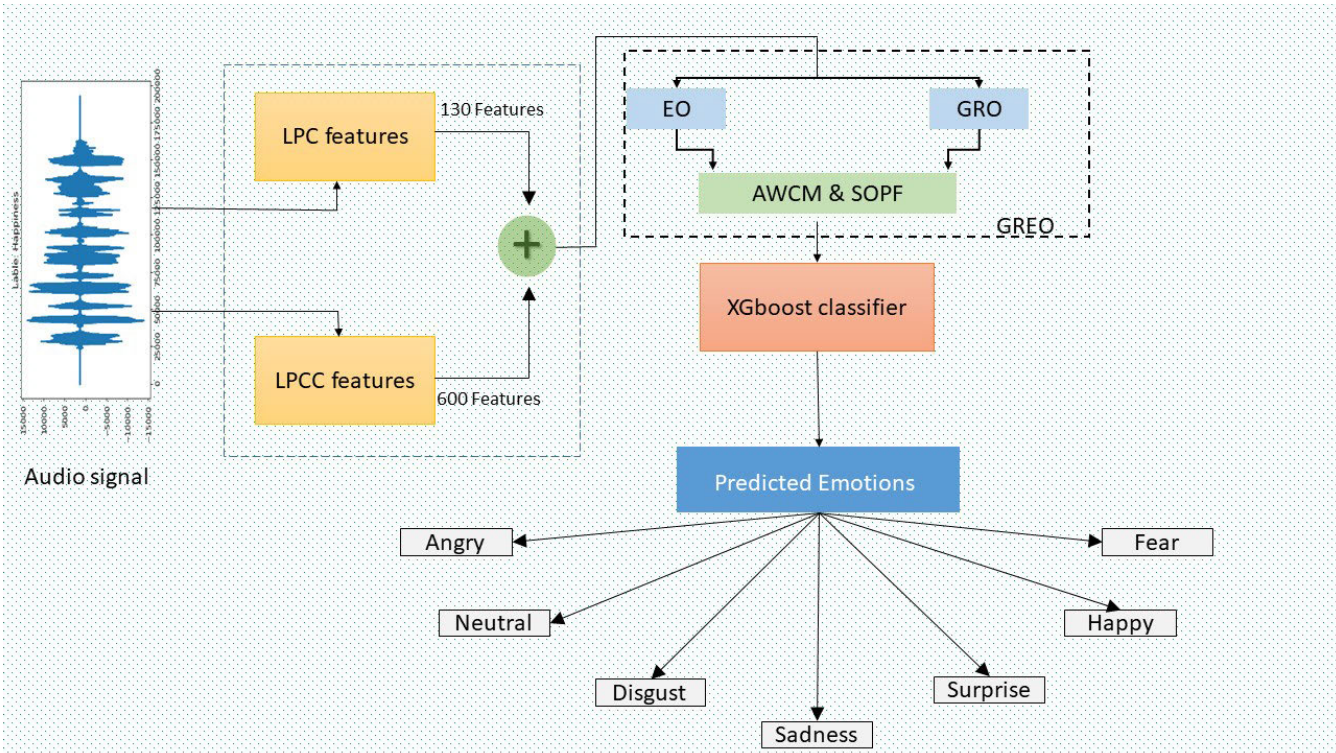

FIGURE 1. Schematic diagram illustrating our proposed FS model used for SER task.

Most of the aforementioned methods are based on the traditional machine learning techniques. Apart from that, a few deep learning techniques have also been proposed in recent times and shown their superiority over machine learning methods for SER tasks. An exploration of Recurrent neural network (RNN) and feed-forward neural network on IEMOCAP dataset is done by Haytham *et al.* [26] and reported 64.78% classification accuracy. Later, Yongming Huang *et al.* [27] implemented deep belief network on noisy ambience to recognize emotions from the speech signals. Later, J. Zhao *et al.* [28] classified human emotions with the help of 1D and 2D convolutional neural network (CNN) and tested on EmoDB dataset. In order to improve classification accuracy on SAVEE and EmoDB datasets, a deep neural network is also proposed in [6]. In 2017, Nicholas Cummins [29] proposed an approach to implement speech signals by spectrographic transformation into image and then classified it with the help of CNN [30]. Mostly, 2-D CNNs are implemented for visual recognition tasks but implementation in the audio signals is found to be unique. A two-layer fuzzy multiple random forest implementation [31] also contributed well in the SER field. Fig. 2 refers to the conventional workflow of the SER task found in the literature.

Moreover, meta-heuristic [65] approaches have become more reliable in the classification task among the researchers. It has numerous contributions in the field of signal processing. A. Das *et al.* [67] have applied Cuckoo optimization algorithm (COA) in the field of signal processing. Cat swarm optimization algorithm (CSO) [68] is also used to recognize emotions from the audio signals. Harris hawks optimization algorithm (HHO) [66] is another well-known optimization algorithm, which tunes the ConvoNet's parameters. Yogesh *et al.* [69] have come up with a simple technique to recognize both emotions and stress levels using a hybrid particle swarm optimization (PSO) algorithm. Researchers find that only a single optimization algorithm might not be sufficient to solve every single problem [73]. That is why, most of the researchers have developed different hybrid optimization algorithms in various fields. Some of the recently proposed optimization algorithm based FS methods are, Cosine Similarity based Harmony Search (HS) Algorithm [5], cooperative Genetic Algorithm (CGA) [60], Binary Bat Algorithm with Late Acceptance Hill-Climbing (BBA-LAHC) [61], hybridization of Mayfly algorithm (MA) and HS named as MA-HS algorithm [62], HS and Naked Mole-Rat Algorithm (HS-NMR) [63], hybridization of GA with PSO and Ant Colony Optimization (ACO) algorithm [64]. Besides, a few multi-objective optimization are also found for solving typical pattern recognition problems like spoken language identification [61], [63], facial emotion recognition [5], handwritten numeral recognition [70], handwritten script classification [71], [72] etc.

A. MOTIVATION

In initial stages of traditional machine learning era, the findings of various feature extraction techniques were the main point of research interest for quite a few years. As a result,

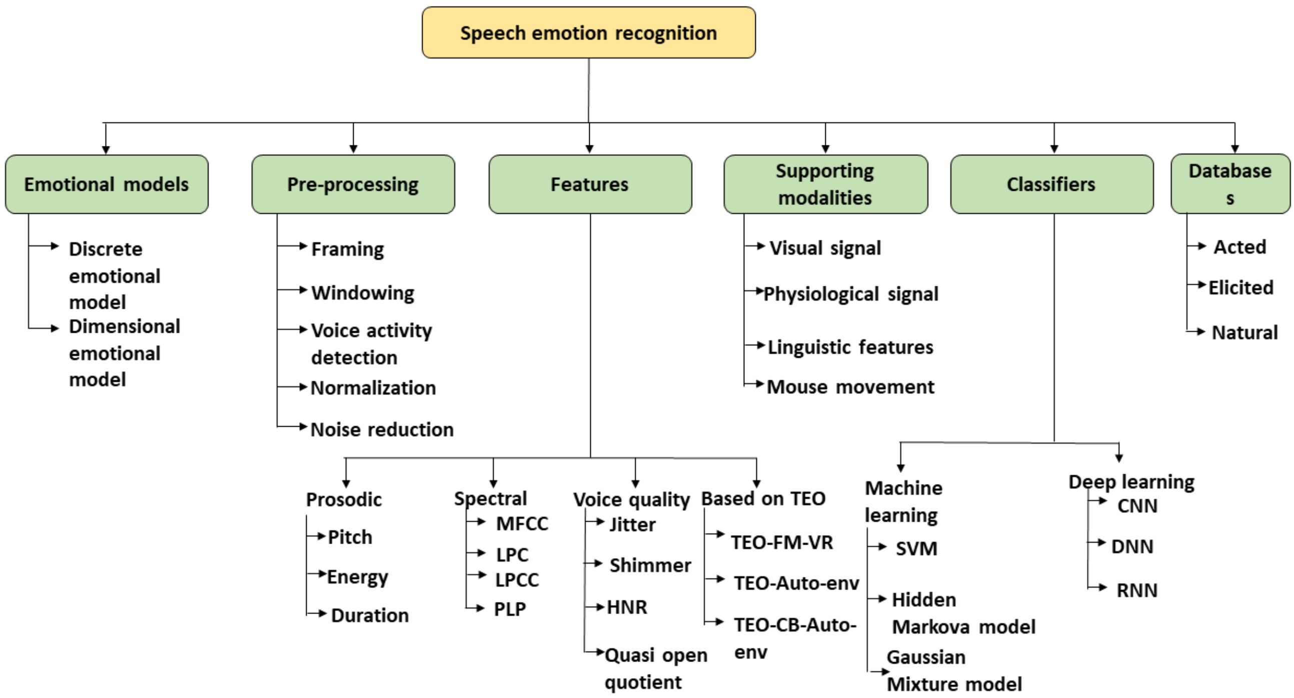

FIGURE 2. Conventional workflow of SER task commonly found in the literature.

various feature extraction methodologies had evolved from different domains of computational intelligence. Similarly, in the SER domain, some popularly used feature extraction methodologies that include MFCC, LPC, LPCC, RAASTA (Relative Spectral-Perceptual Linear Predictive) and so on had been proposed.

However, most of these feature vectors are quite large in terms of number of features and it is very hard to interpret which features contribute mostly in the case of training of an appropriate classifier. Hence, for most of the cases, if an entire feature set is chosen for the classification purpose then the model becomes computationally very costly due to the presence of some redundancy in the feature set. Therefore, to overcome this problem, FS models are applied. These models basically look for the optimal combination of features which results in the best performance of the classifier [32]. In recent times, various methods are introduced to make hybrid models considering different optimization algorithms [42]. These algorithms are used in different research fields such as image classification [33], emotion recognition from facial expressions [5] and so on. However, the ground of SER is not been explored with hybrid optimization algorithms so far. This motivates us to implement the hybrid GREO based FS algorithm in this specific field. To the best of our knowledge, this FS model is proposed for the first time for the SER problem. As mentioned above, for SER task, various feature extraction techniques are already available. So, in order to choose the best combination of feature sets, we have performed some experimentations and found that the combination of LPC and LPCC feature vectors outperforms other combinations with our proposed GREO based FS model and the results of are shown in Table 2.

The sectional review of the entire paper is mentioned as follows: The whole paper has in total five sections, named as I. Introduction, II. Literature survey, III. Motivation, IV. Materials and Methods, V. Results and Discussion and VI. Conclusion. Here, Section IV consists of four subsections, A. Dataset Description, B. Pre-processing, C. Feature extraction, D. Feature selection. In the feature selection subsection, our proposed hybrid GREO is discussed. In section V, there are seven subsections, named as A. XGboost classifier, B. Evaluation matrices used, C. Selection of final feature set, D. Tuning of Hyper-parameters, E. Comparison with other classifiers, F. Tuning of hyper-parameter of XGBoost classifier, G. Comparison with other FS algorithms.

III. MATERIALS AND METHODS

In this section, the workflow of the proposed work has been discussed sequentially. The entire work is divided into different subsections that include dataset description, pre-processing, feature extraction, feature selection using the proposed GREO algorithm and finally, classification using XGBoost classifier.

