Channel Selection for EEG Emotion Recognition via an Enhanced Firefly Algorithm with Brightness-Distance Attraction

Ben Niu, Gemin Liang, Bang Tao, Chao Fu, Shuang Geng(✉), Yang Wang, and Bowen Xue

College of Management, Shenzhen University, Shenzhen 518060, China
lianggemin2021@email.szu.edu.cn, {gs,wangyanghim}@szu.edu.cn

Abstract. Accurate recognition of human emotions through EEG data is of great significance in human-computer interaction, mental health, intelligent medical care and other fields. EEG signal contains a large number of meaningful and extractable features. Therefore, effective feature selection plays an essential role in reducing feature dimensions and avoiding redundancy. In order to select the emotion related features from hundreds of features and achieve better emotion recognition results, we propose an enhanced firefly algorithm (EFA) for EEG emotion recognition, which is based on brightness-distance based attraction and roulette-based local search strategies. Then, we apply EFA to select features for EEG emotion recognition and provide a novel encoding method of fireflies to distinguish the importance of channels and bands respectively. We conduct comparative experiments to evaluate the performance of EFA on DEAP database. The experimental results confirm the superiority of the proposed method in AUC score.

Keywords: Emotion recognition · Feature selection · Swarm intelligence · Fire-fly algorithm · EEG

1 Introduction

Emotion plays an important role in human life. With the development of human-computer interaction (HCI) technology, there are many researches on emotion recognition in recent years. Generally, emotion recognition is mainly based on two kinds of signals: physical signals such as facial expression [1] and speech [2], and physiological signals such as electroencephalogram (EEG) [3] and respiration (RSP) [4]. Among them, EEG has attracted extensive attention of researchers because it is more objective and handier than other signals.

EEG-based emotion recognition usually involves the following steps: preprocessing, feature extraction, feature selection and classification. There are different kinds of features can be extracted from EEG signals, such as time-domain features, frequency-domain features, time-frequency features and nonlinear features [5]. Since most EEG signals are collected in the form of time domain, and it need to transform the signals

Y. Xu et al. (Eds.): ML4CS 2022, LNCS 13657, pp. 182–196, 2023.
https://doi.org/10.1007/978-3-031-20102-8_15

from time domain to frequency domain by some algorithms, like fast Fourier transfer (FFT) [6], short-time Fourier transform [7] and wavelet transform [8]. Furthermore, band power, power spectral density and differential entropy can be calculated from the frequency bands.

Since hundreds of features can be extracted from EEG signals, it is necessary to select features before classification to reduce the feature dimension and avoid redundancy. Principal component analysis (PCA) is the most commonly used dimension reduction technique, which decomposes EEG signals into independent components and removes interference [9, 10]. Swarm intelligence (SI) algorithms have been proposed to solve the problem of feature selection and achieved good results. Particle swarm optimization (PSO) is utilized to select the emotion related features for EEG emotion recognition [11]. Nakisa et al. [12] apply ant colony optimization and PSO for the feature selection of EEG datasets.

Recently, Firefly Algorithm (FA) [13] has been applied in various fields due to its advantages of less parameters and simple operation. In the aspect of distance-based attraction mechanism, FA can automatically divide the whole colony into multiple sub-colonies, which can naturally and effectively deal with nonlinear and multi-modal optimization problems [14]. However, it has the limitations of slow convergence speed, early maturity and low accuracy in the late iterations.

To tackle the limitations of FA and achieve better emotion recognition results for EEG data, we propose an enhanced FA for emotion recognition. Firstly, we transform signals from time domain to frequency domain by FFT and calculate their band power characteristics. Then, we propose an enhanced FA with brightness-distance based attraction and roulette-based local search strategy. Subsequently, we apply this enhanced FA to select the most important features for EEG emotion recognition. In brief, the main contributions of this work are as follows:

1. We propose an Enhanced Firefly Algorithm (EFA) with two strategies: brightness-distance based attraction (BDA) and roulette-based local search strategy (RLS).
2. We apply the enhanced firefly algorithm to select features for EEG emotion recognition and provide a novel encoding method of fireflies to distinguish the importance of channels and bands respectively.
3. We conduct extensive experiments on a public database DEAP [15] to demonstrate the effectiveness of proposed method. The results show that EFA based feature selection algorithm outperforms other competitive feature selection algorithms in EEG-based emotion recognition.

2 Related Work

2.1 EEG Feature Extraction and Selection

Feature extraction and selection play an important role in EEG-based emotion recognition. Feature extraction is mainly to reduce the dimension of EEG data and extract emotion related features from EEG data to study the emotional state of subjects. As a key component of emotion recognition, the quality of features directly determines the performance of emotion recognition model.

EEG Feature Extraction. In the existing research on EEG emotion recognition, there are four kinds of extracted EEG signal features: time-domain features, frequency-domain features, time-frequency features and nonlinear features [5]. Frequency domain features are the most widely used features in emotion recognition based on EEG, such as band power, power spectrum and power spectral density. FFT is used to decompose the EEG signals into five bands: delta, theta, alpha, beta and gamma, and then extract the log band energy of these five bands as features [6]. Li et al. [7] adopt short-time Fourier transform for time-frequency transformation and calculate the power spectral density of four frequency bands respectively.

