

Research Article

Emotion Recognition Based on Framework of BADEBA-SVM

Zhongmin Wang,^{1,2} Zhaoping Zhang ,² and Wenlang Wang^{1,2}

¹School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an Shaanxi 710121, China

²Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University of Posts and Telecommunications, Xi'an Shaanxi 710121, China

Correspondence should be addressed to Zhaoping Zhang; 261430050@qq.com

Received 28 February 2019; Revised 28 May 2019; Accepted 18 June 2019; Published 2 July 2019

Academic Editor: Emilio Turco

Copyright © 2019 Zhongmin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brain-computer interface (BCI) provides a new communication channel between human brain and computer. In order to eliminate uncorrelated channels to improve BCI performance and enhance user convenience with fewer channels, this paper proposes a new framework using binary adaptive differential evolution bat algorithm (BADEBA). The framework uses the important ideas of differential evolution algorithm and bat algorithm to select electroencephalograph (EEG) channels and intelligently optimizes the parameters of support vector machine (SVM). It combines wavelet packet transform (WPT) and common space pattern (CSP) to achieve the goal of using fewer channels to obtain the best classification accuracy. The proposed framework is evaluated with a common data set (DEAP). The results show that, compared with genetic algorithm (GA), binary particle swarm optimization (BPSO) and bat algorithm, the proposed BADEBA in this framework only uses eight channels to improve the classification accuracy by 13.63% in the valence dimension and seven channels to improve the classification accuracy by 15.22% in the arousal dimension. In addition, the spatial distribution of the best channels selected by this method is consistent with the existing knowledge of brain structure and neurophysiology, which shows the accuracy and validity of this method.

1. Introduction

Emotion recognition as an emerging research direction has attracted increasing attention from different fields and is promising for various application domains, especially in the fields of medicine [1], education [2], and road and aviation safety [3]. The literature reveals that researchers have developed a variety of emotion recognition methods based on facial expressions [4–7] and voice [8, 9]. Compared with facial expressions and voice, EEG-based emotion recognition shows greater potential, because internal nerve fluctuations cannot be deliberately hidden or controlled. Other peripheral-physiological signals, such as skin temperature, heart rate, respiratory patterns, blood volume pressure, and galvanic skin response [10, 11], have low resolutions due to the problems of the signal itself. Because of its high classification and recognition accuracy, EEG signals have been introduced into the field of emotion recognition, especially the application of BCI technology, due to its simple equipment, no harm

to the experimenters, and high temporal resolution of the collected brain signals.

BCI is a communication system that transmits information between brain and external devices. One of the main challenges of BCI is its personal dependence. Even if the same experiment is replicated in the same environment, different brain regions will be activated. To solve this problem, more EEG channels need to be added to obtain more decision signals. However, the use of multiple channels can cause additional problems, such as computational complexity, noise, and redundant signals, which will reduce the performance of BCI. In addition, the use of a large number of channels requires a longer preparation time, which directly affects the convenience of BCI. Therefore, the need for performance and convenience can be balanced by choosing the minimum number of channels to obtain the highest or required accuracy. Based on DEAP database, Jianhai Z et al. [12] used Relief method to obtain the most closely related EEG channels, which greatly reduced the number of EEG channels

needed to be collected at the expense of lower recognition rate. Arvaneh M et al. [13] proposed a new EEG channel selection algorithm based on sparse CSP, which can delete those channels that have poor correlations with EEG or even negative effects, and then obtain higher recognition rate while reducing the number of EEG channels. Similar enhancements based on CSP are also used for feature extraction and channel selection [14]. He L et al. [15] used an improved genetic algorithm based on Rayleigh coefficient characteristics to select the optimal subset of EEG channels. Their experiments showed the effectiveness of the method. Lan T et al. [16] proposed a method based on mutual information, which is utilized to rank the importance of channels, and important channels have the priority to be chosen. However, such kind of methods only focuses on the correlation between features and categories and ignores the relationship between features. Ansari-Asl K et al. [17] proposed a channel selection algorithm based on Synchronization Likelihood method in 2007. Five effective channels were selected from 64 EEG channels. When it was used to recognize positive, middle, and negative emotional states, the classification effect was not significantly reduced. Lahiri R et al. [18] used the firefly algorithm to optimize the channel selection for the three objectives of correlation and mutual information. The experimental results showed that the method was applicable. Then they added the classification accuracy as the fourth objective and independently optimized the four objectives. The channel selection was performed through membership ranking. The result was in the optimal solution set, and the channels selected on the basis of the optimization of the classification accuracy were in the majority [19].

Based on the aforementioned reasons, this paper proposes a BADEBA-SVM framework for channel selection, which maximizes classification accuracy and minimizes the number of channels. The framework applies mutation, crossover, and selection mechanisms of the binary adaptive differential evolution algorithm to the bat algorithm, so that mutation mechanism is introduced into the bat algorithm. This can increase the diversity of the bat population and enhance the global optimization ability of the framework by preventing individual population falling into local optimum. Combining with WPT and CSP, the best classification accuracy can be achieved by using fewer channels.