A. DATASET DESCRIPTION

An initial and basic stage for solving any research problem is to collect proper dataset and for our case, we have used two publicly accessible benchmark datasets namely, SAVEE and EmoDB.

1) SAVEE DATASET

The SAVEE dataset [57] contains audio samples of four British male research scholars of University of Surrey aged in between 27 to 31 (DC, JK, JE, KL). In total, 480 samples are taken (4 actors x 120 trials per actor) and emotions are physiologically classified in 6 categories (Anger, Disgust, Fear, Happiness, Sadness and Surprise). Text materials are chosen from 15 lexically transcribed speech of different American dialects sentences (TIMIT) and carefully classified into different emotion classes.

2) EmoDB DATASET

The EmoDB dataset [58] consists of 535 audio data collected from 10 professional actors. There are 7 emotion labels found in this dataset which are: Normal, Anger, Sadness, Happiness, Disgust, Anxiety and Fear.

B. PRE-PROCESSING

For any signal processing task, the pre-processing of sample data plays a vital role in determining the performance of a model. A simple audio pre-processing technique has been used in the present work which is discussed below:

1) PRE-EMPHASIS

The speech signal has both high frequency and low frequency parts, the high frequency part is compensated from the source signal which is stressed during the production of the speech signal. The main idea of this stage is to flatten the high frequency signal by using the high pass finite response (FIR) filter. The equation corresponds to this stage is given below:

$$J(x) = 1 - kx^{-1} \tag{1}$$

where, $J(x)$ is the output after normalization and x is the input signal, k is the pre-emphasis filter coefficient.

2) FRAMING

In this stage of pre-processing, the pre-emphasized signal is divided into small frames so that it can be analyzed independently. There are many framing techniques available, but in this paper, we have used frame shift, which frames on the basis of time difference of two starting points of two consecutive frames and the length of the frame.

3) WINDOWING

After the framing of audio signal, the edges of the signal become quite discontinuous and it reduces the performance. So, in order to get rid of this problem, we have implemented windowing at the edge of the frames. Hamming window is one of the possible ways to do this. The hamming window is used by using the following equation:

$$H_w = y - zcos(\frac{2\pi n}{N} - 1) \tag{2}$$

where, $y = 0.54$ and $z = 0.46$ are constants and N is the number of samples.

C. FEATURE EXTRACTION

For speech analysis from the audio signals, extracting important features is one of the most challenging but an unavoidable tasks. There are different kinds of feature extraction tools available, but for this work, we have used librosa library in Python. In this proposed work, we have extracted both LPC and LPCC features from the audio files.

1) LINEAR PREDICTIVE CEPSTRAL (LPC)

One of the most popular features of audio signal is LPC features [10]. Around 20 LPC features are extracted from time series audio signals, but here in our case we have extracted 130 LPC features by increasing the LPC autocorrelation order and find significant difference in performance when they are concatenated with LPCC features and optimized by proposed algorithm.

LPC analysis is carried out by characterizing each sample in the time-frequency domain and by some linear combination of M, where M is the order of the LPC analysis. In the present work, LPC autocorrelation function of order 130 is used. The frame $J(x)$ is initialized to 0 for $n < 0$ and $n >= N$. It is multiplied with the fast Fourier transform (FFT) parameter $N = 256$. The M^{th} order linear prediction, minimizing the error, is represented by the following equations.

$$\sum_{n=1}^{N} p(n)p(n-i) = L(i) \tag{3}$$

$$\sum_{j=1}^{N} \alpha_j \sum_{n=1}^{N} p(n-j)p(n-i) = L(i) = \sum_{n=1}^{N} p(n)p(n-i) \tag{4}$$

$$\sum_{j=1}^{N} \alpha_j L(j-i) = L(i) \tag{5}$$

where, $i = 1, 2, 3, 4, ..N$ and the coefficients of $L(j-i)$ form an autocorrelation matrix and it is similar to symmetric Toeplitz matrix. The values along the diagonal are same.

where, L stands for autocorrelation matrix. We can find the predictor vector by matrix inversion. Fig. 3 shows the LPC workflow used in the present work.

2) LINEAR PREDICTION CEPSTRAL COEFFICIENTS (LPCC)

LPCC [11] is same as the LPC but it is presented in the cepstrum domain. This method helps to extract features like pitch, vocal tract area function and formants at a low bit rate. LPCC features are extracted from audio signals by calculating the cepstral coefficients of the LPC features of the audio. Then, it is represented by the logarithmic magnitude spectrum which is derived from Fourier Transformation. For LPCC calculation, the LPC vector is necessarily needed and the CC (Cepstral Coefficient) vector is represented by $(b_1.b_2.b_3 \ldots b_N)$ and it is described by $(d_1.d_2.d_3 \ldots d_N)$. This LPC vectors are modified to form CC vector by some series

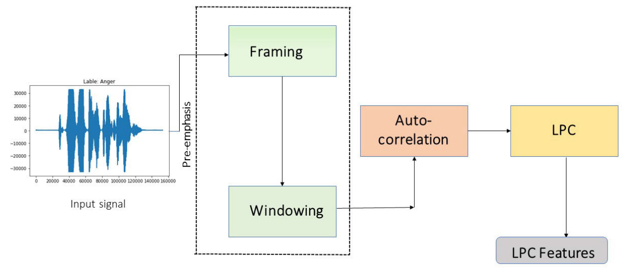

FIGURE 3. Schematic diagram representing the LPC [10] feature extraction methodology.

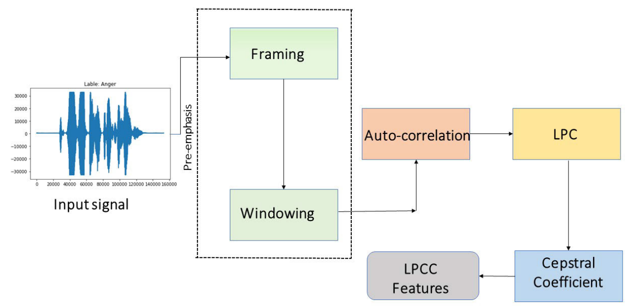

FIGURE 4. Illustration of LPCC [11] feature extraction procedure.

of recursive calls as defined below:

$$b_0 = Fn\sigma^2 \tag{6}$$

$$b_m = d_m + \sum_{1}^{m-1} b_k \cdot d_{k\_1} \, for \; 1 < m < N \tag{7}$$

$$b_m = \sum_{1}^{m-1} b_k \cdot d_{k\_1} \, for \; m > N \tag{8}$$

where, σ^2 stands for LPC gain and b_m stands for the cepstral coefficient and d_m stands for the predictor vector and $j = 1 < j < N - 1$. Fig. 4 shows the illustration of the LPCC feature extraction process. In this work, a feature set consisting of 600 features has been extracted using LPCC.

D. FEATURE SELECTION MODEL

After feature extraction, the most important work is to select optimized features and remove redundant features. For this purpose, we have implemented a hybrid meta-heuristic FS model named as GREO algorithm to improve both exploration and exploitation as well as to choose the optimal feature subset.

1) EQUILIBRIUM OPTIMIZATION ALGORITHM

EO algorithm [18] is a recently introduced meta-heuristic optimization algorithm which tries to maintain a good balance between the exploration and exploitation phases. Exploration seems searching in a globally space but avoiding the local optima, and exploitation seems searching in local space to get a promising solution and increasing the quality of search. EO algorithm gets an inspiration from the dynamic mass balance of a control volume system. A first-order ordinary differential equation expressing the generic mass-balance, in which the change in mass in time is equal to the amount of mass that enters into the system plus the amount being generated inside minus the amount that leaves the system, is described as:

$$V\frac{\mathrm{d}P}{\mathrm{d}t} = QP_{eq} - QP + G \tag{9}$$

Like every meta-heuristic algorithm, EO also starts with an initial population which is created based on the number of particles and the size of the feature dimension. The equation represents the initial randomized population is given below.

$$p_i^{initial} = p_{min} + rand_i(p_{max} - p_{min}) \tag{10}$$

where $p_i^{initial}$ represents the initial concentration vector of the i^{th} particle and p_{min} and p_{max} are the minimum and the maximum concentration of particles respectively, and $rand_i$ belongs to [0, 1] and n is the number of the particles in the population.

The equilibrium state concludes the optimization process as it globally optimizes, and at the starting point of optimization, there is no knowledge of optimization. Let's assume four particles remain the best among all through out the whole optimization procedure. In addition, another candidate is taken into the pool, which happens to be the average of these four candidates. The number of selection of the particles is arbitrary and different for other optimization algorithms. The five selected objects are mentioned below which help to construct a vector named as equilibrium pool.

$$\vec{p_{eq.pool}} = -\vec{p_{eq.(1)}}, \vec{p_{eq.(2)l}}, \vec{p_{eq.(3)}}, \vec{p_{eq.(4)}}, \vec{p_{eq.(avg)}} \tag{11}$$

After this, an exponential term (E) helps to update the concentration and an accurate expression wants to make a balance between the exploration and exploitation, and eventually tries to achieve a good optimization. As the turnout rate always varies in the time volume space so, ϵ, a random vector, ranging between [0, 1] is introduced as shown below:

$$\vec{E} = e^{\vec{\epsilon}(t-t_0)} \tag{12}$$

where t varies with the variation of the iteration (i), which is represented by the equation given as:

$$t = (1 - \frac{i}{maxi})(k_2 \cdot \frac{i}{maxi}) \tag{13}$$

In the above equation, i represents current iteration and maxi represents the maximum number of iterations. k_2 is a variable which manages to develop the exploitation ability. The following equation shows that if the search speed is slowed down by enhancing the exploration and the exploitation abilities then convergence can be achieved easily.

$$t_0 = \frac{1}{\in} ln(-k_1 sign(m - 0.5)[1 - e^{-\epsilon.t}]) + t \tag{14}$$

where, k_1 represents the exploration ability. The more the value of k_2 seems the high exploitation ability and lower

Algorithm 1 **Pseudocode for EO algorithm**

Input: Complete feature space, population size, max iteration
Output: Best combination of features (Final solution)