EEG Feature Selection. In the research of EEG emotion recognition, more electrodes are usually placed on the subject's scalp to obtain more abundant emotional information. However, with the increase of the number of electrodes, the number of features rises sharply, which will lead to excessive calculation and reduce the real-time performance of the system. PCA is used for dimensionality reduction of high-dimensional data, which projects the data onto the principal components, so as to generate the main features of emotion-related EEG and remove useless or noisy information [9, 10]. ReliefF-based channel selection algorithm is applied to reduce the number of channels used in classification task [16]. SI algorithm is widely used in feature selection and its effectiveness has been proved. In the recent years, researchers have applied SI algorithm to feature selection for EEG-based emotion recognition and achieved satisfactory results, such as PSO [11], FA [17], grey wolf optimizer [18] and cuckoo search [19].

2.2 Firefly Algorithm

FA [13] is a heuristic swarm intelligence approach inspired by the flashing behavior of fireflies. The optimization problem is solved by simulating the mutual attraction and movement of fireflies caused by foraging and communication in nature. Fireflies with less brightness are attracted to the brighter one. The brightness of a firefly is determined by the objective function. For the maximum optimization problem, the brightness of a firefly can be simply proportional to the fitness of the objective function. Mutual attraction depends on the light intensity perceived by the firefly, which diminishes with distance. If there is no firefly brighter than a specific firefly, it will move randomly.

As the attractiveness of a firefly is proportional to the light intensity which decrease by distance, the attraction of fireflies is defined as:

$$\beta = \beta_0 e^{-\gamma {r_{ij}}^2} \tag{1}$$

where β_0 is is the attraction of the firefly itself when $r = 0$, γ is light absorption coefficient and r_{ij} is the distance between the fireflies.

The distance between two fireflies can be calculated by Euclidean distance as follow:

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

where x_i and x_j are the position of firefly i and j respectively, $x_{i,k}$ is the k-th component of the firefly i and d is the dimension of the problem.

Each firefly i compares its brightness with that of other firefly j. If firefly j is brighter than i, firefly i moves toward firefly j as Eq. (3). Otherwise, firefly i move randomly as Eq. (4).

$$x_i(t+1) = x_i(t) + \beta\big(x_j(t) - x_i(t)\big) + \alpha r \tag{3}$$

$$x_i(t+1) = x_i(t) + \alpha r \tag{4}$$

where α is a parameter that determines the random search and it decreases with the increase of the number of iterations t, and r is a d-dimensional Gaussian random vector.

For EEG emotion recognition, it is significant to select features highly correlated to emotion from hundreds of extracted features. Although FA has been used in feature selection, it has the limitations of slow convergence speed, early maturity and low accuracy in the late iterations. To achieve better emotion recognition results for EEG data, this paper develops an enhanced FA for feature selection of EEG signals.

3 Enhanced Firefly Algorithm for Emotion Recognition

3.1 General Framework

In this section, we propose an enhanced FA feature selection for EEG emotion recognition. The general framework is presented in Fig. 1. At the first step, we transform signals from time domain to frequency domain by FFT and calculate their band power characteristics. Then, we propose an enhanced FA with brightness-distance based attraction (BDA) and roulette-based local search strategy (RLS). Subsequently, we apply this enhanced FA to select the most important features for EEG emotion recognition. More details are explained as follows.

3.2 Feature Extraction

Data Preprocessing. We leverage DEAP emotion database and its processed EEG signals data of 32 subjects to recognize emotion in the two dimensions of valence and arousal. Each subject participated in 40 trials, and 63 s signals data were collected in each trail. To expand the number of samples per subject without breaking the time continuity, we segment each 63 s trial into 30 samples (4 s long) by sliding window. The size and the step of the window are set 4 s and 2 s respectively. Finally, we get a total of 1200 samples (40 trials $\times$ 30 segments) for each subject. The labels of 30 samples extended from one trail are the same.

Fast Fourier Transform. After data preprocessing, we obtain 1200 samples for each subject, and each sample contains 32 channels of 4-s EEG signals. For each channel of a sample, we use FFT [20] to transform EEG data from time domain to frequency domain, and then use band-pass filter to decompose it into five frequency bands closely related to people's psychological activities, namely theta (4–8 Hz), alpha (8–12 Hz), low beta (12–16 Hz), high beta (16–25 Hz) and gamma (25–45 Hz). Since the collected EEG

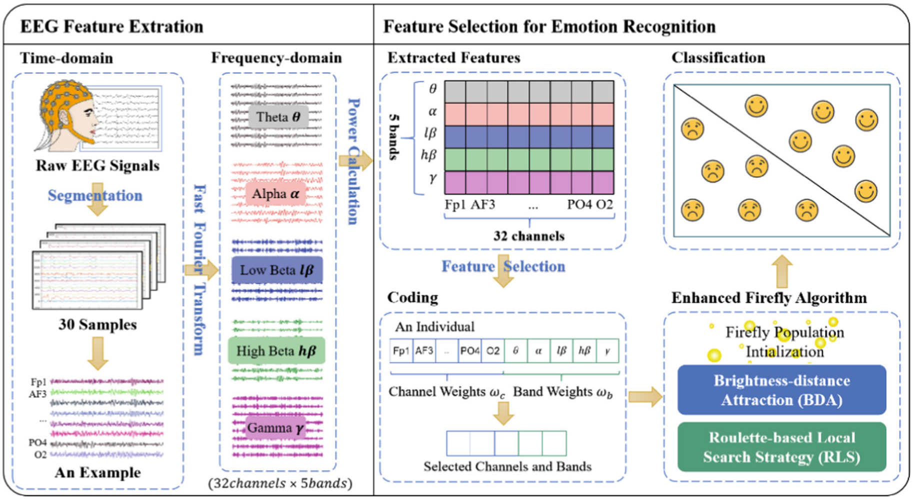

Fig. 1. General framework of proposed method

signal is a discrete sequence $s(n)$, , discrete Fourier transfer (DFT) is often applied to the transformation of EEG data as follow:

$$S(k) = DFT[s(n)] = \sum_{n=0}^{N-1} s(n) W_N^{nk} = \sum_{n=0}^{N-1} s(n) e^{-j\left(\frac{2\pi}{N}\right)nk} (k = 0, 1, \ldots, N-1) \tag{5}$$

where N represents the number of sample points and $W_N = e^{-j\left(\frac{2\pi}{N}\right)}$ is a transform matrix. Due to the high computational complexity of DFT, FFT improves its efficiency by replacing the computation of one larger DFT with the computation of several smaller DFTs.

Band Power. Band Power, a common feature extracted from EEG signals is used to recognize the emotion [5]. The power of a specific frequency band corresponding to channel T in sample i is calculated as

$$p_{iT} = \frac{1}{N} \sum_{k=1}^{N} \left| X_N^{iT}(k) \right|^2 \tag{6}$$

where $X_N^{iT}(k)$ is the FFT of the EEG signals for channel T in sample i, N is the length of FFT and equals the sample length 512 points (4 s).

A 512-point fast Fourier transform (FFT) is used to compute the power of each frequency band and 160 (32 channels × 5 bands) features are obtained for each sample. In order to eliminate the influence of scale differences between features and treat each feature equally, Z-score normalization is applied to each feature. For the feature f_i belonging to sample i, the Z-score normalized value was computed as

$$f_i^{norm} = \frac{f_i - \mu_f}{\sigma_f} \tag{7}$$

where μ_f and σ_f are the mean and the standard deviation of the feature f across all samples respectively.

3.3 Enhanced Firefly Algorithm

FA has the advantages of less parameters, simple operation and easy implementation. However, conventional FA updates its location by moving towards the brighter one and their attractiveness just depends on the light intensity which decrease with the distance as presented in Eq. (1) and (3). In brief, the attraction between two fireflies only depends on their distance. However, we think that the attraction between fireflies depends not only on their distance, but also on their brightness difference. On the other hand, FA has the limitations of slow convergence speed, early maturity and low accuracy in the late iterations. Therefore, we propose two strategies to tackle these challenges, namely brightness-distance based attraction (BDA) and roulette-based local search strategy (RLS). The flow chart of the algorithm is shown in Fig. 2.

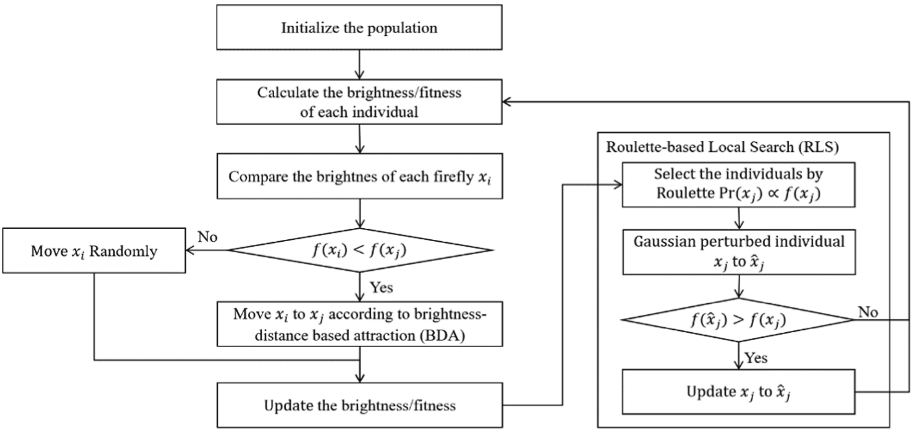

Fig. 2. Flow chart of enhanced firefly algorithm for maximum problem

Brightness-Distance Based Attraction (BDA). For conventional FA, the firefly will move toward the brighter one and their attraction depends on the light intensity which decrease with the distance as presented in Eq. (1) and (3). We think that the attraction between fireflies depends not only on their distance, but also on their brightness difference. If firefly j is much brighter than firefly i, their attraction will be stronger.

To depict this, we propose a relative brightness influence factor c to represent the brightness difference of mutual attraction between fireflies.

$$c = \frac{f(x_j) - f(x_i)}{max[f(x)] - min[f(x)]} \tag{8}$$

where $f(x_i)$ is the brightness value of the firefly i, $max\left[f(x)\right]$ and $min[f(x)]$ are the maximum and minimum brightness values of the current population respectively. Therefore, the movement of fireflies as Eq. (3) can be modified as:

$$x_i(t+1) = x_i(t) + c \cdot \beta\left(x_j(t) - x_i(t)\right) + \alpha r \tag{9}$$

Roulette-Based Local Search Strategy (RLS). To tackle the limitations of slow convergence speed, early maturity and low accuracy of firefly algorithm, we propose Roulette-based Local Search Strategy (RLS). Firstly, we apply Roulette algorithm to select the individuals with higher fitness values. The probability of an individual being selected is proportional to its fitness $\Pr(x_j) \propto f(x_j)$. The selection strategy of roulette algorithm is shown as follow:

$$P\left(x_j\right) = \frac{f\left(x_j\right)}{\sum_{j=1}^{n} f(x_j)} (j = 1, 2 \ldots n) \tag{10}$$

where $P\left(x_j\right)$ is the probability of individual x_j selection, $f(x_j)$ is the fitness value of individual x_j and n is the total number of fireflies.