2. Related Work

2.1. Bat Algorithm. Bat algorithm is a new heuristic swarm intelligence algorithm [20] proposed by Xinshe Yang in 2010. The algorithm searches the solution space of the optimization problem by simulating the bat's foraging behaviour to achieve the goal of optimization. Compared with intelligent algorithms such as genetic algorithm and particle swarm optimization, it has two advantages. One is to enhance the diversity of the population in the optimization process using the frequency tuning method, and the other is to improve the convergence of the algorithm by adjusting the pulse loudness and frequency adaptively.

In bat algorithm, a bat represents a feasible solution, while the prey of a bat represents the optimal solution. The position

of Bat i at t time is x_i^t and the velocity is v_i^t , and the updating formula of Bat i at $t+1$ time in d -dimensional foraging space is as follows:

$$f_i = f_{\min} + (f_{\max} - f_{\min}) \beta \quad (1)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x_*) f_i \quad (2)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (3)$$

where β is the random number within $[0, 1]$, f_i is the frequency of sound waves emitted by Bat i at the current time, and $f_i \in [f_{\min}, f_{\max}]$. x_* is the current global optimal solution.

When searching locally, a bat is randomly selected from the bat population, and according to (4), a new authority solution is created by free flight of the bat.

$$x_{\text{new}} = x_{\text{old}} + \varepsilon A^t \quad (4)$$

where A^t is the average loudness of the bat population at t time, $\varepsilon \in [-1, 1]$ is a random vector, and x_{old} is a random solution selected from the current optimal solution. During the iteration of bat algorithm, the following equations are utilized to update loudness A_i and emission pulse rate r_i .

$$A_i^{t+1} = \alpha A_i^t \quad (5)$$

$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)] \quad (6)$$

where α and γ are constants. In most applications, we usually set $0 < \alpha < 1$, $\gamma > 0$, and r_i^0 is the initial emission rate.

2.2. Differential Evolution Algorithm. Differential evolution algorithm is an evolutionary algorithm based on the population difference. It was proposed by Rainer Storn and Kenneth Price in 1997 to solve Chebyshev polynomials [24]. The mutation operation is the most prominent feature of the algorithm. The mutation is accomplished by adding the weighted difference between two individuals to the individual. Like GA, the algorithm includes mutation, crossover, and selection operations. The mutation operation is as follows:

$$v_{ij}(t) = x_{b1}(t) + F(x_{b2}(t) - x_{b3}(t)) \quad (7)$$

where $x_{b1}(t)$, $x_{b2}(t)$, and $x_{b3}(t)$ are three individuals randomly selected from the population, and $b1 \neq b2 \neq b3 \neq i$, $b1, b2, b3 \in \{1, 2, \dots, NP\}$. F is a scaling factor and one of the main control parameters of differential evolution algorithm, $F \in [0, 1]$.

Cross-operation can increase the diversity of the population. Interindividual cross-manipulation of the t -generation population and its variant intermediates is performed based on (8).

$$u_{ij}(t+1) = \begin{cases} v_{ij}(t), & \text{if } (\text{rand}(0, 1) \leq CR \\ & \text{or } j = \text{rand}(1, n) \\ x_{ij}(t), & \text{if } (\text{rand}(0, 1) > CR \\ & \text{or } j \neq \text{rand}(1, n) \end{cases} \quad (8)$$

where $\text{Rand}(0,1)$ is a random number obeying uniform distribution on $[0, 1]$, $\text{rand}(1, n) \in [1, 2, \dots, d]$ is a random

integer, $j=[1, 2, \dots, d]$. This crossover strategy ensures that at least one component of $u_{ij}(t+1)$ is contributed by $v_i(t)$.

Selection operation is to determine whether individual $u_{ij}(t+1)$ is superior to $x_i(t)$ based on fitness function to generate a new generation of individual $x_i(t+1)$. The equation is as follows:

$$x_i(t+1) = \begin{cases} u_{ij}(t+1), & \text{if } (f(u_{ij}(t+1)) < f(x_i(t))) \\ x_i(t), & \text{others} \end{cases} \quad (9)$$

3. Proposed Methods

This section provides a detailed overview of the proposed framework (Figure 1). The proposed framework consists of datasets (a training dataset and a validation dataset), signal preprocessing, channel selection, feature extraction, optimization of support vector machines, and classification.

In this framework, a novel evolutionary method is used to evaluate the fitness function value by validation set and optimize the channel subset and support vector machine parameters to obtain the highest classification accuracy and the minimum number of channels. The first step is to decompose the EEG signal into wavelet packet in frequency domain and then reconstruct the signal in the $(\beta+\gamma)$ band. The second step is to extract the channel feature of position 1 through CSP. Here we regard the electrode space as solution space. Each bat's position vector consists of 0 and 1. 0 means that the channel is not selected, and 1 means that the channel is selected. The third step is to optimize the parameters of SVM. The parameters needed to be determined for SVM model are penalty coefficient C and kernel parameter σ . We selected the radial basis function as the kernel function of SVM, which is expressed as follows:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (10)$$

where σ is kernel parameter.

