BIO Web of Conferences **59**, 03012 (2023) https://doi.org/10.1051/bioconf/20235903012
ICBB 2023

Feature Optimization of EEG Signals Based on Ant Colony Algorithm

Zhang Shengjie[1,a], Pan Rongkai[1,b], Liu Guanglu[1,c]

[1]School of electronic Information, Xi' an Polytechnic University, China,

Abstract: EEG signal can be understood as a kind of bioelectrical signal, which can reflect emotional information when the body is in different emotional states. However, the data collected are often high-dimensional. including many irrelevant or redundant features. The high-dimensional features make the space cost increase exponentially, which brings many difficulties to the research. Ant colony optimization algorithm, a swarm intelligence algorithm, can be used for feature selection. Ant colony optimization algorithm is used for feature selection of EEG signals. The feature subset to be selected is trained cooperatively and learned actively. The classification accuracy is evaluated through convolutional neural network, and the optimal subset is selected from the iterative local optimal solution. The results show that the ant colony optimization algorithm can effectively reduce the time complexity and calculation cost, Improve the accuracy of classification.

1. INTRODUCTION

With the rapid development of modern science, researchers pay more and more attention to the study of EEG as a physiological signal, but its high dimension, instability, poor anti-interference, nonlinear and other characteristics lead to great difficulties in the study of EEG.However, the emergence of swarm intelligence algorithm can solve this kind of problem well. By reducing the dimension of data, it can effectively reduce the time complexity and calculation cost in the process of data processing.

2. BASIC CHARACTERISTICS OF EEG SIGNALS

Electroencephalogram (EEG) is a kind of weak biological signal, which is considered as a nonlinear complex connected by multiple units, and is vulnerable to interference. At the same time, EEG is one of the very important physiological signals. EEG signals are usually divided into spontaneous EEG signals and evoked EEG signals according to whether they respond to external stimuli. Spontaneous electroencephalogram (EEG) refers to the spontaneous electrical activity of cortical neurons rather than the brain being stimulated by the outside world. The signal amplitude of spontaneous EEG signal is larger than that of evoked EEG signal, which is generally between $0 \sim 75\mu V$ and has a large frequency range, and the potential will change with the change of conditions, so it is a non time locked signal; The amplitude of evoked EEG signal is relatively small, generally between $2 \sim 10\mu V$ and frequency coverage is lower than the former. It has the characteristics of constant waveform and constant latency, and is a time locked signal[1].

Specifically, EEG signals have the following characteristics:

2.1 Poor Anti-Interference, Low Signal-to-Noise Ratio and Weak Signal

The EEG signal itself is weak, so it is very vulnerable to environmental and external factors during the acquisition process, such as internal interference from eye movement, muscle movement, heartbeat and external interference from acquisition equipment and other electronic devices .

2.2 Non Stationarity and Strong Randomness

Because there are many factors that affect the generation of EEG signals, and the change law has not been fully understood, there are natural differences between different individuals, and different brain regions of the same individual also affect each other, making EEG signals show the characteristics of changing with time , so it is called unstable signal[2].

2.3 Nonlinear

The brain responds conditionally to external stimuli from

[a]17651722678@163.com
[b]1160372028@qq.com
[c]2472127689@qq.com

the body, so the collected EEG signals are generated by the superposition of multiple signals, showing nonlinear characteristics.

2.4 Rhythmicity

The frequency band of EEG signals covers a wide range. EEG signals of different individuals have certain common characteristics in amplitude, frequency and phase. In the frequency domain, they can be divided into four bands, namely α rhythmic wave, β rhythmic wave, δ rhythmic wave, and θ rhythmic wave[3].