After selecting a subset of fireflies that perform well, we take Gaussian perturbations on these individuals and move them around themselves. The position of the movement can be expressed as:

$$\hat{x}_j = x_j + \alpha r \tag{11}$$

where x_j is the current position of firefly j, $\hat{x}_j$ is the position after the update, r is an n-dimensional Gaussian random vector and α is a parameter that determines the random search. Finally, we choose the position with the best performance as follow:

$$new\_x_j = best(\hat{x}_j, x_j) \tag{12}$$

3.4 Enhanced Firefly Algorithm for Feature Selection

Encoding of Fireflies. As presented in Sect. 3.2, each sample contains 32 channels, including Fp1, AF3,…, PO4 and O2. Each channel can be transformed from time domain to frequency domain by FFT, and then decomposed into five frequency bands. After that, we calculate the band power of each frequency band. Therefore, we can obtain 160 features (32 channels × 5 bands) for each sample. We adopt enhanced FA to select some important features from these 160 features. Due to this specific method of feature extraction, each firefly can be decoded into 37 dimensions, of which 32 dimensions are channel importance weights $\omega_c \in R^{1\times 32}$, and 5 dimensions are band importance weights $\omega_b \in R^{1\times 5}$.

The encoding of a firefly x is presented in Fig. 3. We can obtain importance weights W for 160 features as follow:

$$W = {\omega_b}^T \cdot \omega_c \tag{13}$$

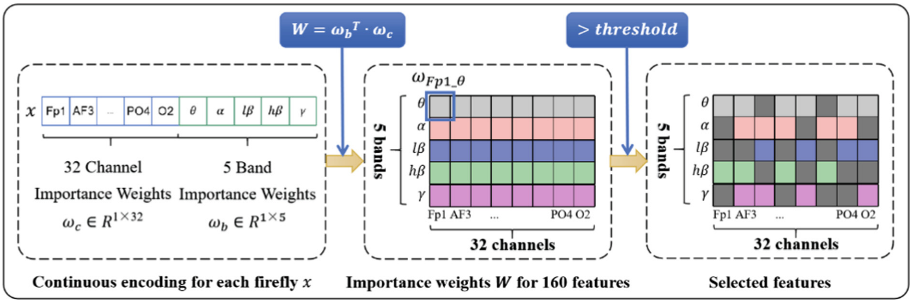

Fig. 3. Encoding scheme for a firefly

where ω_{Fp1_θ} represent the importance weight for band power that extracted from frequency bands θ of channel Fp1.

After that, we select features whose important weights are higher than the threshold, and the threshold is set to the average of all feature weights. The selection coefficient of feature f is calculated as:

$$C_f = \begin{cases} 0, w_f < \overline{w} \\ 1, w_f \geq \overline{w} \end{cases} \tag{14}$$

where w_f is the weight of the feature f and $\overline{w}$ is the average of all feature weights. The strategy for choosing the features as follow:

$$feature\_list = \{f | C_f = 1\} \tag{15}$$

Fitness Function. Fitness function is used to evaluate the performance of classification practice after selecting the most important features. We select K-Nearest Neighbor algorithm [21] as the classifier because of its advantages of simplicity and high precision. We adopt widely used area under ROC curve (AUC) as model performance measurement metrics. Therefore, the fitness is as follows:

$$fitness = AUC = \frac{1}{m^+ m^-} \sum\nolimits_{x^+ \epsilon D^+} \sum\nolimits_{x^- \epsilon D^-} I(x^+, x^-) \tag{16}$$

$$I\left(x^+, x^-\right) = \begin{cases} 1, & P_{x^+} > P_{x^-} \\ 0.5, & P_{x^+} = P_{x^-} \\ 0, & P_{x^+} < P_{x^-} \end{cases} \tag{17}$$

where m^+ and m^- are the number of positive samples and negative samples respectively, D^+ and D^- represent the set of positive samples and negative samples respectively and P_x is the prediction of sample x.

4 Experiments

4.1 Experimental Settings

Datasets. We conducted experiments on DEAP, a database for emotion analysis using physiological signals, which was published by Koelstra [15]. DEAP is based on the three-dimensional emotion model of Valence-Arousal-Dominance (VAD), and its EEG data can be used for emotion recognition research. In the process of data collection, a total of 32 subjects were selected for the experiment, including 16 males and 16 females, ranging in age from 19 to 37. In the data collection experiment, the physiological signals of the subjects were collected through 40 channels, among which the top 32 were EEG channels and the last eight channels were peripheral physiological signals. As this paper is about the correlation analysis of EEG data, our analysis is mainly based on the EEG data collected from the first 32 EEG channels.

Evaluation Protocols. We conducted subject-dependent experiments on DEAP database. Since each subject participated on 40 trials that was not enough for our experiments, we segmented each trial into 30 samples with the sliding window (size of 4 s and step of 2 s). Therefore, we totally obtained 1200 samples (40 trials $\times$ 30 segments) for each subject. The labels of 30 samples segmented from one trail are the same. For each subject, 960 samples were taken as training set and 240 samples as testing set. We evaluated our proposed model on 2 dimensional emotions, namely Valence and Arousal, and the threshold to divide samples into two classes was set to 5. We applied AUC as a criterion to evaluate the accuracy of the algorithm in emotion recognition.