Initialize the particle's population, $j = 1, 2, 3, \ldots, n$
Assign equilibrium candidates' fitness a large number
Assign free parameters $k_1 = 2, k_2 = 1, GP = 0.5$;
while $i < maxi$ **do**

  for $I = 1, \ldots$, number of particles (n) **do**
    Calculate fitness of i^{th} particle
    if fit(p_i) $<$ fit($p_{eq(1)}$) **then**
      Replace $p_{eq(1)}$ with p_i and fit($p_{eq(1)}$) with fit(p_i)
    else if fit(p_i) $>$ fit($p_{eq(1)}$) and fit(p_i) $<$ fit($p_{eq(2)}$) **then**
      Replace $p_{eq(2)}$ with p_i and fit($p_{eq(2)}$) with fit(p_i)
    else if fit(p_i) $>$ fit($p_{eq(1)}$) and fit(p_i) $>$ fit($p_{eq(2)}$) and fit(p_i) $<$ fit($p_{eq(3)}$) **then**
      Replace $p_{eq(3)}$ with p_i and fit($p_{eq(3)}$) with fit(p_i)
    else if fit(p_i) $>$ fit($p_{eq(1)}$) and fit(p_i) $>$ fit($p_{eq(2)}$) and fit(p_i) $>$ fit($p_{eq(3)}$) and fit(p_i) $<$ fit($p_{eq(4)}$) **then**
      Replace $p_{eq(4)}$ with p_i and fit($p_{eq(4)}$) with fit(p_i)
    end if
  end for
  $P_{avg} = \frac{(p_{eq(1)} + p_{eq(2)} + p_{eq(3)} + p_{eq(4)})}{4}$
  Equilibrium pool $P_{eq.pool} = (p_{eq(1)}, p_{eq(2)}, p_{eq(3)}, p_{eq(4)}, p_{eq(avg)})$
  Accomplish memory saving (if $i > 1$)
  assign $t = (1 - \frac{i}{maxi})^{(\frac{k_2 \cdot i}{maxi})}$

  for $I = 1, \ldots$, number of particles (n) **do**
    Choose a candidate randomly from the equilibrium pool
    Generate random number ϵ and m
    $E = k_1 \times sign(m - 0.5) \times [\exp^{-\epsilon \cdot i} - 1]$
    Construct $GCP = \begin{cases} 0.5 \cdot m & \text{if } m > GP \\ 0 & \text{else} \end{cases}$
    Construct $F_0 = GCP(P_{eq} - \epsilon \cdot P)$
    Construct $F = F_0 \cdot E$
    Update concentration $P = p_{eq} + (P - p_{eq}) \cdot E + \frac{F}{\epsilon \cdot V} \times (1 - E)$
  end for
  $i = i + 1$
end while
Output: Final Solution

the exploration. $sign(m - 0.5)$ shows the direction of the exploration and exploitation. The value of m is in between 0 and 1. The revised form of Eqn. (12) is represented as follows.

$$\vec{E} = k.sign\,(m - 0.5)\left[e^{-\in t} - 1\right] \tag{15}$$

After that, the next important stage is generation rate which helps to give a exact solution of the optimization task by ensuring a good exploitation phase. There are many models to compute generation rate among those one of the well known models for 1-D space is as follows.

$$\vec{H_G} = \vec{H_0}.e^{\vec{\epsilon}(t-t_0)} \tag{16}$$

where, H_0 is the initial value and ϵ is the decay constant. To get a more symmetric search pattern and controlled result, Eqn. (16) can be modified as follows:

$$\vec{H_G} = \vec{H_0}.\vec{E} \tag{17}$$

$$E \vec{-} 0 = GCP(Peq - \in P) \tag{18}$$

$$GCP = \begin{cases} 0.5 \cdot m & \text{if } m > GP \\ 0 & \text{else} \end{cases} \tag{19}$$

Here, GCP stands for generation control parameter which is the actual probability of the generation term in the updation process. Finally, the equation represents the EO updation rule which is as follows:

$$P = p_{eq} + \left(P - p_{eq}\right) E + \frac{F}{\in V\,(1 - E)} \tag{20}$$

The pseudocode of EO algorithm is explained in Algorithm 1.

Algorithm 2 Pseudocode for GRO algorithm

Input: Complete feature space, population size, max iteration
Output:Best combination of features (Final solution)

Initialize the particle's population $j = 1, 2, 3, \ldots, n$
Calculate the fitness function
while Convergence criterion is not satisfied **do**
  Obtain Yav, the mean value of all possible solution
  Set the worst fitness as Y_{worst}
  if fit(Y_{avg}) < fit(Y_{worst}) **then**
    Replace Y_{avg} with Y_{worst}
  end if

  for $I = 1, \ldots$, number of particles **do**
    choosing a population randomly from the Y_j
    Compare Y_I, Y_j, Y_{avg} and rank them according to their fitness values and the best will be sorted in Y_{best} and worst will be updated in Y_{worst}

$$Fib\,(n) = GF \cdot \frac{\left(\text{Ø}^n - \left(1 - \text{Ø}^{-n}\right)\right)}{\sqrt{5}} \ \ where\ GF = 1.618 \tag{21}$$

    Check the constraints
    $Y_t = Y_{median} - Y_{worst}$
  end for

  for $i = 1, \ldots$, number of particles **do**

    for $j = 1, \ldots$, number of variables **do**
      Update the solution $Y_{new} = (1 - Fib_t)Y_{best} + rand.Y_t.Fib_t$
      Check the constraints
    end for
  end for
end while
Output: Final Solution

2) GOLDEN RATIO OPTIMIZATION ALGORITHM

Though there is a diversity in the nature and the natural components, everything have unique shapes and sizes and follow fixed patterns, which become more visible from the gift of advanced science. Every physical phenomenon is witnessed in the form of a fixed proportion, called golden ratio [19]. The idea of golden ratio was first initiated by Fibonacci, and he introduced a series of numbers which are made by calculating the sum of previous two numbers and the ratio of the consecutive two numbers is 1.618, known as golden ratio. The idea behind this algorithm came from this property. Fibonacci numbers can be obtained from the following equation.

$$Fib\,(n) = GF \cdot \frac{\left(\text{Ø}^n - \left(1 - \text{Ø}^{-n}\right)\right)}{\sqrt{5}} \quad where\ GF = 1.618 \tag{22}$$

In the optimization process, everything is relating to the vector and the direction of the vector to fetch the best target. Initially the mean value of the population is calculated and then fitness is calculated. After calculating the fitness, the solution is compared with the mean solution and if it has a better fitness then the worst solution is replaced by it. Further the worst solution is calculated again and algorithm will proceed one step towards convergence. Then one solution vector is chosen at random from the population and the impact of that particular vector upon movements of another two solution vector from the entire population set is calculated. In addition to it, the direction of the solution vector is determined by considering the resultant of the directions of the two vectors. Now to denote the direction of the new vector, it is necessary to compare it with rest of the two previously chosen vectors. The vector which has the lowest value of the objective function is considered as the main vector.

$$Fib_{best} > Fib_{medium} > Fib_{worst} \tag{23}$$
$$Y_t = Y_{medium} - Y_{worst} \tag{24}$$

The above equation gives the information about the modulus value of the movement and the corresponding direction, in search of global minimum. Thereby to perform the global and local search operation Fibonacci's formula is used. The most important thing is to update the solution to achieve best one. To perform the global search from the whole space, it is better to add a random movement to add a new solution. The equation which is used to update the solution is given below.

$$Y_{new} = (1 - Fib_t)\, Y_{best} + rand \cdot Y_t \cdot Fib_t \tag{25}$$

Now, the new solution is updated and if the boundary condition is satisfied then the new solution will be replaced with the previous one. Algorithm 2 presents the pseudocode of the GRO algorithm.

3) PROPOSED GREO ALGORITHM

Both EO and GRO algorithms are meta-heuristic optimization algorithms. Both EO and GRO have the ability to maintain proper exploration and exploitation, so, the hybrid model gives more optimized solution on combination. In the first stage, both the EO and GRO algorithms are implemented separately, which finally, produce their final state of population having best solutions. Then, the combination of their population is prepared by evaluating the importance of all features belonging to any of the two sets of population. This process is known as average weighted combination method (AWCM) [15]. Thereafter, a local search method is applied on the provincial population outputted from both the subsets. For better results, we have implemented sequential one-point flipping (SOPF) which enhances both the subsets' discriminative nature. In AWCM, the sum of all the accuracies of all the solutions are calculated initially. For an example, if a solution from EO algorithm having an accuracy of 89% and if a solution from GRO algorithm having an accuracy of 90% are considered, then the importance of the feature is calculated as sum of both (that is, $0.89 + 0.90 = 1.79$). The AWCM cutoff (as shown in Table 1) is calculated as the mean of these importance values. The features which have higher importance than the AWCM will be finally included. If the size of each feature set is found to be N then, after calculating AWCM, it will become $2N$. The features are taken as the binarized form ('1' or '0'), and finally, we get a provincial population. The most significant issue is to cancel out the redundant features from the population outputted from the AWCM. This is done by applying a local search called SOPF. It is a non-greedy algorithm. SOPF sequentially checks every solution. SOPF considers each neighbour of the final solution which evolves from AWCM algorithm and calculates its fitness. If any neighbour results better fitness than that of the original solution, then the solution is replaced by its neighbour.

E. XGBoost CLASSIFIER

XGBoost or eXtreme Gradient Boosting, proposed by Chen *et al.* [36] is a recently developed and very widely used classifier.

XGBoost is advantageous not only for accurate performance but also in terms of classification speed. The main features the classifier provides are various types of boosting approaches such as **1. Gradient Boosting**, which includes learning rates only, **2. Stochastic Gradient Boosting**, which consists of row, column and column per split levels subsampling and **3. Regularized Gradient Boosting**, having the advantage of L1 and L2 regularization.

Algorithm of XGBoost classifier automatically handles the missing data value using **Sparse Aware** facility. This **Block Structured** algorithm supports tree constructed parallelization. An already fitted model of XGBoost classifier can further be boosted by **continuous training**.

The XGBoost classifier uses **Gradient Boosting decision tree algorithm** to boosting of gradients. The gradient boosting, popularly known as **multiple additive regression tree** is a type of ensemble learning technique which rectifies the error made by existing model with the newly introduced dataset. This kind of ensemble learning enables the idea of sequential embedding of model until the performance reaches to its saturation. Gradient Boosting a recently evolved approach where newly generated models take residuals and errors of prior models into account and add the experiences as a whole for the final prediction. This ensemble idea uses gradient descent to reduce the loss, and this is why it is called the Gradient Booster. The main characteristics which make it significantly popular over other classifiers are the fast execution time, parallelizable core and wide variety of changeable hyper-parameters making it more robust. In addition, it consistently outperforms other traditional classifiers for both classification and regression tasks, which is quite evident from our experiments discussed below.

IV. RESULTS AND DISCUSSION

In our current work, we have evaluated the proposed model on two benchmark SER datasets, namely SAVEE and EmoDB. For this purpose, we have performed several experiments to optimize our final results and and for evaluation, we have chosen some commonly used Evaluation Metrics.

A. EVALUATION METRICS

As mentioned above, in this present work, we have prepared a hybrid FS model of EO and GRO algorithms to achieve best combination of feature subset out of the entire feature set. To estimate the performance of our proposed model, we have relied on four popularly considered evaluating criterion, such as Accuracy, Precision, Recall and F1 Score.

These parameters are calculated depending upon some basic elementary measures, which can be found from the confusion matrix. These are the True positive, False positive, True Negative and False Negative values. These parameters are defined specially for binary class prediction but can be calculated from multi-class classification tasks also.

On the basis of above elementary parameters, we have calculated previously mentioned evaluation metrics with the following mathematical formulae: **Accuracy:**

$$Accuracy_i = \frac{\sum_i M_{i_i}}{\sum_i \sum_j M_{i_j}} \tag{26}$$

Precision:

$$Precision_i = \frac{\sum_i M_{i_i}}{\sum_i \sum_j M_{j_i}} \tag{27}$$

Recall:

$$Recall_i = \frac{\sum_i M_{i_i}}{\sum_j M_{i_j}} \tag{28}$$

F1 Score:

$$F1Score_i = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{29}$$

Now, using above formulae, we have estimated the performance of our model and compared it with some other traditional models to conclude the stand of our model in the field of SER.