It is found that the choice of C and σ have great effects on the performance of SVM. The value of penalty coefficient C will affect the complexity of model and the penalty degree of the training error. Thus it will affect the efficiency of the performance. Learning is overfitting if the value of C is too large, whereas extra learning is needed if the value is too small. The kernel parameter σ precisely defines the structure of high-dimensional feature space, thus controlling the complexity of the ultimate solution. A larger value of σ (it will end up with a hypersurface that is almost flat) causes underfitting, whereas a smaller value of σ (it will end up with a spiky hypersurface) causes overfitting. The fourth step is to evaluate the fitness function. We set the fitness function according to the importance of the two factors: average classification accuracy and the number of channels used. Since the maximum classification accuracy is equal to

the minimum classification error, the fitness function is set as follows:

$$\text{Fitness} = 0.8 \times (1 - \text{ACC}) + 0.2 \times \frac{\sum_{j=1}^n f_j}{n} \quad (11)$$

$$\text{ACC} = \frac{\# \text{ of correctly classified data}}{\# \text{ of validation data}} \quad (12)$$

where n is the number of channels, $f_j=1$ indicates that this channel is selected, $f_j=0$ indicates that this channel is not selected, and ACC is the average classification accuracy of SVM for validation set.

3.1. Binary Adaptive Differential Evolution Algorithm. In order to solve the optimization problem of discrete space such as channel selection, the differential evolution algorithm must be binary coded and the new mutation operator must be designed by using logic operation instead of arithmetic operation. In this paper, the following operations are defined first.

$$A | F \cdot B = \begin{cases} A | B, & \text{rand}() \leq F \\ A, & \text{rand}() > F \end{cases} \quad (13)$$

Random selection of three individuals is conducive to global search, but reduces the convergence of the algorithm. In order to improve the convergence performance of the algorithm, the vector difference between the optimal individual and the suboptimal individual of the three individuals randomly selected is added. The vector difference of two randomly selected individuals is added to the optimal individual of the current algebra, which has a good performance on convergence but is easy to fall into local optimum. In order to make the algorithm possess both global search ability and convergence performance, (14) is modified as follows: using “|”, “ \oplus ”, or exclusive or operation, respectively.

$$v_{ij}(t) = \begin{cases} X_b(t) | F \cdot (X_{i1}(t) \oplus X_{i2}(t)), & \text{rand}() \leq 1 - \left(\frac{t}{T}\right)^2 \\ X_{\text{best}}(t) | F \cdot (X_{b1}(t) \oplus X_{b2}(t)), & \text{others} \end{cases} \quad (14)$$

where $X_b(t)$ is the best of the three individuals randomly selected in the current generation. $X_{i1}(t)$ and $X_{i2}(t)$ are the other two suboptimal individuals. $X_{\text{best}}(t)$ is the best individual in current generation, and $X_{b1}(t)$ and $X_{b2}(t)$ are two randomly selected individuals. t is the number of iterations and T is the maximum number of iterations.

Differential evolution algorithm has three control parameters: population size N , difference vector scaling factor F , and crossover probability CR . F and CR have significant influences on the performance of the algorithm. F is set to be large in the early stage and small in the late stage, and CR is set to be small in the early stage and large in the late stage. This can improve the performance of the algorithm. The modified equation of the scaling factor F is as follows:

$$F = F_{\min} + (F_{\max} - F_{\min}) \left| \frac{T-t}{T} \right| \quad (15)$$

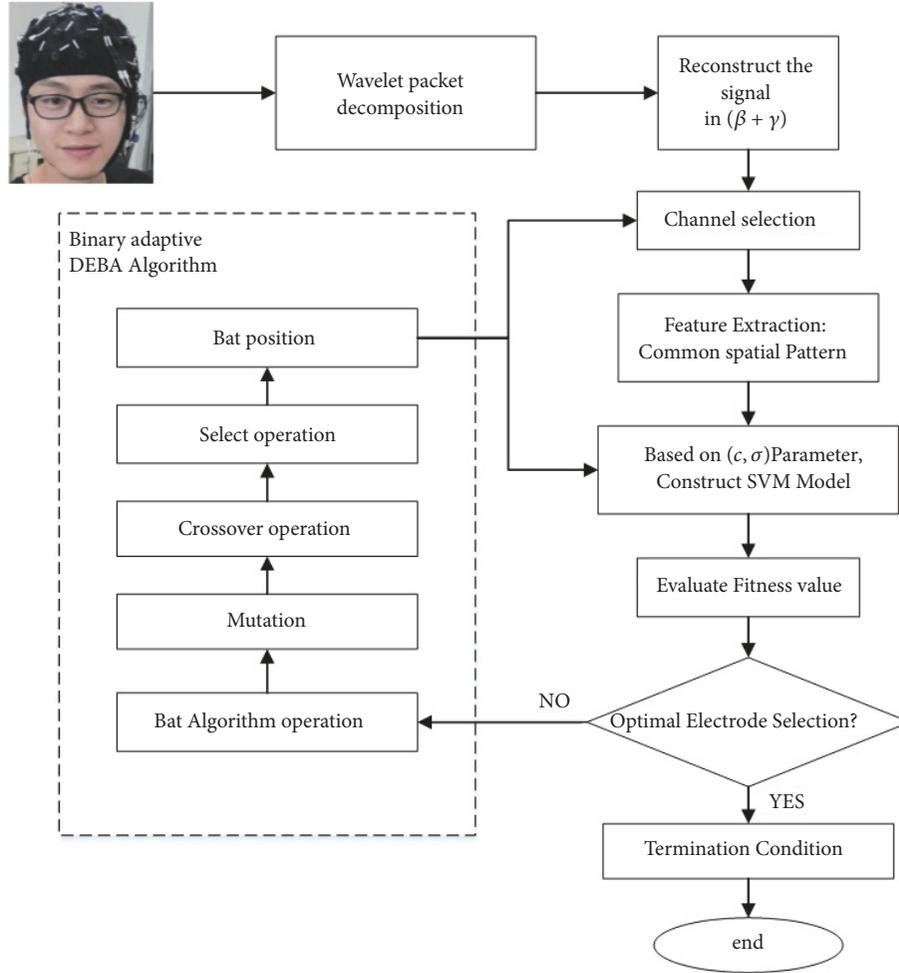


FIGURE 1: BADEBA-SVM framework for emotion recognition.

where F_{\max} is the upper limit of the scaling factor and F_{\min} is the lower limit of the scaling factor. Generally, $F_{\max}=0.9$ and $F_{\min}=0.1$.

Dynamic CR can make the algorithm converge to the appropriate position with a higher probability in the early stage of global search. In this study, the strategy of linearly decreasing weighted CR values is adopted, and the equation is as follows:

$$CR = CR_{\min} + \frac{t \times (CR_{\max} - CR_{\min})}{T} \quad (16)$$

where CR_{\min} is the minimum crossover probability and CR_{\max} is the maximum crossover probability.

3.2. Binary Adaptive Differential Evolution Bat Algorithm. This paper proposes a binary adaptive differential evolution bat algorithm, which introduces the mutation, crossover, and selection mechanism of binary adaptive differential evolution algorithm into bat algorithm. Compared with the bat algorithm, the difference between the two algorithms is that, in each evolutionary process, the evolved bat position $X_i(t)$ does not directly go to the next iteration, but carries

out mutation, crossover, and selection operations among individuals in the population to obtain a new bat position. Then it goes to $(t+1)$ iteration. The specific implementation steps are shown in Algorithm 1.

4. Experimental Data and Results

DEAP is a data set for emotion analysis using EEG, physiological and video signals [25]. In this section, only EEG signals (including 32 channels in 10-20 systems) are used for emotion recognition. The data set consisted of 32 subjects, each of whom provided 40 samples for self-evaluation of valence, arousal, dominance, and liking (ranging from 1 to 9 points). In this experiment, valence and arousal were chosen as two separate classifications. We selected the processed matlab version of the data set. Each sample lasted for 63S (3 seconds rest and 60 seconds EEG data) and contained 8064 data points, with a sampling frequency rate of 128 HZ. In order to improve the signal-to-noise ratio (SNR), the first 3 seconds of the data set were deleted, leaving only 60S useful EEG data. In data set splitting, 24×40 samples were randomly selected as the training set, and 8×40 samples were left as the

Require: Objective function $\text{Fitness}(x_{ij})$, $x_{ij} = (x_1, x_2, \dots, x_d)^T$

- (1) Initialize the bat population x_{ij} ($i = 1, 2, \dots, NP$), v_i , F_{\max} , F_{\min} , CR_{\min} and CR_{\max}
- (2) Define pulse frequency f_i at x_{ij}
- (3) Initialize pulse rate r_i and the loudness A_i
- (4) **while** $t < \text{Max number of iterations}$ do
- (5) Generate new solutions by adjusting frequency, and updating velocities and locations/solutions according to Eq.(1)-Eq.(3)
- (6) **if** $\text{rand} > r_i$ **then**
- (7) Select a solution among the best solutions
- (8) Generate a local solution around the selected best solution
- (9) **end if**
- (10) Generate a new solution by flying randomly according to Eq.(4)
- (11) **if** $\text{rand} < A_i$ & $f(x_{ij}) < f(x^*)$ **then**
- (12) Accept the new solutions
- (13) Increase r_i and reduce A_i according to Eq.(5) and Eq.(6)
- (14) **end if**
- (15) Mutate using Eq. (14)
- (16) Crossover operation using Eq.(8)
- (17) Select operation using Eq.(9), find the current best x^*
- (18) **end while**
- (19) Output Result

ALGORITHM 1: Binary adaptive differential evolution bat algorithm.

test set. In order to improve the generalizability of the model, the training set was divided into 10 parts, each containing 96 samples and used as the validation set of the model. When training the model, we adopted the method of 10-fold cross validation. In label making, we chose the value 5 as the cut-off point. Values greater than 5 were categorized into one type of emotion, and values less than or equal to 5 were categorized into the other type. The running environment of the system was matlab 2016b, win10, memory 12G, CPU model i7-7500U, 64bit operating system.