Comparison Algorithms. In order to evaluate the performance of enhanced FA (EFA), we selected three heuristic algorithms as our comparison algorithms, namely PSO, GA and FA. We selected features by these four SI-based algorithms and then adopted KNN algorithm for emotion classification. Additionally, ReliefF-PNN [16], an algorithm that selected channels by ReliefF algorithm and classified the emotion by probabilistic neural network (PNN), was selected as a comparison algorithm to verify the effectiveness of SI-based feature selection in classification tasks.

Parameter Settings. For four SI-based algorithms, the candidate solutions were initialized between 0 and 1, and the lower and upper boundaries were set to 0 and 1. The population size of SI-based algorithms was set to 50 and the maximum number of iterations was set to 100. We searched the existing literature and found the most commonly used and recommended parameter settings for SI algorithms. EFA and FA shared the same parameters. For ReliefF-PNN, we traversed the sigma values from 0.1 to 0.9 with a step size of 0.1, and finally selected 0.1 with the highest accuracy.

4.2 Experimental Results

Classification Results and Number of Selected Features. We conducted 10 independent runs for all methods, and calculate average AUC as overall performance. Figure 4 displays the overall performance of our proposed EFA and the other four algorithms. It can be seen that four SI-based algorithms perform better than ReliefF-PNN in most subjects. It confirms the effectiveness of SI-based feature selection for emotion recognition. Among the four SI-based algorithms, EFA achieves the highest accuracy on both Valence and Arousal, evident from the positive effect of our proposed BDA and LRS strategies of EFA. Figure 5 depicts the number of features selected by five algorithms for each subject. As shown in the figure, the feature numbers of GA and ReliefF-PNN for all subjects are the least, followed by EFA and FA algorithm, and the number of features of PSO is much more than that of other algorithms. Table 1 shows the average performance (AUC) of 10 runs for each subject in more details. It can be found that EFA achieves the best accuracy in almost all subjects.

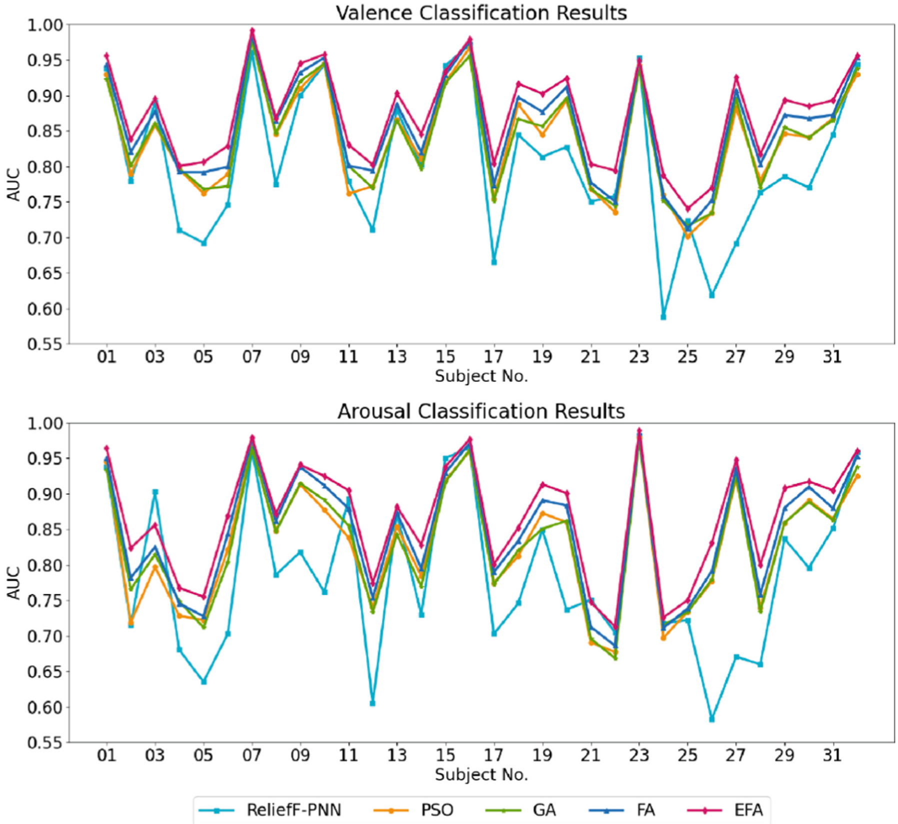

Fig. 4. Testing performance (AUC) of comparison algorithms in valence and arousal

Optimization Results. Figure 6 shows the fitness optimization for subject 01 and 32 with iteration process. Compared with FA, EFA is featured by the faster convergence speed in the first 10 iterations and converge to higher fitness. We believe that the improved performance of EFA benefits from the BDA and RLS strategies. For BDA strategy, the attraction of fireflies depends on brightness and distance, which help fireflies move faster to the brighter position. For RLS strategy, it provides opportunity for fireflies with higher fitness to achieve the better results. Furthermore, PSO has fast convergence while it obtains the lower AUC value, which may be caused by trapping into local optimum. The convergence of GA is slow and converge to lower fitness.