Usually, accuracy is a good measure to evaluate the performance of a model in the domain of data science, but it can be considered as a sufficient measure only when, we have symmetric datasets. For a symmetric dataset, the false positive and the false negative values are almost same. Therefore, to generalize the evaluation task of the model, we have considered other parameters such as precision, recall and F1 score too. F1 score can be considered as more useful than accuracy, especially when the dataset has uneven distribution of classes.

B. SELECTION OF FINAL FEATURE COMBINATION

Prolific feature space selection is the most important task in the domain of machine leaning based classification. In the

TABLE 1. Example illustrating the AWCM cutoff calculation for obtaining the final optimized feature vector.

	Population	F1	F2	F3	F4	F5	Accuracy	WF1	WF2	WF3	WF4	WF5
EO	**EO1**	1	1	0	1	0	0.85	0.85	0.85	0	0.85	0
	EO2	0	1	1	0	1	0.93	0	0.93	0.93	0	0.93
	EO3	0	1	0	0	1	0.94	0	0.94	0	0	0.94
	EO4	1	0	0	0	1	0.73	0.73	0	0	0	0.73
	EO5	0	0	0	1	0	0.78	0	0	0	0.78	0
GRO	**GRO1**	1	1	1	0	0	0.87	0.87	0.87	0.87	0	0
	GRO2	1	0	0	1	0	0.92	0.92	0	0	0.92	0
	GRO3	0	0	1	0	1	0.65	0	0	0.65	0	0.65
	GRO4	0	1	1	1	0	0.90	0	0.90	0.90	0.90	0
	GRO5	1	0	1	0	1	0.82	0.82	0	0.82	0	0.82
Feature importance								**4.1**	**4.49**	**4.17**	**3.45**	**4.07**
AWCM cutoff								**4.056**				
Final feature vector								**1**	**1**	**1**	**0**	**1**

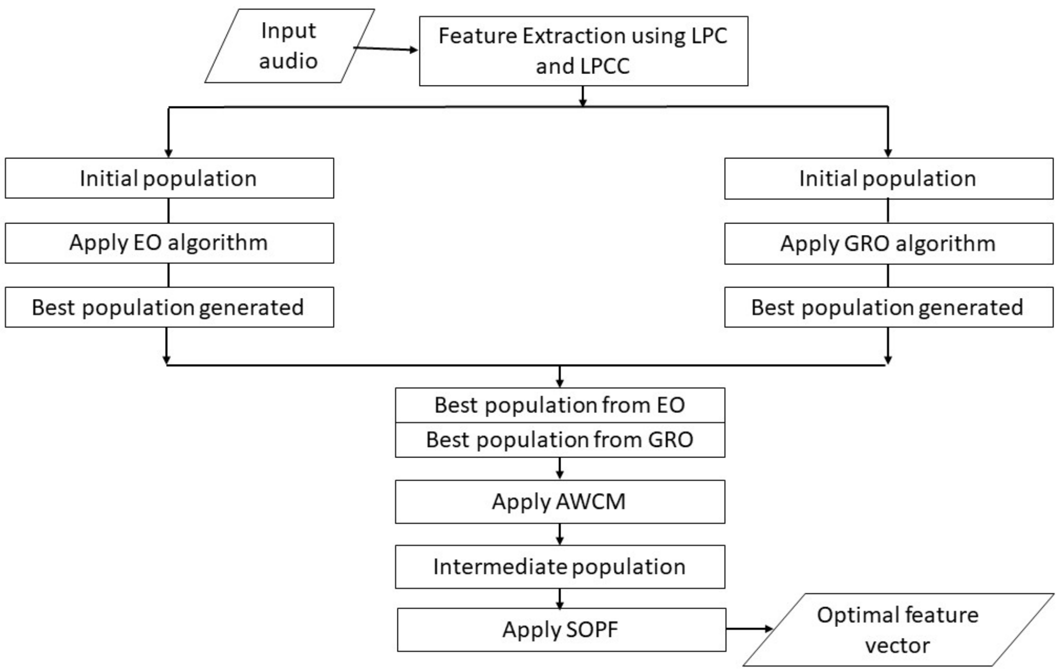

FIGURE 5. Schematic diagram representing the overall framework of the proposed GREO based FS model.

notion of this, we have extracted three different feature vectors, namely MFCC, RAASTA and LPC+LPCC from both datasets and tried different combinations by concatenating them. We have confirmed our feature set on the basis of final classification accuracy obtained on both datasets. The overall results are detailed in Table 2.

It is very much intuitive that the combination of LPC and LPCC features gives the best results among all other combinations because features from similar category with larger numbers often contain less number of redundant feature vectors.The feature space of LPC and LPCC achieves the best classification accuracies of 97.31% and 98.46% for SAVEE and EmoDB datasets respectively. Furthermore, the combination of all four types of features gives second best accuracy of 94.46% for SAVEE dataset, whereas for EmoDB dataset, the combination of LPC, LPCC and MFCC gives the second best result with 94.63% accuracy. Thereafter, it is also observed that the combination of only RAASTA and MFCC features gives the worst classification accuracies among all other combinations with 78.32% and 86.33% classification rates on SAVEE and EmoDB datasets respectively. Therefore, we have selected LPC and LPCC features as our final feature set, which is to be optimized. As a whole, we have extracted 600 features using LPCC feature descriptor and 130 features using LPC feature descriptor and concatenated them forming a total feature set of 730 elements containing final feature space.

C. TUNING OF HYPER-PARAMETERS

We have also performed above optimization algorithms along with our proposed hybrid model with various stages of hyper-parameters. Fig. 10 and Fig. 11 show the variations of

TABLE 2. Comparison of various combinations of feature sets with GERO model and XGBoost classifier applied on both the datasets.

Feature Combination	SAVEE				EmoDB			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
LPC+LPCC	**97.31%**	**95%**	**98%**	**99%**	**98.46%**	**97%**	**99%**	**96%**
MFCC	87.12%	82%	86%	88%	91.44%	94%	90%	91%
RAASTA	82.31%	79%	80%	85%	88.37%	90%	91%	93%
MFCC+LPC+LPCC	93.41%	91%	95%	94%	94.63%	95%	96%	92%
RAASTA+LPC+LPCC	85.44%	84%	88%	86%	90.22%	91%	91%	92%
MFCC+RAASTA	78.32%	80%	80%	79%	86.33%	90%	93%	87%
MFCC+RAASTA+LPC+LPCC	94.46%	94%	96%	93%	92.112%	91%	90%	92%

final classification accuracies with varying initial population sizes for SAVEE and EmoDB datasets respectively. From Fig. 10 and Fig. 11, it can be seen that the best classification accuracies are first achieved with population size equals to 30 for both datasets with our proposed algorithm. From Fig. 10, we can see that, with initial population size equals to 5, the accuracy is below 95% and when it gradually increases to the highest accuracy when population size is made 30. But with population size equals to 20 the accuracy decreased a little than that of 10. Thereafter, it first strikes 97.31% or the highest, at population size of 30. After that we can see the best accuracy achieved remains same till population size equals to 50. Then, again it decreases when population size is further increased to 60. Whereas for EmoDB dataset, the variation of 7 class classification accuracy of our proposed framework is a bit more stable and gradual. Here, it is also observed that the classification accuracy increases for population size varying from 5 to 10 and thereafter, the change is found to be minimal from 10 to 20. Similarly, for SAVEE dataset, the highest accuracy (measured as 97.31%) is first achieved with population size equals to 30. However, no change in accuracy is found for population size varying from 30 to 40 after which the accuracy gradually decreases.

Similar to above experiment, we have also performed several such experiments to find the optimum hyper-parameters which give the best classification accuracy on both of the datasets. We have finalized the hyper-parameters of both algorithms on the basis of final classification accuracy obtained by the hybrid model of EO and GEO algorithms. In Table 3, we have shown our final hyper-parameter values of EO algorithm on both datasets. From Table 3, it can be seen that the initial population size (30) and the maximum iteration value (20), the values of a1 (2) and a2 (1) for both datasets are same but the Omega differs from 0.85 to 0.9 for SAVEE and EmoDB datasets. In addition, the Pool size is also different with numerical values of 4 for SAVEE and 3 for EmoDB.

Similarly, the final hyper-parameters of GRO algorithm for both datasets are illustrated in Table 4. Here, also the population size and maximum number of iterations for both datasets are fixed to 30 and 15 respectively. Similar to that of EO algorithm, the Omega value is different but for SAVEE dataset, it is 0.95 whereas for EmoDB dataset, it is 0.85. Along with that, the Golden Value also differs a little. Referring Table 4, it is to be noted that that the golden value for SAVEE dataset is 1.75 and for EmoDB dataset, it is 1.95.

TABLE 3. Final hyper-parameters of EO algorithm on SAVEE and EmoDB datasets.

Parameter	SAVEE	EmoDB
Population Size	30	30
Maximum Iteration	20	20
Omega	0.85	0.9
a1	2	2
a2	1	1
Pool Size	4	3

TABLE 4. Final hyper-parameters of GRO algorithm on SAVEE and EmoDB datasets.

Parameter	SAVEE	EmoDB
Population Size	30	30
Maximum Iteration	15	15
Omega	0.95	0.85
Golden Value	1.75	1.95

This is to be mentioned that the optimum values of the hyper-parameters of each optimization algorithms are determined on the basis of the performance of the final hybrid model and not individual algorithm's performances.

In this work, we have plotted the ROC curves obtained for both SAVEE and EmoDB datasets (shown in Fig. 12 and Fig. 13 respectively) using our proposed GREO algorithm and XGBoost as the classifier.

For both datasets, our proposed model gives 100% classification accuracy for some specific emotion classes and near about 95% for rest of the emotion classes. The emotion class which results to higher accuracy, is considered as a prolific class. Such emotion classes up-hold the final classification accuracy of the model. Thus, due to our robust GREO based FS algorithm and efficient classifier, our proposed framework brings about state-of-the-art results for both SAVEE and EmoDB datasets.