4.1. Signal Preprocessing and Feature Extraction. WPT includes the decomposition and reconstruction of wavelet packet coefficients. Wavelet packet decomposition has been extensively used in the field of signal processing. Compared with wavelet decomposition, it cannot only decompose the signal orthogonally in the whole frequency band, but also select the corresponding frequency band adaptively according to the characteristics of the signal, so that it matches the signal spectrum and has higher time-frequency resolution. Based on the multiresolution characteristic of WPT, the optimal component combination relation of EEG signal can be selected, and the signal in the useful information frequency range can be extracted and reconstructed. The decomposition algorithm for the coefficients is obtained by (17).

$$\begin{aligned} d_k^{j+1,2n} &= \sum_k h_{0(2l-k)} d_1^{j,n}, \\ d_k^{j+1,2n+1} &= \sum_k h_{1(2l-k)} d_1^{j,n}, \end{aligned} \quad (17)$$

where $d_1^{j,n}$, $d_k^{j+1,2n}$, and $d_k^{j+1,2n+1}$ are all the wavelet packet coefficients, and $h_{0(2l-k)}$ and $h_{1(2l-k)}$ are the low-pass and high-pass filter coefficients for decomposition, respectively.

The reconstruction algorithm for the wavelet packet coefficients is deduced as

$$d_1^{j,n} = \sum_k g_{0(l-2k)} d_k^{j+1,2n} + \sum_k g_{1(l-2k)} d_k^{j+1,2n+1} \quad (18)$$

where $g_{0(l-2k)}$ and $g_{1(l-2k)}$ are the low-pass and high-pass filter coefficients for reconstruction, respectively.

Research [26] shows that the high frequency components of EEG signals can reflect people's emotional and cognitive state, and β and γ bands are better than low frequency bands in distinguishing emotional state transition [27]. In order to improve the SNR of EEG signals, the emotion-related β (13-30Hz) and γ (30-50Hz) band signals are obtained. The wpcocf function Db4 is used to decompose 7-layer wavelet packets to obtain 128 nodes. Since the sampling frequency rate of EEG signals is 128HZ, the range of each subband is 1HZ. The signals ($\beta+\gamma$) bands are reserved and reconstructed to obtain the final EEG signals. Then, the ($\beta+\gamma$) band signals of the selected channel are constructed by the spatial filter based on CSP algorithm, which maximizes the energy difference of the spatial components of EEG data after spatial filtering. The variance of each line of the eigenvector matrix of two kinds of signals is extracted as EEG features.

4.2. Optimization Algorithms for Channel Selection. For swarm intelligence optimization algorithm, parameters determine the performance of the algorithm. In terms of parameter setting, we have carried out multivariate sensitivity tests. For example, the range of bat population size is set to from 10 to 100, incrementing by 5 each time; the range of pulse A_0 is set to from 0.05 to 0.9, incrementing by 0.05 each time. By changing the parameters at the same time to observe its impact on the classification results, the following parameters are finally determined to achieve better classification results.

In BCI system, BADEBA was used to select electrodes. Equation (11) was used as fitness function. Fifty bats were selected and used. We set the maximum number of iterations to 500, the maximum pulse A_0 to 0.25, the maximum pulse rate r_0 to 0.5, the pulse frequency range to $[0, 1]$, and the pulse frequency enhancement coefficient γ to 0.05, the pulse frequency attenuation coefficient α to 0.95, the upper limit of scale factor F_{\max} to 0.9, the lower limit of scale factor F_{\min} to 0.1, the maximum crossover probability CR_{\max} to 0.9, and the minimum of crossover probability CR_{\min} to 0.1.

In BCI system, GA was used to select electrodes. Equation (11) was also used as fitness function. The parameters were as follows: 50 individuals were used, and the maximum number of iterations was set to 500, the crossover probability to 0.8, and the mutation probability to 0.01.

In BCI system, the binary particle swarm optimization algorithm [28] was used to select the electrode. Equation (11) was also used as fitness function. The parameters were as follows: 50 particles were used and the maximum number of iterations was set to 500. We set c_1 and c_2 to 2, r_1 , r_2 and τ to be random numbers between 0 and 1. The t -th inertia weight is as follows:

$$w(t) = w_{\max} - (w_{\max} - w_{\min}) \frac{t}{T} \quad (19)$$

where w_{\max} was set to 0.9, w_{\min} was set to 0.5, and T represents the maximum number of iterations.

4.3. Experimental Results. Using binary adaptive differential evolution bat algorithm for channel selection, experiments were conducted on the two dimensions of valence and arousal. The results are shown in Figure 2: the left chart is the optimal channel in the valence dimension and the right chart is the optimal channel in the arousal dimension. As can be seen from Figure 2, in addition to some other channels, the prefrontal and occipital lobes are selected as the best channels. This suggests the effectiveness of the proposed framework. From neurophysiological knowledge, the brain regions of the prefrontal and occipital lobes are related to emotions.

Figure 2 shows that by minimizing the fitness function, the framework chooses eight channels in the valence dimension and seven channels in the arousal dimension. In order to verify the effectiveness of the framework, GA, BPSO, BA, and BADEBA are used for channel selection and comparison with the whole channel. The experimental results on the valence dimension are shown in Table 1.