3. ANT COLONY ALGORITHM

3.1 Ant Colony Algorithm Source

In 1992, Italian scholar Marco Dorigo proposed the concept of "ant colony system", which is a swarm intelligence algorithm, by simulating the collective path finding behavior of ants in nature. The feasible solution of the problem to be optimized is represented by the ant 's walking path, and all paths of the entire ant colony constitute the solution space of the problem to be optimized. Ants with shorter routes release more pheromones. As time goes on, the concentration of pheromones accumulated on the shorter routes gradually increases, and the number of ants choosing this route also increases. Finally . the whole ant will focus on the best path under the effect of positive feedback, and the corresponding optimal solution of the problem to be optimized[4]. Ant colony algorithm has the characteristics of distributed computing, positive feedback and heuristic search. Different from the real ant colony, the ants in the ant colony algorithm have the ability to remember the current local optimal solution and the global optimal solution, and can optimize the selection by changing the environmental factors.

3.2 Application of Ant Colony Algorithm

Ant colony algorithm was initially used to solve space problems such as TSP (Traveling Salesman Problem). Later, it was found that ant colony algorithm had advantages in solving complex optimization problems, and showed great potential for feature optimization[5]. The ant colony algorithm simulates the feature selection problem as the minimum path problem of the search graph. Each feature is regarded as a node to construct a search graph, and the nodes are connected with each other. By constructing a search graph model, ants can access features one by one to realize feature selection. The probability transfer rules between features are determined by pheromone information and local heuristic information[6]. When using the ant colony optimization algorithm to solve the feature selection problem, the commonly used binary coded number string list candidate feature subset is different from the basic binary ant colony optimization algorithm. This paper uses the random binary ant colony optimization algorithm to improve its search space on the basis of retaining the binary ant colony optimization algorithm, and proposes the concept of access probability. The access concept is not based on any prior knowledge. In this paper, the access probability is the same, and the two-step feature node selection is carried out by combining the maximum correlation minimum redundancy rule[7][8]. Select the next access feature node through the access probability, and then select the child node according to the transition probability. 1 means to select the corresponding feature sub node, and 0 means not to select the sub node. The selection mechanism of feature nodes can be described as follows:

Step 1: the beginning, the ant randomly selects a feature node;

Step 2: The ant observes the node to be accessed in the search space at node i_x, and selects the feature j for the next visit through the access concept;

Step 3: After confirming the next access feature, select the feature sub node through the transition probability between i_x and j_y;

Step 4: Repeat Step 2 and Step 3 until all features are traversed. Since there is no need to compare the transfer probability between feature sub nodes, the efficiency of feature selection is greatly improved and the calculation cost is reduced.

4. FEATURE OPTIMIZATION

4.1 Optimization Method

In order to solve the problems of redundancy feature mixing and lack of sensitive features in feature extraction of EEG signals, this paper proposes a new feature optimization method, which first uses ant colony algorithm to optimize feature subsets, and then improves the existing semi supervised learning algorithm through collaborative training, active learning and other methods. Combined with ant colony optimization features, reduce the demand of the algorithm for samples and improve the accuracy of classification[9]. EEG researchers divide the regular characteristics of EEG into four categories: time-domain characteristics, frequency -domain characteristics, time-frequency characteristics and nonlinear characteristics. A large number of experimental studies have shown that single feature parameter extraction or multi domain feature parameter extraction is not ideal for classification, ignoring the fact that there is some overlap and correlation between the various feature parameters. The trained classifier model often reduces the reliability and accuracy of the classifier due to over fitting, or the speed of classification due to the excessive amount of extracted feature parameters. The low efficiency of classification leads to the phenomenon of "dimension disaster"[10]. Therefore, the ant colony optimization algorithm is used to reduce the dimensions of the original feature set to improve the accuracy of classification. With this algorithm, the ant colony can obtain a group of local optimal solutions every time it traverses, including multiple features. The quality of features is evaluated through the fitness function. Finally, the optimal feature subset is obtained from multiple local optimal solutions through information element update and multiple iterations[11][12].

4.2 Implementation Step

Input: original EEG characteristic parameter set, initial parameters of ant colony algorithm, maximum iteration number, etc.

Output: local optimal solution and optimal feature subset of each iteration.