D. COMPARISON WITH OTHER CLASSIFIERS

In the present work, to finalize our model we have performed experiments on both datasets by feeding the best solution obtained from the hybrid model to different classifiers like

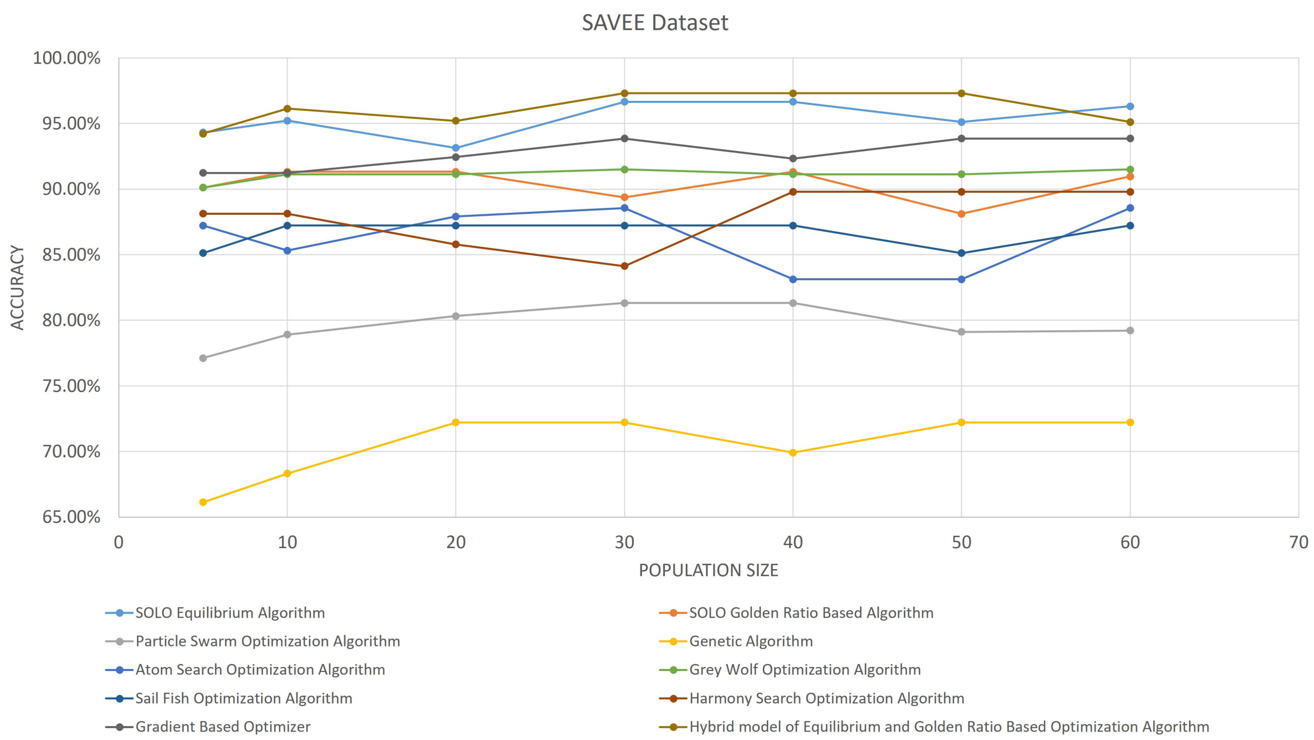

FIGURE 6. **Variation of classification accuracy with respect to population size on SAVEE dataset.**

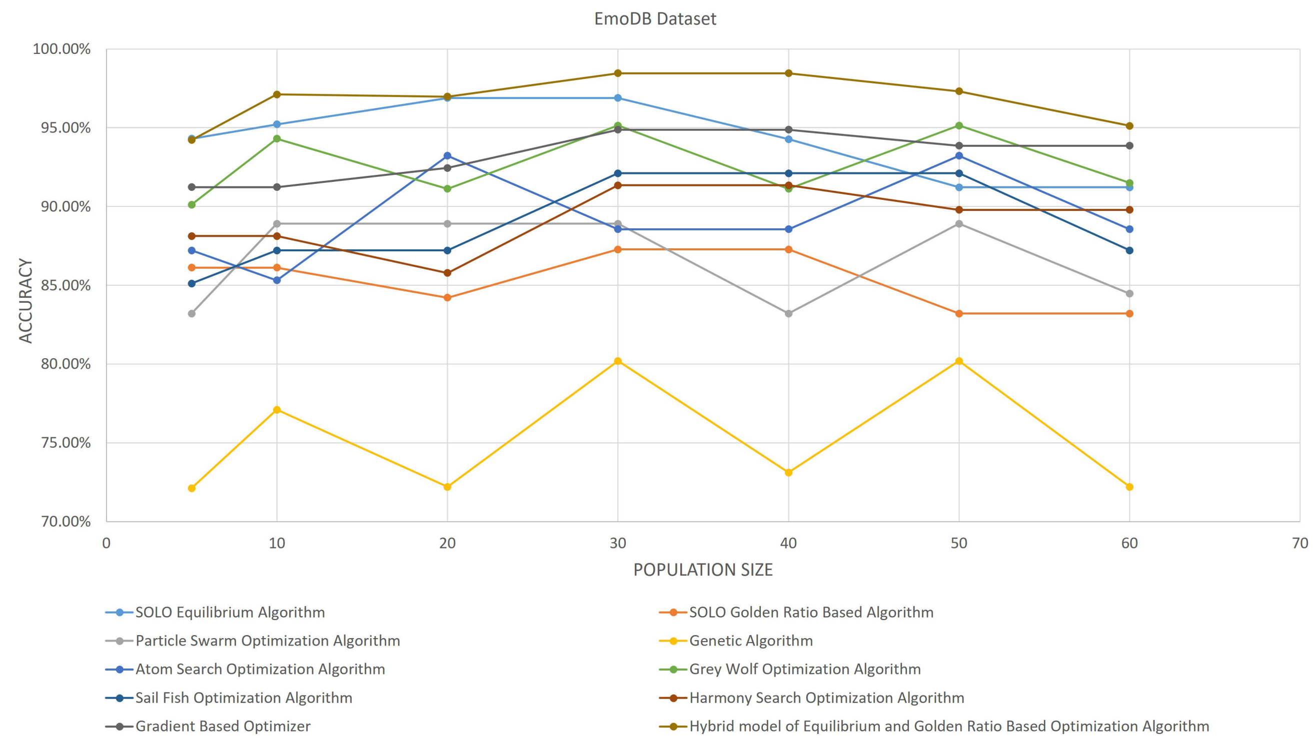

FIGURE 7. **Variation of classification accuracy with respect to population size on EmoDB dataset.**

KNN, MLP and XGBoost and have monitored the classification accuracies. The detailed results for SAVEE and EmoDB datasets are tabulated in Table 5. In Table 5, along with accuracies, we have also given the precision, recall and F1 score values for comparison purpose. It is observed from these experimentations that XGboost classifier has thoroughly outperformed other classifiers by reaching the state-of-the-art results over both datasets. It can be seen from Table 5 that

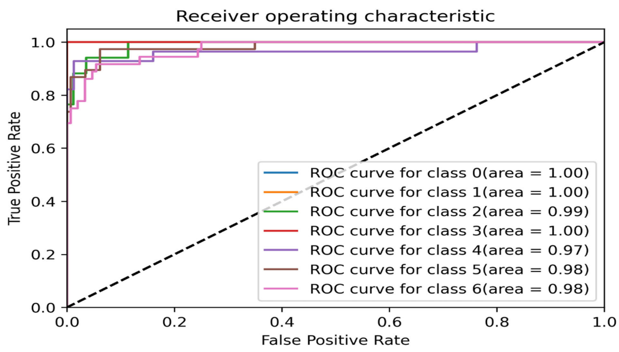

FIGURE 8. ROC curve obtained on SAVEE dataset using our proposed model.

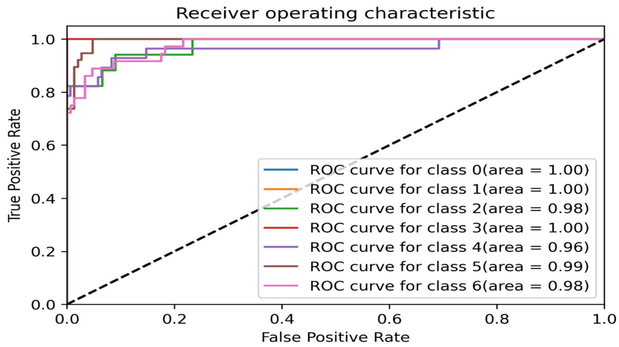

FIGURE 9. ROC curve obtained on EmoDB dataset using our proposed model.

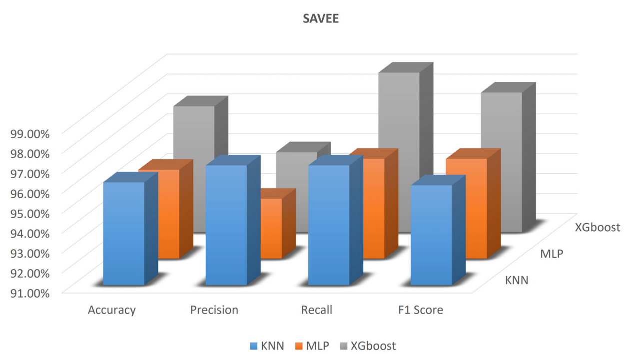

FIGURE 10. Performance comparison of the proposed GREO algorithm using three different classifiers over SAVEE dataset.

the classification accuracy of XGBoost classifier is 97.31% for SAVEE dataset and 98.46% for EmoDB dataset. It is also observed that after XGBoost classifier, the performance of KNN classifier is better than MLP classifier. The KNN classifier attains 96.15% and 97.13% classification accuricies on SAVEE and EmoDB datasets respectively. The MLP classifier performs the worst not only in terms of classification accuracies but also in case of training and testing times. We also have given a comparison bar-diagram indicating the performances of different classifiers on SAVEE and EmoDB datasets shown in Fig. 14 and Fig. 15 respectively.

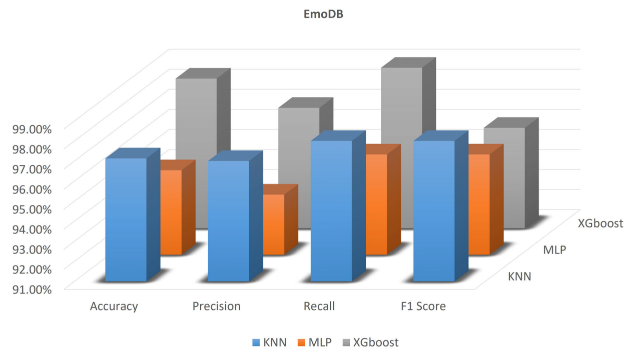

FIGURE 11. Performance comparison of the proposed GREO algorithm using three different classifiers over EmoDB dataset.

TABLE 5. Comparison of performance of our proposed GREO based FS model using different classifiers on SAVEE and EmoDB datasets.

Parameter	SAVEE Dataset			EmoDB Dataset		
	XGBoost	KNN	MLP	XGBoost	KNN	MLP
Accuracy	**97.31%**	96.15%	95.45%	**98.46%**	97.13%	95.21%
Precision	95%	97%	94%	97%	97%	94%
Recall	99%	97%	96%	99%	98%	96%
F1 Score	98%	96%	96%	96%	98%	95%

TABLE 6. Final set of parameters of XGBoost classifier for both the datasets giving optimum result.

Parameter	Value
Maximum Depth (‘max_{depth}’)	3
‘eta’ value	0.3
Number of ‘steps’	30
‘nthread’	5
‘objective’ function	‘multi:softprob’
Evaluation Matrix(‘$eval_{metric}$’)	‘auc’

E. TUNING OF HYPER-PARAMETERS OF XGBoost CLASSIFIER

To finalize the hyper-parameters of XGBoost classifier, different experiments are performed by varying a single parameter and fixing others. The most significant and performance determining parameters of XGBoost classifier are **Maximum Depth**, **Number of steps** and **eta value**. Apart from these, there are some other different hyper-parameters which effect the final performance such as the *objective function*, n^{th} *read* and so on. Fig. 16, Fig. 17 and Fig. 18 show the variation of final accuracy with respect to the variation of **maximum depth**, **eta value** and **number of steps** of the classifier respectively for both EmoDB and SAVEE datasets. It is quite evident from aforementioned figures that we get best classification results for both datasets with **maximum depth**, **eta** and **steps** equal to 3, 0.3 and 30 respectively. In addition, the final values of all other hyper-parameters are given in Table 6.