As can be seen from Table 1, when validation sets are used for binary classification in the dimension of valence, 75.26% of the average classification accuracy is achieved when only 8 channels are used with BADEBA algorithm. When using GA, only 61.63% average classification accuracy is obtained by choosing 15 channels. As can be seen from Table 2, when validation sets are used for binary classification in the dimension of arousal, 75.98% of the average classification accuracy is achieved when only 7 channels are used with BADEBA algorithm. When using GA, only 60.76% average classification accuracy is obtained by choosing 14 channels.

The results in Tables 1 and 2 clearly show that the proposed BADEBA is superior to other evolutionary algorithms.

According to the channels selected by the four algorithms, in order to find out which band of signals contributes most to classification results, we reconstruct the original signals in the β band, γ band, and $(\beta+\gamma)$ band after decomposing the wavelet coefficients. We use the four optimized models to classify the emotions. The experimental results (Figure 3) show that the classification accuracy of signals in the $(\beta+\gamma)$ band is better than that of signals in the β band or in the γ band alone. This shows that for certain emotions, the relationship between the two bands is neglected by using β band signal or γ band signal alone, thus reducing the classification accuracy. In contrast, the $(\beta+\gamma)$ band signal contains some relationship between β band and γ band, thus improving the accuracy of classification.

In order to verify the performance of the proposed framework, we use the test set to evaluate the performance of the proposed model. The classification accuracy of the model optimized by BADEBA algorithm is 74.86% and 75.61% in the two dimensions of valence and arousal, respectively. The results of DEAP data classification are compared with those of other literature in Table 3.

As can be seen from Table 3, compared with the models proposed by Chung and Yoon, the proposed model in this paper not only reduces the channels, but also increases the classification accuracy. The number of channels selected by Dai's model is less than that of the model proposed in this paper, but the accuracy is reduced. Therefore, the model proposed in this paper has a great potential value for application.

5. Conclusions

Based on the emotion recognition of EEG signals, a novel evolutionary optimization method is proposed to select the important channels and optimize the parameters of support vector machine to achieve higher classification accuracy. This paper has two main contributions. One is to combine the binary adaptive differential evolution algorithm with the bat algorithm to improve the diversity of the population, which not only has a strong global search ability in the early iteration, but also has a strong local search ability in the late iteration. The other is to improve the signal-to-noise ratio of emotion-related EEG signals through wavelet packet decomposition and reconstruction and then extract EEG features by CSP filtering for classification. The results show that the combination of signals in the $(\beta+\gamma)$ band has higher classification accuracy than that in the single β band or γ band. Finally, the improved BADEBA algorithm is applied to the channel selection problem of emotion recognition based on EEG signals, and the validity and feasibility of the algorithm are proved by tests. How to use other intelligent algorithms to combine with binary differential evolution algorithm, such as grey wolf algorithm [29], and how to use other intelligent algorithms to combine with bat algorithm, such as flower pollination algorithm [30], to improve the performance of the model, can be used for future research directions.

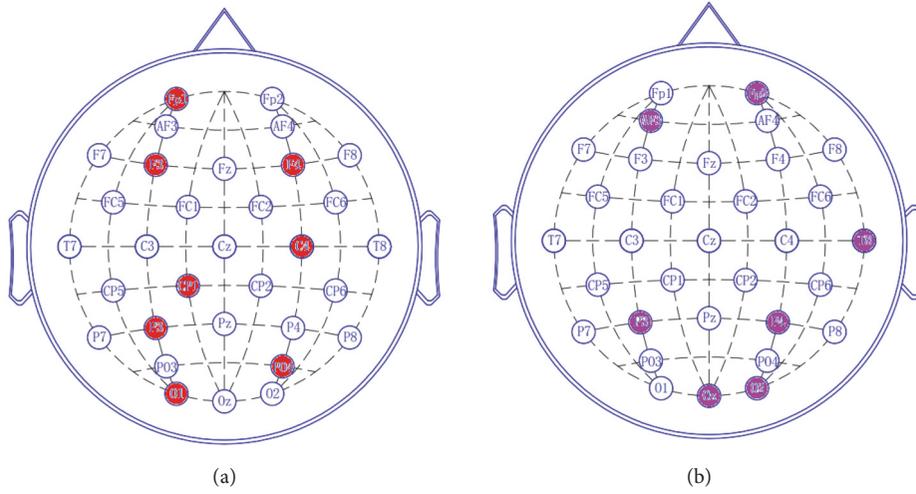


FIGURE 2: Channels selected based on valence (a) and arousal (b).

TABLE 1: Channels selected, average accuracies, and standard deviations (%) by four optimal algorithms in valence.

Methods	# Channel	Channel location	Mean \pm S.D.
Origin(Baseline)	32	All	55.97 \pm 13.83
GA	15	Fp1,AF3,F3,Fz,F4,FC1,FC2,T7, CP1,P3,Pz,PO3,PO4,O1,Oz	61.63 \pm 12.54
BPSO	12	Fp1,Fp2,AF3,F3,F4,FC2,T7,CP1, P3,PO3,O1,Oz	69.38 \pm 10.76
BA	10	Fp1,Fp2,F3,F4,C4,T8,P3,O1,Oz,O2	72.42 \pm 9.93
BADEBA	8	Fp1,F3,F4,C4,CP1,P3,PO4,O1	75.26 \pm 9.72

TABLE 2: Channels selected, average accuracies, and standard deviations (%) by four optimal algorithms in arousal.