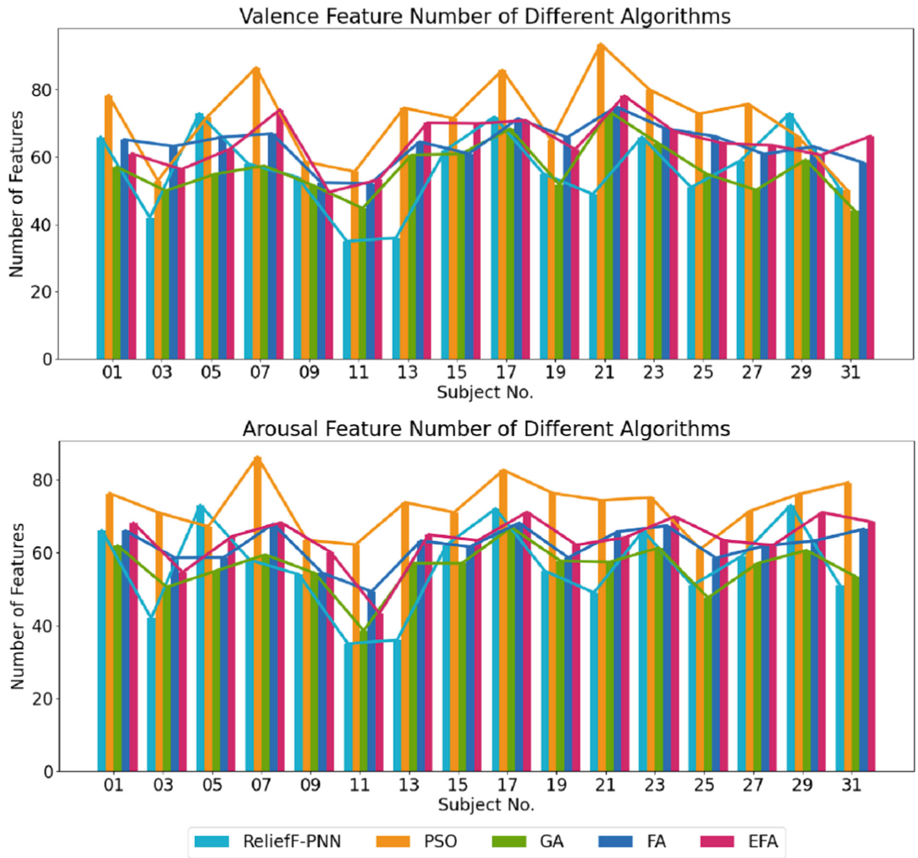

Fig. 5. The number of selected features in valence and arousal

Table 1. Average AUC performance (%) of 10 runs for each subject

Sub	Valence					Arousal				
	ReliefF	PSO	GA	FA	EFA	ReliefF	PSO	GA	FA	EFA
01	93.83	92.97	92.31	94.38	**95.61**	93.82	94.51	93.38	95.05	**96.43**
02	77.88	78.96	80.15	82.05	**83.85**	71.54	71.89	76.57	78.18	**82.38**
03	88.99	85.86	86.11	87.74	**89.55**	**90.31**	79.70	81.46	82.52	85.61
04	70.99	79.57	79.71	79.20	**80.06**	68.03	72.83	74.86	74.56	**76.75**
05	69.25	76.20	76.82	79.19	**80.60**	63.51	72.24	71.17	72.75	**75.51**
06	74.59	78.88	77.21	79.95	**82.87**	70.26	82.14	80.39	84.34	**86.94**

(*continued*)

Table 1. (*continued*)

Sub	Valence					Arousal				
	ReliefF	PSO	GA	FA	EFA	ReliefF	PSO	GA	FA	EFA
07	96.14	97.81	97.88	98.63	**99.17**	96.15	96.59	96.47	97.72	**97.97**
08	77.49	84.51	84.67	86.44	**86.78**	78.59	84.82	84.66	86.21	**87.21**
09	90.01	91.00	92.00	93.25	**94.50**	81.79	91.31	91.49	93.81	**94.09**
10	94.37	94.50	94.51	95.33	**95.79**	76.17	87.78	89.14	91.18	**92.51**
11	77.92	76.16	80.09	80.11	**83.04**	89.31	83.83	85.54	87.93	**90.49**
12	71.03	77.21	76.94	79.42	**80.24**	60.48	74.03	73.36	75.35	**77.39**
13	87.98	86.62	86.47	88.75	**90.30**	87.19	85.36	84.25	87.36	**88.20**
14	80.22	81.03	79.65	81.99	**84.55**	73.02	78.44	76.94	79.47	**82.81**
15	**94.21**	91.89	91.77	92.87	93.42	**95.03**	91.90	91.66	93.00	93.81
16	97.16	96.73	95.59	97.69	**97.92**	96.48	96.00	96.19	97.13	**97.70**
17	66.52	75.41	75.10	77.37	**80.36**	70.26	77.45	77.23	78.98	**80.08**
18	84.47	88.73	86.70	89.75	**91.66**	74.57	81.23	82.02	83.36	**85.22**
19	81.34	84.45	85.66	87.70	**90.27**	84.96	87.27	85.06	89.09	**91.34**
20	82.75	89.31	89.64	91.21	**92.42**	73.70	86.02	86.21	88.37	**90.11**
21	75.03	76.86	76.82	77.75	**80.32**	**75.07**	69.08	69.61	71.25	74.73
22	75.81	73.55	74.43	75.05	**79.44**	70.46	67.72	66.84	68.62	**71.31**
23	**95.30**	94.25	93.82	95.08	94.99	98.29	97.90	97.72	98.65	**98.94**
24	58.84	76.02	75.15	75.84	**78.77**	71.80	69.72	71.54	71.12	**72.61**
25	72.39	70.13	71.63	71.25	**74.04**	72.23	73.40	73.36	73.86	**75.03**
26	61.83	73.42	73.42	75.23	**76.99**	58.26	77.66	77.95	79.14	**83.10**
27	69.16	88.23	89.57	90.74	**92.58**	67.06	92.26	92.64	93.61	**94.79**
28	76.31	78.07	77.03	80.32	**81.69**	65.98	73.98	73.43	75.84	**80.00**
29	78.57	84.63	85.49	87.26	**89.37**	83.67	85.88	85.97	88.03	**90.81**
30	77.02	84.08	84.07	86.78	**88.49**	79.59	89.07	88.85	91.01	**91.72**
31	84.43	86.77	86.55	87.27	**89.29**	85.07	86.52	86.34	87.97	**90.51**
32	94.41	92.97	93.80	95.35	**95.63**	95.84	92.50	93.83	95.31	**96.06**
Ave	80.51	83.96	84.09	85.65	**87.33**	78.70	82.84	83.00	84.71	**86.63**