F. COMPARISON WITH OTHER OPTIMIZATION ALGORITHM BASED FS METHODS

We have given a comparison table of our proposed hybrid GREO based FS algorithm with 9 popularly used optimization algorithms for FS such as EO [18], GRO [50], PSO [53],

TABLE 7. Comparison of the proposed GREO based FS model with some state-of-the-art FS algorithms on SAVEE dataset.

Algorithm	Number of selected features	Accuracy	Precision	Recall	F1 Score
EO algorithm	123	96.66%	93%	97%	97%
GRO algorithm	98	96.32%	92%	98%	95%
PSO algorithm	67	81.32%	80%	83%	79%
GA	420	72.21%	73%	78%	72%
ASO algorithm	97	88.56%	88%	89%	89%
GWO algorithm	105	91.50%	92%	93%	92%
SFO algorithm	230	87.22%	88%	90%	90%
HS algorithm	187	89.79%	90%	93%	93%
GBO algorithm	170	93.86%	94%	94%	96%
Proposed GREO algorithm	**87**	**97.31%**	**95%**	**99%**	**98%**

TABLE 8. Comparison of the proposed GREO based FS model with some state-of-the-art FS algorithms on EmoDB dataset.

Algorithm	Number of selected features	Accuracy	Precision	Recall	F1 Score
EO algorithm	150	96.89%	95%	98%	97%
GRO algorithm	132	90.28%	91%	92%	90%
PSO algorithm	77	88.91%	87%	90%	88%
GA	219	80.20%	78%	84%	84%
ASO algorithm	66	93.22%	93%	91%	89%
GWO algorithm	106	95.14%	92%	96%	94%
SFO algorithm	159	92.11%	90%	93%	93%
HS algorithm	238	91.35%	93%	93%	92%
GBO algorithm	324	94.88%	95%	96%	94%
Proposed GREO Algorithm	**98**	**98.46%**	**97%**	**99%**	**96%**

TABLE 9. Performance comparison of our proposed GREO based FS model with some state-of-the-art works for SAVEE dataset.

Sl No.	Researchers	Feature Set Used	Method	Achieved Accuracy
1.	Mao et al. (2014) [37]	Deep Features learnt by CNN itself	CNN	73.60%
2.	Zhen – Tao Liu et al. (2018) [38]	MFCC features	GA-BEL Model	76.40%
3.	Dung Nguyen et al. (2018) [39]	Neural Network Learns its own features	PathNet	93.75%
4.	Noushin Hajarolasvadi et al. (2019) [40]	Deep Features of Neural Network	3D CNN-Based approach with K-Means Clustering	81.05%
5.	P. Barros et al. (2015) [41]	Deep Features	Cross Channel Deep Neural Architecture	92.00%
6.	Ingryd Pereira et al.(2018) [59]	Spectrogram conversion using Short Time Fourier Transform(STFT)	Pre-Trained BEGAN	40%
7.	E. Avtos et al. (2018) [43]	MFCC	SVM	77.4%
8.	**Our Proposed work**	**Concatenation of LPC and LPC correlation features**	**FS using GREO model and classification with XGBoost Classifier**	**97.31%**

TABLE 10. Performance comparison of our proposed GREO based FS model with some state-of-the-art works for EmoDB dataset.

Sl No.	Researchers	Feature Set USed	Method	Achieved Accuracy
1.	Mao et al. (2014) [37]	Deep Features learnt by CNN itself	CNN	85.20%
2.	Deng et al.(2013) [44]	LLD features, like ZCR, RMS, energy , MFCC, HNR, frequency of pitch	De-noising Autoencoder	57.9%
3.	Taner Danisman et al. [45]	MFCC, total energy and F0	SVM	63.5%
4.	Albornoz et al. (2011) [46]	MFCC, Spectral, prosodic features and log spectrum	SVM, MLP, GMM, HMM and Hierarchical classifier	71.5%
5.	Shen et al. (2011) [47]	LPCC, MFCC, pitch , Energy and LPCMCC	SVM	82.5%
6.	Wang et al. [48]	MFCC and fourier parameters	SVM	88.88%
7.	Wu et al. (2011) [24]	Prosodic features, Speaking rate features, features based on TEO and ZCR	SVM	91.3%
8.	**Our Proposed Work**	**Concatenation of LPC and LPC Correlation featrues**	**FS using GREO model and classification with XGBoost Classifier**	**98.46%**

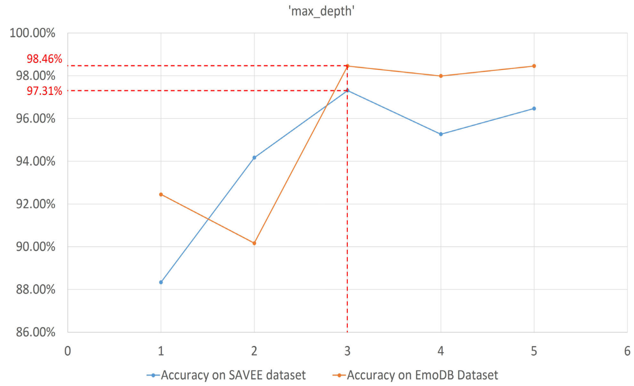

FIGURE 12. Variation of performance (in terms of accuracy) with respect to the depth of the XGBoost classifier for both SAVEE and EmoDB datasets.

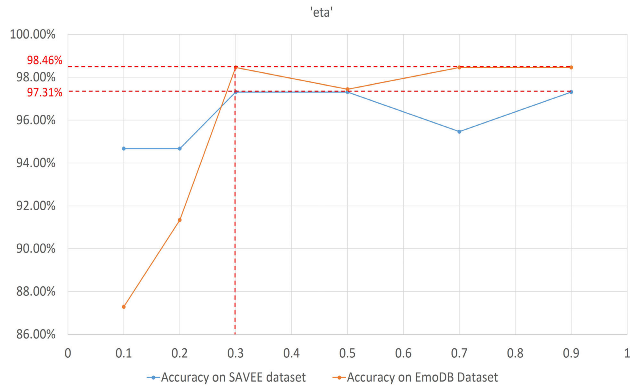

FIGURE 13. Variation of accuracy with respect to 'eta' value for both SAVEE and EmoDB datasets.

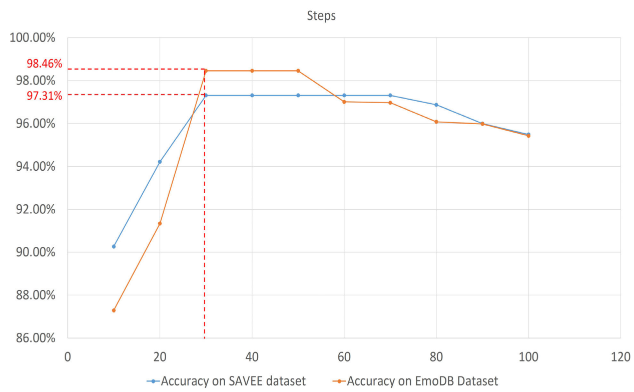

FIGURE 14. Variation of accuracy on SAVEE and EmoDB datasets with respect to steps or the number of iterations for XGBoost classifier.

GA [49], Atom Search optimization (ASO) algorithm [51], Gray Wolf Optimization (GWO) algorithm [52], Sail Fish Optimization (SFO) algorithm [54], HS algorithm [55] and Gradient Based optimization (GBA) algorithm [56]. The final comparative results are given in the Table 7 and Table 8 for SAVEE and EmoDB datasets respectively. In Tables 7 and 8, along with classification accuracy and other evaluation metrics (like Precision, Recall and F1 Score), we have also given the number of optimal features selected at the final stage of optimization.

It is quite evident from the comparison tables that our proposed hybrid model has outperformed other optimization algorithms not only in terms of accuracy but also with respect to the number of features selected as the best solution. From Table 7, it is evident that when the EO and GRO algorithms are applied individually, they achieved 96.66% and 96.32% classification accuracies using 123 and 98 number of features as the finally chosen feature subset on SAVEE dataset. Whereas when they are hybridized using our proposed approach, the GREO algorithm achieves 97.31% accuracy while utilizing only 87 features as the final feature set. However, it is also observed that the PSO algorithm selects only 67 features as the final feature space but the classification accuracy is compromised, which is found to be 81.32%.

Similarly, from Table 8, in case of EmoDB dataset, it is found that the EO and GRO algorithms individually achieves 96.89% and 90.28% using 150 and 132 number of selected features respectively. On the other hand, our proposed hybrid model attains 98.46% accuracy while utilizing only 87 features containing final feature set. Therefore, it is clearly seen that our proposed GREO based FS model selects less number of optimal features as the final feature set while giving much better classification accuracy, which effectively concludes the efficiency of our model.

G. COMPARISON WITH PAST SER METHODS

Table 9 and Table 10 represent a comparative study of currently revealed SER works with our proposed method which positively shows that our proposed model has outperformed all of them and able to establish state-of-art results on both SAVEE and EmoDB datasets by achieving 97.31% and 98.46% classification accuracies respectively.

From Table 9, it is evident that Nguyen *et al.* [39] introduce PathNet structure which gets 93.75% accuracy and it holds the second position in the list. P. Barros *et al.* [41] also reach to a quite decent classification accuracy of 92.00% with Cross Channel Deep Neural Architecture. Table 10 also shows similar results. Wu *et al.* [48] have used machine learning techniques to extract handcraft features and classified using traditional SVM classifier and been able to achieved quite promising result with 91.3% accuracy on EmoDB dataset. Table 9 and Table 10 clearly indicate that our model performs not only superior to other models but also with a good margin of difference.

V. CONCLUSION

In this work, we have proposed a hybrid meta-heuristic FS method named as GREO which is actually based on two recently introduced optimization algorithms, EO and GRO. The proposed FS method has been evaluated on two well-known publicly available SER datasets, namely SAVEE and EmoDB, and our proposed method has achieved recognition accuracies of 97.31% and 98.46% respectively. The proposed FS algorithm has been compared with eight popular optimization algorithms such as EO, GRO, PSO, GA, ASO, GWO, SFO and HS, and the obtained results have proven the superiority of GREO algorithm over those methods. Moreover, research in the field of SER has been a key interest in recent times, and many deep learning and machine learning based models have been proposed by the researchers to recognize the emotions from speech. Usually, deep learning based models perform better than the machine learning based models. However, in our task, we have achieved the state-of-the-art results on two open-access datasets, and obtained better results than some deep learning based models also.

Though our results are quite satisfactory, still there are some rooms for improvement of the proposed model which are listed below:

- We have considered the features obtained from traditional feature extraction methods (i.e., LPC and LPCC). In future, we can use feature vectors obtained from some deep learning based models.
- Here, we have implemented GREO on a randomly generated population, however, using any clustering algorithm, we can choose the population on the basis of certain properties of the dataset. This may help us to increase performance of the FS model.
- No Free Lunch(NFL) algorithm clearly explains that there is no optimization algorithm, which can optimize every single problem. So, hybrid approach of other optimization algorithms can be tried out to improve performance the overall system.