Step l: parameter initialization;

Step 2: m ants were randomly distributed at n feature points to build a tabu table $tabu_k$, and the visited features were placed in the tabu table $tabu_k$;

Step 3: Calculate the transition probability of ant k accessing the next feature according to Equation (1);

$$p_{ij}^{k}(t)=\begin{cases}\dfrac{\tau_{ij}^{\alpha}(t)\cdot\eta_{ij}^{\beta}(t)}{\sum\tau_{is}^{\alpha}(t)\cdot\eta_{is}^{\beta}(t)} & j,s\notin tabu_k\\ 0 & \text{otherwise}\end{cases}\tag{1}$$

Where, $tabu_k$ ——Ant k Tabu List

$\tau_{ij}(t)$ ——Pheromone concentration from feature i to feature j at t

$\eta_{ij}(t)$ ——Heuristic function, whose size is determined by the classification accuracy

p_{ij}^{k} ——Probability of the k ant from feature i to feature j

Step 4: After ant k traverses to obtain feature subsets, it evaluates each feature subset through the trained classifier to obtain the corresponding classification accuracy;

Step 5: Calculate the classification accuracy and fitness function value of each ant traversal feature through Equation (2). The higher the function value, the better the subset will be . Select the highest fitness value as the local optimal solution in this iteration process ;

$$f(k)=\frac{R(k)}{1+\lambda n(k)}\tag{2}$$

Where , R ——correct rate of feature subset classification

f ——Fitness value of ant k

n ——Number of subset features

λ ——Weight ratio of feature number

Step 6: update pheromone according to Equation (3) and (4), and conduct the next iteration:

$$\tau_{ij}(t+n)=(1-\rho)\times\tau_{ij}(t)+\Delta\tau_{ij}\tag{3}$$

$$\Delta\tau_{ij}=\sum_{k=1}^{n}\Delta\tau_{ij}^{k}\tag{4}$$

Where , $\Delta\tau_{ij}$ ——pheromone change in iteration

ρ ——volatilization coefficient

Step 7: Judge whether the number of iterations meets the maximum number of iterations. If yes, output the results. If not, skip to step 2.

From this, the local optimal solution. classification accuracy, fitness value and running time after each iteration can be obtained. and then the optimal feature subset can be obtained based on the running time and accuracy. The execution flow chart is shown in 3.1.

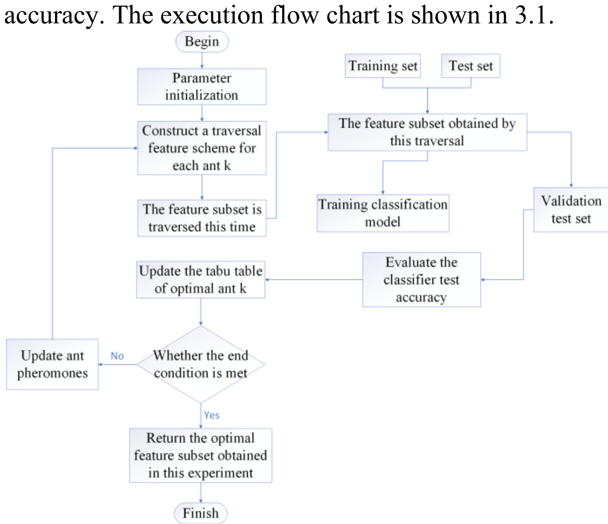

Figure 1 Flow chart of feature optimization based on ant colony algorithm

5. SIMULATION RESULTS

In order to verify the advantages of ant colony algorithm in eliminating redundant features and improving classification accuracy. DEAP data set is used for experimental verification. The original feature set extracted is: average value, standard deviation, skewness, kurtosis, maximum value, minimum value, power spectral density of four rhythmic waves (α , β , δ and θ bands), sample entropy and fuzzy entropy, a total of 12 features are optimally extracted through ant colony algorithm. Set the initial parameters as : α =0.4, β =2.7, $\tau_{ij}(0)$ =1, ρ =0.25, λ =0.01, m =8(set according to the ratio of the number of features to the number of ants is 1.5), extract 1280 samples from the DEAP database, and take 75% as the training set (matrix size is 960*12), and 25% as the test set (matrix size is 320*12). The feature subset obtained by each traversal is tested through CNN convolutional neural network. This experiment has tested 10 times in total, and the number of iterations is set to 30. The experimental results are shown in Table 1:

Table 1 Result of feature optimization

Number of experiments	Classification accuracy	Number of Characteristics	Operation time/s
1	78.66%	4	27.398
2	79.33%	5	36.365
3	78.46%	3	15.961
4	76.25%	4	24.254
5	79.88%	4	25.862
6	75.33%	3	16.358

7	80.21%	4	21.127
8	79.33%	5	29.628
9	78.43%	5	32.643
10	77.82%	4	22.368
Original feature set	67.25%	12	68.354

6. Conclusions

The experimental results show that in the seventh experiment, the classification accuracy is the highest, 80.21%, 12.96% higher than the original feature set. The optimal feature subset is: skewness, β, δ band power spectral density and fuzzy entropy, which account for two frequency domain features in the optimal feature subset, indicating that the power spectral density of rhythmic waves plays a dominant role in this EEG change rule. At the same time, it shows the superiority of ant colony algorithm in EEG feature optimization, which can improve the calculation efficiency and accuracy of the algorithm, and reduce the time complexity and calculation cost.The algorithm is improved on the basis of binary ant colony algorithm, and the access probability is proposed to study the feature selection in EEG emotion recognition, so as to improve the reliability of emotion recognition.The setting of access probability is not based on any prior knowledge, and the setting probability is the same. Different access probabilities can also be further set for control experiment to find the optimal access probability so as to optimize the algorithm.In this paper, the algorithm is improved from the aspect of search mode. In addition, it can be improved and perfected by optimizing search operator and designing new objective function, so as to find the optimal subset more efficiently, improve the operation efficiency of the algorithm and increase the practicability of the algorithm.

Acknowledgements

National innovation and entrepreneurship training program for college students

References

[1] Yue Zongtian. Research on preprocessing and sleep staging based on multi lead EEG signals [D]. Nanjing University of Posts and Telecommunications, 2017.

[2] Navarro X, Pore F. Beuche A, etal. Denoising preterm EEG by signal decomposition and adaptive filtering : a comparative study [J]. Medical Engineering & Physics ,2015,37(3):315-320.

[3] Aslanyan EV, Kiroi VN, Lazurenko DM, EEG spectral characteristics during voluntary motor activity [J]. Neuroscience and Behavioral Physiology, 2015,45(9):1029-1037.

[4] Li Xuan, Wu Xiaobing, Liu Qiong. Logistics distribution center location problem solving based on adaptive elite improved ant colony algorithm with mutation [J]. Journal of Chengdu University (Natural Science Edition). 2022.41(01):46-51.

[5] Zhang Jiehui, He Zhongshi, Wang Jian, Huang Xuequan. Combined feature selection algorithm based on adaptive ant colony algorithm [J]. Journal of System Simulation, 2009,21(06):1605-1608+1614.

[6] Sun Qian, Zhang Jin, Wang Yuxiang.A review of ant colony algorithm optimization strategies[J].Information Security and Technology, 2014, 5（02）：22-23+27.

[7] Chen Bianna. Research on Feature Selection Algorithm Based on Evolutionary Computing [D]. South China University of Technology, 2020.

[8] Liang Benlai, Zhu Lei. Intrusion Detection Method based on improved Ant colony solving feature subset[J].Computer Application Software, 2021, 38（07）：323-331.

[9]Hou Yuanshao. Feature gene selection algorithm based on ant colony optimization[J].Journal of Zhongzhou University, 2019, 36（06）：120-123.

[10] Xiao Yunshuang. Random binary full connection ant colony optimization algorithm and its dimension reduction application in high - dimensional medical data [D]. Chongqing University, 2021.

[11]Li Zhanshan, Liu Zhaogeng, Yu Yin, Yan Wenhao. Quantized pheromone ant colony optimization feature selection algorithm[J].Journal of Northeastern University (Natural Science Edition), 2020, 40（01）：17-22.

[12]Li Kaiqi, Diao Xingchun, Cao Jianjun, Li Feng. High precision text feature selection method based on improved ant colony algorithm[J].Journal of PLA University of Science and Technology (Natural Science Edition), 2010,11(06):634-639.