Channel Importance and Band Importance. We obtained the optimal solution for each subject through 4 SI-based algorithms. According to our encoding method, the optimal solution contains 37 dimensions, among which the top 32 are the importance weights for channels and the last 5 dimensions are the importance weights for bands. Table 2 shows the average importance of five frequency bands of all subjects in Valence and Arousal classification tasks. It reveals that frequency bands with higher frequency

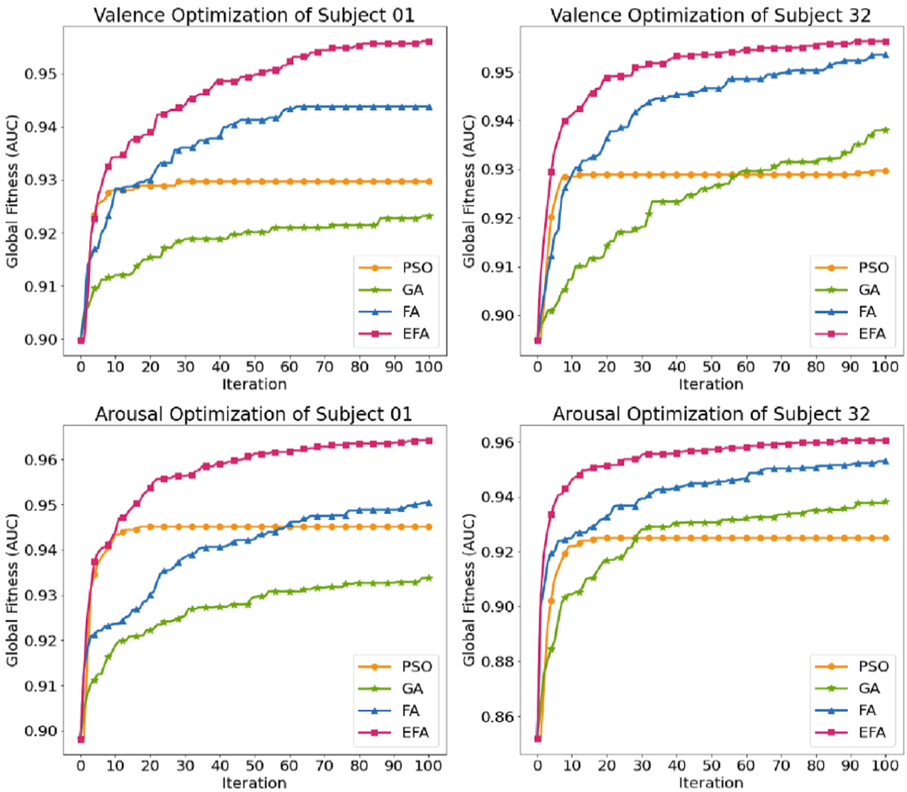

Fig. 6. Fitness optimization for subject 01 and 32 in valence and arousal

are more significant for predicting emotion classification. We calculated the average importance of channels of all subjects and found out the top 10 important channels as shown in Fig. 7. It shows that the top 10 channels of valence are mainly located in frontal and parietal brain regions related to emotion processing, while top 10 channels of arousal are mainly located in parietal and occipital brain regions.

Table 2. The importance of different bands in valence dimension

Band name	Valence				Arousal			
	PSO	GA	FA	EFA	PSO	GA	FA	EFA
Theta	0.4449	0.1421	0.1977	0.2010	0.4358	0.1477	0.1996	0.1840
Alpha	0.5563	0.2316	0.3192	0.3116	0.5732	0.2316	0.3000	0.3004
Low-beta	0.7590	0.4312	0.4859	0.4843	0.7510	0.4306	0.4861	0.4949
High-beta	0.9012	0.5233	0.6333	0.6216	0.7943	0.5383	0.5720	0.5556
Gamma	**0.9246**	**0.7125**	**0.7582**	**0.8305**	**0.8847**	**0.7518**	**0.7433**	**0.8058**