REFERENCES

[1] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649.
[2] P. Dreuw, D. Rybach, T. Deselaers, M. Zahedi, and H. Ney, "Speech recognition techniques for a sign language recognition system," in *Proc. 8th Annu. Conf. Int. Speech Commun. Assoc.*, 2007, pp. 1–4.
[3] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Trans. Audio, Speech Language Process.*, vol. 22, no. 10, pp. 1533–1545, Oct. 2015.
[4] T.-S. Chi, L.-Y. Yeh, and C.-C. Hsu, "Robust emotion recognition by spectro-temporal modulation statistic features," *J. Ambient Intell. Humanized Comput.*, vol. 3, no. 1, pp. 47–60, Mar. 2012.
[5] S. Saha, M. Ghosh, S. Ghosh, S. Sen, P. K. Singh, Z. W. Geem, and R. Sarkar, "Feature selection for facial emotion recognition using cosine similarity-based harmony search algorithm," *Appl. Sci.*, vol. 10, no. 8, p. 2816, Apr. 2020.
[6] S. Latif, R. Rana, S. Younis, J. Qadir, and J. Epps, "Transfer learning for improving speech emotion classification accuracy," 2018, *arXiv:1801.06353*. [Online]. Available: http://arxiv.org/abs/1801.06353
[7] C. C. Chibelushi and F. Bourel, "Facial expression recognition: A brief tutorial overview," *CVonline, On-Line Compendium Comput. Vis.*, pp. 1–5, Jan. 2003.
[8] M. G. S. Ortega, L. F. Rodríguez, and J. O. Gutierrez-Garcia, "Towards emotion recognition from contextual information using machine learning," *J. Ambient Intell. Humanized Comput.*, vol. 11, pp. 1–21, Sep. 2019.
[9] H. Mukherjee, S. M. Obaidullah, K. C. Santosh, S. Phadikar, and K. Roy, "A lazy learning-based language identification from speech using MFCC-2 features," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 1, pp. 1–14, Jan. 2020.
[10] S. Shetty and S. Hegde, "Automatic classification of carnatic music instruments using MFCC and LPC," in *Data Management, Analytics and Innovation*. Singapore: Springer, 2020, pp. 463–474.
[11] B. Karan, S. S. Sahu, and K. Mahto, "Parkinson disease prediction using intrinsic mode function based features from speech signal," *Biocybernetics Biomed. Eng.*, vol. 40, no. 1, pp. 249–264, Jan. 2020.
[12] J. C. Saldanha and M. Suvarna, "Perceptual linear prediction feature as an indicator of dysphonia," in *Advances in Control Instrumentation Systems*. Singapore: Springer, 2020, pp. 51–64.
[13] H. Gupta and D. Gupta, "LPC and LPCC method of feature extraction in speech recognition system," in *Proc. 6th Int. Conf.-Cloud Syst. Big Data Eng. (Confluence)*, Jan. 2016, pp. 498–502.
[14] M. Ghosh, R. Guha, R. Sarkar, and A. Abraham, "A wrapper-filter feature selection technique based on ant colony optimization," *Neural Comput. Appl.*, vol. 32, pp. 1–19, Apr. 2019.
[15] M. Ghosh, R. Guha, I. Alam, P. Lohariwal, D. Jalan, and R. Sarkar, "Binary genetic swarm optimization: A combination of GA and PSO for feature selection," *J. Intell. Syst.*, vol. 29, no. 1, pp. 1598–1610, Sep. 2019.
[16] K. K. Ghosh, S. Ahmed, P. K. Singh, Z. W. Geem, and R. Sarkar, "Improved binary sailfish optimizer based on adaptive β-Hill climbing for feature selection," *IEEE Access*, vol. 8, pp. 83548–83560, 2020.
[17] B. Chatterjee, T. Bhattacharyya, K. K. Ghosh, P. K. Singh, Z. W. Geem, and R. Sarkar, "Late acceptance hill climbing based social ski driver algorithm for feature selection," *IEEE Access*, vol. 8, pp. 75393–75408, 2020.
[18] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: A novel optimization algorithm," *Knowl.-Based Syst.*, vol. 191, Mar. 2020, Art. no. 105190.
[19] A. F. Nematollahi, A. Rahiminejad, and B. Vahidi, "A novel meta-heuristic optimization method based on golden ratio in nature," *Soft Comput.*, vol. 24, no. 2, pp. 1117–1151, Jan. 2020.
[20] R. Nakatsu, A. Solomides, and N. Tosa, "Emotion recognition and its application to computer agents with spontaneous interactive capabilities," in *Proc. IEEE Int. Conf. Multimedia Comput. Syst.*, Oct. 1999, pp. 343–351.
[21] B. Adams, C. Breazeal, R. A. Brooks, and B. Scassellati, "Humanoid robots: A new kind of tool," *IEEE Intell. Syst.*, vol. 15, no. 4, pp. 25–31, Jul. 2000.
[22] E. H. Kim, K. H. Hyun, S. H. Kim, and Y. K. Kwak, "Emotion interactive robot focus on speaker independently emotion recognition," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics*, Sep. 2007, pp. 1–6.
[23] B. Schuller, G. Rigoll, and M. Lang, "Hidden Markov model-based speech emotion recognition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 2, Apr. 2003, p. II-1.
[24] C.-H. Wu and W.-B. Liang, "Emotion recognition of affective speech based on multiple classifiers using acoustic-prosodic information and semantic labels," *IEEE Trans. Affect. Comput.*, vol. 2, no. 1, pp. 10–21, Jan. 2011.
[25] J. Rong, G. Li, and Y.-P.-P. Chen, "Acoustic feature selection for automatic emotion recognition from speech," *Inf. Process. Manage.*, vol. 45, no. 3, pp. 315–328, May 2009.
[26] H. M. Fayek, M. Lech, and L. Cavedon, "Evaluating deep learning architectures for speech emotion recognition," *Neural Netw.*, vol. 92, pp. 60–68, Aug. 2017.
[27] Y. Huang, K. Tian, A. Wu, and G. Zhang, "Feature fusion methods research based on deep belief networks for speech emotion recognition under noise condition," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 5, pp. 1787–1798, May 2019.
[28] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," *Biomed. Signal Process. Control*, vol. 47, pp. 312–323, Jan. 2019.
[29] N. Cummins, S. Amiriparian, G. Hagerer, A. Batliner, S. Steidl, and B. Schuller, "An image-based deep spectrum feature representation for the recognition of emotional speech," in *Proc. 25th ACM Int. Conf. Multimedia*, 2017, pp. 478–484, doi: 10.1145/3123266.3123371.
[30] Mustaqeem and S. Kwon, "A CNN-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 20, no. 1, p. 183, Dec. 2019.
[31] L. Chen, W. Su, Y. Feng, M. Wu, J. She, and K. Hirota, "Two-layer fuzzy multiple random forest for speech emotion recognition in human-robot interaction," *Inf. Sci.*, vol. 509, pp. 150–163, Jan. 2020.
[32] K. K. Ghosh, R. Guha, S. K. Bera, N. Kumar, and R. Sarkar, "S-shaped versus V-shaped transfer functions for binary manta ray foraging optimization in feature selection problem," Jadavpur Univ., Kolkata, India, Tech. Rep., 2020, doi: 10.21203/rs.3.rs-28684/v1.
[33] K. K. Ghosh, P. K. Singh, J. Hong, Z. W. Geem, and R. Sarkar, "Binary social mimic optimization algorithm with X-shaped transfer function for feature selection," *IEEE Access*, vol. 8, pp. 97890–97906, 2020.
[34] S. Ahmed, K. K. Ghosh, P. K. Singh, Z. W. Geem, and R. Sarkar, "Hybrid of harmony search algorithm and ring theory-based evolutionary algorithm for feature selection," *IEEE Access*, vol. 8, pp. 102629–102645, 2020.