Methods	# Channel	Channel location	Mean \pm S.D.
Origin(Baseline)	32	All	55.24 \pm 14.21
GA	14	Fp1,Fp2,AF3,Fz,F4,T7,Cz,T8,CP2, P3,Pz,PO4,Oz,O2	60.76 \pm 12.86
BPSO	11	Fp1,AF3,Fz,F4,FC2,T7,CP2,CP6, P4,O1,Oz	69.95 \pm 11.38
BA	9	Fp2,AF3,F3,FC2,C3,T8,P3,PO3,O2	73.51 \pm 10.17
BADEBA	7	Fp2,AF3,T8,P3,P4,Oz,O2	75.98 \pm 10.85

TABLE 3: Comparison with related work using EEG signals in DEAP dataset.

Authors	# Channel	Approach in the model	Accuracy(%)
Chung et al. [21]	32	Power spectral analysis with Bayes classifier	66.6(valence,2-class)
	32		53.4(valence,3-class)
Yoon et al. [22]	32	FFT enhanced feature extraction and classification	70.9(valence,2-class)
	32		70.1(arousal,2-class)
Dai et al. [23]	5	Sparse constrained differential evolution enabled hybrid selection	73.0(valence,2-class)
	5		74.5(arousal,2-class)
Our research	8	BADEBA optimal channel and classification	74.86(valence,2-class)
	7		75.61(arousal,2-class)

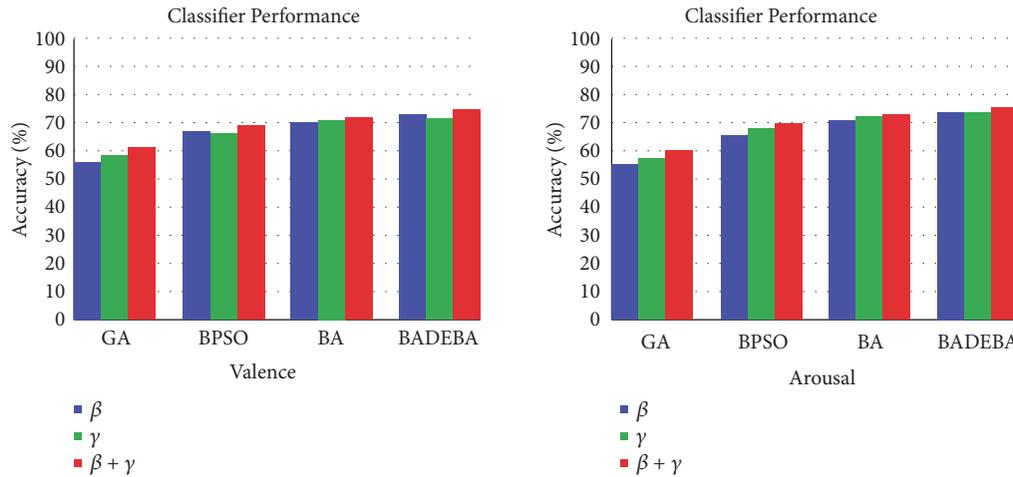


FIGURE 3: Classification results in β band, γ band, and $(\beta + \gamma)$ band.

Data Availability

The Deap dataset is a multimodal dataset for the analysis of human affective states. The electroencephalogram (EEG) and peripheral-physiological signals of 32 participants were recorded as each watched 40 one-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. The data used to support the findings of this study were supplied by Queen Mary University of London, United Kingdom, under license and so cannot be made freely available. Requests for access to these data should be made to Queen Mary University of London, United Kingdom, <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html> or email: i.patras@eecs.qmul.ac.uk.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61373116 and Grant 61572399 and the Project of Science and Technology Department of Shaanxi Province of China (Grant No. 2019ZDLGY07-08, 2018GY-013).