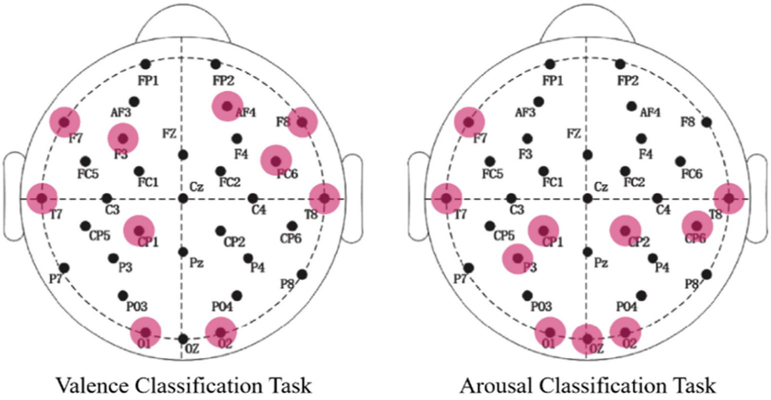

Fig. 7. Top 10 important channels in valence and arousal

5 Conclusion

In this work, we propose Enhanced Firefly Algorithm (EFA) with brightness-distance based attraction (BDA) and roulette-based local search strategy (RLS), which provide faster convergence and higher accuracy. We apply EFA to EEG-based emotion recognition and provide a novel encoding method of fireflies, which can distinguish the importance of channels and bands respectively. We conducted subject-dependent experiments on DEAP database, and the experimental results show that the EFA achieves the highest accuracy among the competitive feature selection methods in emotion classification of arousal and valence. Our proposed algorithm requires further improvement to reduce the number of selected features. In future studies this problem will be handled by considering multi-objective optimization for feature selection.

Acknowledgement. This study is supported by National Natural Science Foundation of China (71901150, 71901143), Natural Science Foundation of Guangdong (2022A1515012077), Guangdong Province Innovation Team "Intelligent Management and Interdisciplinary Innovation" (2021WCXTD002), Shenzhen Higher Education Support Plan (20200826144104001).

References

1. Ko, B.C.: A Brief review of facial emotion recognition based on visual information. Sensors. **18**(2), 401 (2018)
2. Ingale, A.B., Chaudhari, D.S.: Speech emotion recognition. Int. J. Soft Comput. Eng. (IJSCE). **2**(1), 235–238 (2012)
3. Bos, D.O.: EEG-based emotion recognition. The Influence of Visual and Auditory Stimuli **56**(3), 1–17 (2006)
4. Wei, C.Z.: Stress emotion recognition based on RSP and EMG signals. In: Advanced Materials Research, vol. 709, pp. 827–831. Trans Tech Publications Ltd. (2013)

5. Jenke, R., Peer, A., Buss, M.: Feature extraction and selection for emotion recognition from EEG. IEEE Trans. Affect. Comput. **5**(3), 327–339 (2014)
6. Nie, D., Wang, X.W., Shi, L.C., Lu, B.L.: EEG-based emotion recognition during watching movies. In: 2011 5th International IEEE/EMBS Conference on Neural Engineering, pp. 667–670 (2011)
7. Li, D., Wang, Z., Wang, C., et al.: The fusion of electroencephalography and facial expression for continuous emotion recognition. IEEE Access. **7**, 155724–155736 (2019)
8. Piho, L., Tjahjadi, T.: A mutual information based adaptive windowing of informative EEG for emotion recognition. IEEE Trans. Affect. Comput. **11**(4), 722–735 (2018)
9. Zhang, Q., Lee, M.: A hierarchical positive and negative emotion understanding system based on integrated analysis of visual and brain signals. Neurocomputing **73**(16–18), 3264–3272 (2010)
10. Rahman, M.A., Hossain, M.F., Hossain, M.F.: Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal. Egypt. Inform. J. **21**(1), 23–35 (2020)
11. Li, Z., Qiu, L., Li, R.: Enhancing BCI-based emotion recognition using an improved particle swarm optimization for feature selection. Sensors **20**(11), 3028 (2020)
12. Nakisa, B., Rastgoo, M.N., Tjondronegoro, D.: Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. Expert Syst. Appl. **93**, 143–155 (2018)
13. Yang, X.S.: Nature-inspired Metaheuristic Algorithms. Luniver Press (2010)
14. Yang, X.-S., Deb, S., Zhao, Y.-X., Fong, S., He, X.: Swarm intelligence: past, present and future. Soft. Comput. **22**(18), 5923–5933 (2017). https://doi.org/10.1007/s00500-017-2810-5
15. Koelstra, S., Muhl, C., Soleymani, M.: Deap: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2011)
16. Zhang, J., Chen, M., Hu, S.: PNN for EEG-based emotion recognition. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 002319–002323. IEEE (2016)
17. He, H., Tan, Y., Ying, J., Zhang, W.: Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm. Appl. Soft Comput. **94**, 106426 (2016)
18. Yildirim, E., Kaya, Y., Kiliç, F.: A channel selection method for emotion recognition for EEG based on swarm-intelligence algorithms. IEEE Access **9**, 109889–109902 (2021)
19. Sreeshakthy, M., Preethi, J.: Classification of human emotion from deap EEG signal using hybrid improved neural networks with cuckoo search. BRAIN Broad Res. Artif. Intell. Neurosci. **6**(3–4), 60–73 (2016)
20. Nussbaumer, H.J.: The fast Fourier transform. In: Fast Fourier Transform and Convolution Algorithms. Springer, Berlin, Heidelberg (1981). https://doi.org/10.1007/978-3-662-00551-4_4
21. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **13**(1), 21–27 (1967)