[35] L. O. Sawada, L. Y. Mano, J. R. T. Neto, and J. Ueyama, "A module-based framework to emotion recognition by speech: A case study in clinical simulation," *J. Ambient Intell. Humanized Comput.*, pp. 1–10, Apr. 2019.
[36] Z. Chen, F. Jiang, Y. Cheng, X. Gu, W. Liu, and J. Peng, "XGBoost classifier for DDoS attack detection and analysis in SDN-based cloud," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jan. 2018, pp. 251–256.
[37] Q. Mao, M. Dong, Z. Huang, and Y. Zhan, "Learning salient features for speech emotion recognition using convolutional neural networks," *IEEE Trans. Multimedia*, vol. 16, no. 8, pp. 2203–2213, Dec. 2014.
[38] Z.-T. Liu, Q. Xie, M. Wu, W.-H. Cao, Y. Mei, and J.-W. Mao, "Speech emotion recognition based on an improved brain emotion learning model," *Neurocomputing*, vol. 309, pp. 145–156, Oct. 2018.
[39] D. Nguyen, K. Nguyen, S. Sridharan, I. Abbasnejad, D. Dean, and C. Fookes, "Meta transfer learning for facial emotion recognition," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 3543–3548.
[40] N. Hajarolasvadi and H. Demirel, "3D CNN-based speech emotion recognition using K-Means clustering and spectrograms," *Entropy*, vol. 21, no. 5, p. 479, May 2019.
[41] P. Barros, C. Weber, and S. Wermter, "Emotional expression recognition with a cross-channel convolutional neural network for human-robot interaction," in *Proc. IEEE-RAS 15th Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2015, pp. 582–587.
[42] K. H. Sheikh, S. Ahmed, K. Mukhopadhyay, P. K. Singh, J. H. Yoon, Z. W. Geem, and R. Sarkar, "EHHM: Electrical harmony based hybrid meta-heuristic for feature selection," *IEEE Access*, vol. 8, pp. 158125–158141, 2020.
[43] E. Avots, T. Sapiński, M. Bachmann, and D. Kamińska, "Audiovisual emotion recognition in wild," *Mach. Vis. Appl.*, vol. 30, no. 5, pp. 975–985, Jul. 2019.
[44] J. Deng, Z. Zhang, E. Marchi, and B. Schuller, "Sparse autoencoder-based feature transfer learning for speech emotion recognition," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, Sep. 2013, pp. 511–516.
[45] T. Danisman and A. Alpkocak, "Emotion classification of audio signals using ensemble of support vector machines," in *Proc. Int. Tutorial Res. Workshop Perception Interact. Technol. Speech-Based Syst.* Berlin, Germany: Springer, Jun. 2008, pp. 205–216.
[46] E. M. Albornoz, D. H. Milone, and H. L. Rufiner, "Spoken emotion recognition using hierarchical classifiers," *Comput. Speech Lang.*, vol. 25, no. 3, pp. 556–570, Jul. 2011.
[47] P. Shen, Z. Changjun, and X. Chen, "Automatic speech emotion recognition using support vector machine," in *Proc. Int. Conf. Electron. Mech. Eng. Inf. Technol.*, vol. 2, Aug. 2011, pp. 621–625.
[48] K. Wang, N. An, B. N. Li, Y. Zhang, and L. Li, "Speech emotion recognition using Fourier parameters," *IEEE Trans. Affect. Comput.*, vol. 6, no. 1, pp. 69–75, Jan. 2015.
[49] D. E. Goldberg and M. P. Samtani, "Engineering optimization via genetic algorithm," in *Electronic Computation*. Reston, VA, USA: American Society of Civil Engineer (ASCE), Feb. 1986, pp. 471–482.
[50] A. F. Nematollahi, A. Rahiminejad, and B. Vahidi, "A novel meta-heuristic optimization method based on golden ratio in nature," *Soft Comput.*, vol. 24, no. 2, pp. 1117–1151, Jan. 2020.
[51] W. Zhao, L. Wang, and Z. Zhang, "A novel atom search optimization for dispersion coefficient estimation in groundwater," *Future Gener. Comput. Syst.*, vol. 91, pp. 601–610, Feb. 2019.
[52] E. Emary, H. M. Zawbaa, C. Grosan, and A. E. Hassenian, "Feature subset selection approach by gray-wolf optimization," in *Proc. Afro-Eur. Conf. Ind. Advancement*. Cham, Switzerland: Springer, 2015, pp. 1–13.
[53] P. C. Fourie and A. A. Groenwold, "The particle swarm optimization algorithm in size and shape optimization," *Struct. Multidisciplinary Optim.*, vol. 23, no. 4, pp. 259–267, May 2002.
[54] S. Shadravan, H. R. Naji, and V. K. Bardsiri, "The sailfish optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems," *Eng. Appl. Artif. Intell.*, vol. 80, pp. 20–34, Apr. 2019.
[55] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, Feb. 2001.
[56] I. Ahmadianfar, O. Bozorg-Haddad, and X. Chu, "Gradient-based optimizer: A new Metaheuristic optimization algorithm," *Inf. Sci.*, vol. 540, pp. 131–159, Nov. 2020.
[57] P. Jackson and S. Haq, "Surrey audio-visual expressed emotion (SAVEE) database," Univ. Surrey, Guildford, U.K., Tech. Rep., 2014.
[58] F. Burkhardt, A. Paeschke, M. Rolfes, W. F. Sendlmeier, and B. Weiss, "A database of German emotional speech," in *Proc. 9th Eur. Conf. Speech Commun. Technol.*, 2005, pp. 1–4.
[59] I. Pereira, D. Santos, A. Maciel, and P. Barros, "Semi-supervised model for emotion recognition in speech," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, Oct. 2018, pp. 791–800.
[60] R. Guha, A. H. Khan, P. K. Singh, R. Sarkar, and D. Bhattacharjee, "CGA: A new feature selection model for visual human action recognition," *Neural Comput. Appl.*, 2020, doi: 10.1007/s00521-020-05297-5.
[61] A. Das, S. Guha, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "A hybrid meta-heuristic feature selection method for identification of Indian spoken languages from audio signals," *IEEE Access*, vol. 8, pp. 181432–181449, 2020, doi: 10.1109/ACCESS.2020.3028241.
[62] T. Bhattacharyya, B. Chatterjee, P. K. Singh, J. H. Yoon, Z. W. Geem, and R. Sarkar, "Mayfly in harmony: A new hybrid meta-heuristic feature selection algorithm," *IEEE Access*, early access, Oct. 16, 2020, doi: 10.1109/ACCESS.2020.3031718.
[63] S. Guha, A. Das, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "Hybrid feature selection method based on harmony search and naked mole-rat algorithms for spoken language identification from audio signals," *IEEE Access*, vol. 8, pp. 182868–182887, 2020, doi: 10.1109/ACCESS.2020.3028121.
[64] M. Ghosh, R. Guha, P. K. Singh, V. Bhateja, and R. Sarkar, "A histogram based fuzzy ensemble technique for feature selection," *Evol. Intell.*, vol. 12, no. 4, pp. 713–724, Dec. 2019.
[65] I. H. Osman and J. P. Kelly, "Meta-heuristics theory and applications," *J. Oper. Res. Soc.*, vol. 48, no. 6, pp. 657–657, Dec. 1997.
[66] N. A. Golilarz, A. Addeh, H. Gao, L. Ali, A. M. Roshandeh, H. M. Munir, and R. U. Khan, "A new automatic method for control chart patterns recognition based on ConvNet and Harris hawks meta heuristic optimization algorithm," *IEEE Access*, vol. 7, pp. 149398–149405, 2019.
[67] A. Das, H. S. Das, and H. S. Das, "Impact of cuckoo algorithm in speech processing," in *Applications of Cuckoo Search Algorithm and its Variants*. Singapore: Springer, 2020, pp. 207–228.
[68] A. Singh, "Speech emotion recognition using enhanced cat swarm optimization algorithm," *Int. J. Inf. Technol.*, vol. 6, no. 5, pp. 1–9, 2020.
[69] C. K. Yogesh, M. Hariharan, R. Ngadiran, A. H. Adom, S. Yaacob, C. Berkai, and K. Polat, "A new hybrid PSO assisted biogeography-based optimization for emotion and stress recognition from speech signal," *Expert Syst. Appl.*, vol. 69, pp. 149–158, Mar. 2017.
[70] R. Guha, M. Ghosh, P. K. Singh, R. Sarkar, and M. Nasipuri, "M-HMOGA: A new multi-objective feature selection algorithm for handwritten numeral classification," *J. Intell. Syst.*, vol. 29, no. 1, pp. 1453–1467, Jun. 2019.
[71] I. Chatterjee, M. Ghosh, P. K. Singh, R. Sarkar, and M. Nasipuri, "A clustering-based feature selection framework for handwritten indic script classification," *Expert Syst.*, vol. 36, no. 6, p. e12459, Dec. 2019, doi: 10.1111/exsy.12459.
[72] P. K. Singh, S. Das, R. Sarkar, and M. Nasipuri, "Feature selection using harmony search for script identification from handwritten document images," *J. Intell. Syst.*, vol. 27, no. 3, pp. 465–488, Jul. 2018.
[73] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.

ARIJIT DEY is currently pursuing the bachelor's degree in computer science and engineering with the Maulana Abul Kalalm Azad University of Technology, Kolkata, India. His research interests include machine learning, deep learning, and optimization algorithms.

SOHAM CHATTOPADHYAY is currently pursuing the bachelor's degree in electrical engineering with Jadavpur University, Kolkata, India. His research interests include machine learning, deep learning, and optimization algorithms.

PAWAN KUMAR SINGH (Member, IEEE) received the B.Tech. degree in information technology from the West Bengal University of Technology, in 2010, and the M.Tech. degree in computer science and engineering and the Ph.D. (engineering) degree from Jadavpur University (J.U.), in 2013 and 2018, respectively. He also received the RUSA 2.0 Fellowship for pursuing his postdoctoral research from J.U. in 2019. He is currently working as an Assistant Professor with the Department of Information Technology, J.U. He has published more than 50 research articles in peer-reviewed journals and international conferences. His current research interests include computer vision, pattern recognition, handwritten document analysis, image and video processing, feature optimization, machine learning, deep learning, and artificial intelligence. He is also a member of The Institution of Engineers (India) and the Association for Computing Machinery (ACM) and a Life Member of the Indian Society for Technical Education (ISTE), New Delhi, and Computer Society of India (CSI).

ALI AHMADIAN (Member, IEEE) received the Ph.D. degree (Hons.) from Universiti Putra Malaysia (UPM), in 2014. He is currently a Fellow Researcher with the Institute of Industry Revolution 4.0, UKM. As a Young Researcher, he is dedicated to research in applied mathematics. In general, his primary mathematical focus is the development of computational methods and models for problems arising in AI, biology, physics, and engineering under fuzzy and fractional calculus (FC); in this context, I have worked on projects related to drug delivery systems, acid hydrolysis in palm oil frond, and carbon nanotubes dynamics, Bloch equations, and viscosity. He could successfully receive 15 national and international research grants and selected as the 1% top reviewer in the fields of mathematics and computer sciences recognized by Publons, from 2017 to 2019. He is also a member of editorial board of *Progress in Fractional Differentiation and Applications* (Natural Sciences Publishing) and a Guest Editor of *Advances in Mechanical Engineering* (SAGE), *Symmetry* (MDPI), *Frontier in Physics* (Frontiers), and the *International Journal of Hybrid Intelligence* (Inderscience Publishers). He is the author of more than 80 research articles published in the reputed journals, including the IEEE Transactions on Fuzzy Systems, *Fuzzy Sets and Systems*, *Communications in Nonlinear Sciences and Numerical Simulations*, *Computational Physics*, and so on. He also presented his research works in 38 international conferences held in Canada, Serbia, China, Turkey, Malaysia, and UAE. He was a member of programme committee in a number of International conferences in fuzzy field at Japan, China, Turkey, South Korea, and Malaysia. He is also serving as a referee in more than 80 reputed international journals.

MASSIMILIANO FERRARA received the master's degree *(cum laude)* in economics and the Ph.D. degree *(cum laude)* from the University of Messina, and the double master's degree *(cum laude)* from the University of Naples Federico II and the Scuola Normale Superiore di Pisa. He was a General Counsel of the Fondazione Banco di Napoli, the Vice Rector with the University for foreigners "Dante Alighieri" of Reggio Calabria, and the Head of the Regione Calabria Department for Cultura, Research and Education. He was a Visiting Professor with Harvard University, Cambridge, MA, USA, Western Michigan University, USA, Morgan State University, Baltimora, USA, and the Northeastern University di Boston, USA. He has been a Research Affiliate with the Invernizzi Center for Research on Innovation, Organization, Strategy and Entrepreneurship (ICRIOS), University Bocconi of Milan, since 2013, the President of the Scientific Committee of MEDAlics Research Center, and the Scientific Director of DECISIONS Laboratory. He is currently with the Center for Dynamics, Dresden University of Technology, Germany. He is also a Full Professor with the Mathematical Economics, Statistics, Business Analytics & Decision Theory, Applied Economics, Mediterranea University of Reggio Calabria, where he is also the Chairman of the Department of Law, Economics & Human Sciences and a member of the Academic Senate. He is the author and coauthor of up 200 articles on peer-review and ISI journal and ten research monographs. He has also been Knight Order of Merit of the Italian Republic, since 2010, "for international scientific merits," and he is in the prestigious US Encyclopedia Hmolpedia on thermodynamics and the theoretical and applied physics, for offering a decisive contribution to the creation and development of the scientific theory called "Economic geometric dynamics." He is also an Editor, a Co-Editor, an Associate Editor, and a Referee of reputable international scientific journal in economics, and pure and applied mathematics.

RAM SARKAR (Senior Member, IEEE) received the B.Tech. degree in computer science and engineering from the University of Calcutta, in 2003, and the M.E. degree in computer science and engineering and the Ph.D. (engineering) degree from Jadavpur University, in 2005 and 2012, respectively. He also received Fulbright-Nehru Fellowship (USIEF) for pursuing his postdoctoral research from the University of Maryland, College Park, MD, USA, from 2014 to 2015. He joined the Department of Computer Science and Engineering, Jadavpur University, as an Assistant Professor, in 2008, where he is currently working as an Associate Professor. His current research interests include image processing, pattern recognition, machine learning, and bioinformatics.