References

- [1] P. Purnamasari, A. Ratna, and B. Kusumoputro, "Development of Filtered Bispectrum for EEG Signal Feature Extraction in Automatic Emotion Recognition Using Artificial Neural Networks," *Algorithms*, vol. 10, no. 2, p. 63, 2017.
- [2] M. L. Menezes, A. Samara, L. Galway et al., "Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset," *Personal and Ubiquitous Computing*, vol. 21, no. 6, pp. 1003–1013, 2017.
- [3] J. Chen, B. Hu, P. Moore, X. Zhang, and X. Ma, "Electroencephalogram-based emotion assessment system using ontology and data mining techniques," *Applied Soft Computing*, vol. 30, pp. 663–674, 2015.
- [4] Z. Liu, M. Wu, W. Cao et al., "A facial expression emotion recognition based human-robot interaction system [J]," 2017.
- [5] M. I. Rashid, M. Hasan, N. Yeasmin et al., "Emotion recognition based on vertical cross correlation sequence of facial expression images," in *Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS '17)*, pp. 1344–1347, IEEE, 2017.
- [6] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, and V. Palade, "Stacked deep convolutional auto-encoders for emotion recognition from facial expressions," in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN '17)*, pp. 1586–1593, IEEE, 2017.
- [7] D. V. Sang and P. T. Ha, "Discriminative deep feature learning for facial emotion recognition," in *Proceedings of the 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR '18)*, pp. 1–6, IEEE, 2018.
- [8] R. Panda, R. M. Malheiro, and R. P. Paiva, "Novel audio features for music emotion recognition," *IEEE Transactions on Affective Computing*, 2018.
- [9] A. Mooman and S. Powers, "A novel approach to neural network based intelligent systems for emotion recognition in audio signals," in *Proceedings of the 2017 IEEE International Conference on Information Reuse and Integration (IRI '17)*, pp. 115–120, IEEE, 2017.
- [10] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.
- [11] F. Nasoz, C. L. Lisetti, K. Alvarez et al., "Emotion recognition from physiological signals for user modeling of affect," in *Proceedings of the 3rd Workshop on Affective and Attitude User Modelling*, Pittsburgh, Pa, USA, 2003.
- [12] J. Zhang, M. Chen, S. Zhao, S. Hu, Z. Shi, and Y. Cao, "Relief-Based EEG Sensor Selection Methods for Emotion Recognition," *Sensors*, vol. 16, no. 10, p. 1558, 2016.
- [13] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Optimizing the channel selection and classification accuracy in EEG-based

- BCI,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1865–1873, 2011.
- [14] H. Fauzi, M. I. Shapiai, N. A. Setiawan, J. Jaafar, and M. Mustafa, “Channel selection for common spatial pattern Based on energy calculation of motor imagery EEG signal,” in *Proceedings of the 2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC '17)*, pp. 33–39, IEEE, 2017.
- [15] L. He, Y. Hu, Y. Li, and D. Li, “Channel selection by Rayleigh coefficient maximization based genetic algorithm for classifying single-trial motor imagery EEG,” *Neurocomputing*, vol. 121, pp. 423–433, 2013.
- [16] T. Lan, D. Erdogmus, A. Adami, S. Mathan, and M. Pavel, “Channel selection and feature projection for cognitive load estimation using ambulatory EEG,” *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 74895, 12 pages, 2007.
- [17] K. Ansari-Asl, G. Chanel, and T. Pun, “A channel selection method for EEG classification in emotion assessment based on synchronization likelihood,” in *Proceedings of the 2007 15th European Signal Processing Conference*, pp. 1241–1245, IEEE, 2007.
- [18] R. Lahiri, P. Rakshit, A. Konar, and A. K. Nagar, “Evolutionary approach for selection of optimal EEG electrode positions and features for classification of cognitive tasks,” in *Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC '16)*, pp. 4846–4853, IEEE, 2016.
- [19] R. Lahiri, P. Rakshit, and A. Konar, “Evolutionary perspective for optimal selection of EEG electrodes and features,” *Biomedical Signal Processing and Control*, vol. 36, pp. 113–137, 2017.
- [20] X.-S. Yang, “A new metaheuristic bat-inspired algorithm,” in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*, pp. 65–74, Springer, Berlin, Germany, 2010.
- [21] S. Y. Chung and H. J. Yoon, “Affective classification using Bayesian classifier and supervised learning,” in *Proceedings of the 2012 12th International Conference on Control, Automation and Systems*, pp. 1768–1771, IEEE, 2012.
- [22] H. J. Yoon and S. Y. Chung, “EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm,” *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2230–2237, 2013.
- [23] Y. Dai, X. Wang, P. Zhang, W. Zhang, and J. Chen, “Sparsity constrained differential evolution enabled feature-channel-sample hybrid selection for daily-life EEG emotion recognition,” *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 21967–21994, 2018.
- [24] R. Storn and K. Price, “DE—a simple and efficient adaptive scheme for global optimization over continuous space,” *Technical Report*, vol. 25, no. 6, pp. 95–102, 1995.
- [25] S. Koelstra, C. Mühl, M. Soleymani et al., “DEAP: a database for emotion analysis; using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [26] W.-L. Zheng, H.-T. Guo, and B.-L. Lu, “Revealing critical channels and frequency bands for emotion recognition from EEG with deep belief network,” in *Proceedings of the 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER '15)*, pp. 154–157, IEEE, 2015.
- [27] Y. Peng, J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu, “EEG-based emotion recognition with manifold regularized extreme learning machine,” in *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 974–977, IEEE, 2014.
- [28] J. Kennedy and R. C. Eberhart, “A discrete binary version of the particle swarm algorithm,” in *Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 5 of *Computational cybernetics and simulation*, pp. 4104–4108, IEEE, 1997.
- [29] H. Liu, G. Hua, H. Yin, and Y. Xu, “An intelligent grey wolf optimizer algorithm for distributed compressed sensing,” *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 172319, 10 pages, 2018.
- [30] R. Wang and Y. Zhou, “Flower pollination algorithm with dimension by dimension improvement,” *Mathematical Problems in Engineering*, vol. 2014, Article ID 481791, 9 pages, 2014.

