Improving Physiological Emotion Recognition Using Hybrid SVM-Firefly and Random Forest Models

Yufei Wang[1]*, Quan Wen[2], Weiwei Xu[1]
[1]School of Electrical and Information Engineering, JiLin Engineering Normal University, Changchun 130052, China
[2]College of Communication Engineering, Jilin University, Changchun, 130012, China
E-mail: yufeiwang202404@163.com

Keywords: machine learning, emotional recognition, emotional signals, feature extraction

Received: May 30, 2025

Aiming at the problems of low recognition accuracy of physiological and emotional signals and susceptibility to interference during the recognition process, an improved algorithm based on machine learning is proposed. The support vector machine algorithm and firefly algorithm model are used for data classification and recognition of physiological signals. The random forest algorithm is added to improve the recognition and decision-making performance of the algorithm model. The experiment is based on the publicly available DEAP dataset, which includes 32 subjects, 1,280 3-minute multi-modal physiological models, and a sampling rate of 360Hz. Taking five emotions of happiness, sadness, anger, fear, and neutrality as tasks, the accuracy and F1 value are used for evaluation, and comparisons are made with support vector machine, random forest algorithm, convolutional neural network, and long short-term memory network fusion algorithm. The results showed that among the five types of emotion changes, sample 2 had the highest recognition accuracy for different emotion signals, with an overall accuracy of over 90.00%, indicating that this model could effectively extract features from the signals. The highest accuracy of the proposed model reached 94.60%, which was approximately 3.10% higher than that of the sample with the lowest accuracy. Its average F1-macro reached as high as 90.3%, outperforming the comparison model in some emotion categories. However, the lowest F1 value of 84.9% occurred in emotion state 4, slightly lower than that of the comparison model. The designed model has better accuracy and higher recognition performance in emotion recognition. This has good guiding value for the research on improving the accuracy of emotion recognition for different signals at present.

Povzetek:

1 Introduction

With the rapid development of the social economy, social pressure is increasing dramatically. More people are influenced by negative emotions. The incidence rate of mental illness is gradually increasing [1]. With the advancement of intelligent technology, the methods for mental health prevention and detection are becoming increasingly sophisticated [2]. More people with mental problems can receive timely detection and psychological treatment. Meanwhile, poor performance in emotion recognition became the biggest issue [3]. Machine learning, as an advanced intelligent model, is often used for intelligent recognition and detection analysis due to its rapid data processing and analysis capabilities [4]. However, for the recognition of human physiology and emotions, pure machine learning methods are subject to varying degrees of limitations, resulting in a single algorithm model being unable to achieve good emotion analysis and recognition in emotion recognition research. Therefore, a new model based on machine learning is proposed to address the low accuracy and poor performance in emotion recognition. The machine learning is utilized to improve the classification analysis of emotion data and enhance the recognition performance of emotion recognition. The innovation of this study lies in proposing an optimized ensemble support vector machine (SVM) based on Random Forest (RF) and Firefly Algorithm (FA). The FA is used to address the limitations of SVM hyperparameter optimization, and then the RF ensemble decision-making is driven by fused features to break through the shortcomings of a single model. The key parameters for physiological signal processing are clarified and the parameter ambiguity in existing research has been compensated for. The contribution of this study lies in the fact that the proposed model effectively enhances the anti-interference ability and overall performance of physiological and emotional signal recognition, providing feasible technical ideas and theoretical guidance for current research on improving the accuracy of physiological and emotional recognition.

The study is divided into four parts. The first part elaborates on current domestic and international research, analyzing the problems that have arisen in the current

research. The second part is to construct a research model and optimize the traditional algorithm through different methods. The third part is to test the performance of the current model. The fourth part summarizes the current research.

2 Related works

Emotions play an important role in people's daily lives, usually reflecting their cognitive and decision-making abilities, which have an important expressive effect on human activities and information transmission. Gao H et al. proposed a distribution-based uncertainty graph convolutional network method in response to the challenges of existing Electroencephalogram (EEG) emotion recognition methods in long-distance feature learning, modeling fuzzy topological relationships between EEG channels, and noise label interference. This method employed a graph convolutional network architecture to represent the spatial dependence and time-spectral correlation in EEG signals. Graph blending technology was adopted to enhance potential connected edges and alleviate noise labels. The results showed that this method was superior to previous methods [5]. To improve the accuracy and efficiency of emotion classification using EEG signals, Bardak F K et al. proposed a new hybrid emotion classification algorithm based on EEG signals. It used multiple classifiers of an adaptive neuro fuzzy inference system designed in parallel. The selected optimal feature sets were used as inputs, trained separately to capture the nonlinear feature distributions of various emotions, and the outputs of each classifier were integrated into a new feature vector, which was input into the adaptive network for final emotion classification. The results showed that the accuracy of the algorithm on European Emotion signal Database Database for Emotion Analysis using Physiological Signals (DEAP) dataset was 73.49%, and the accuracy on feeling emotions dataset was 95.97% [6]. Ju X et al. proposed an emotion recognition method based on the time-difference minimization neural network to effectively utilize the knowledge of emotional activities. This method employed the prior knowledge of emotional changes over time, used the maximum mean difference technique to evaluate the electroencephalogram features, and minimized the differences through a multi-branch convolutional recurrent network. The results showed that this method achieved the best performance on the SEED, SeED-IV, DEAP, and DREAMER datasets [7]. Hajarolasvadi N et al. proposed a new method to obtain representative video frames in the feature space domain to improve the efficiency of facial emotion recognition in videos. The principal component analysis method extracted features from a single emotional video, extracting the most important feature frames representing the temporal motion variations embedded in the video. The experimental results showed that the new method outperformed traditional methods by 8% and 4% in performance, while reducing computational time [8].

Lin Y H et al. developed a new method for emotion recognition in social networks. The partial least squares method was used for analyzing and identifying network video data. The research results indicated that it could improve the effectiveness and accuracy of video emotion recognition [9]. Liu C et al. proposed a theoretical result based on the five major personalities and self-preferences to conduct a more in-depth analysis for emotional composition and psychological construction in current social platforms. The convolutional neural network was used to analyze human emotions. The research results indicated that the data platform used among users could reflect their true psychological situation [10]. Hariri W et al. proposed a new method for emotion and facial expression recognition using 3D data and 2D depth images to address the efficiency issue of facial emotion recognition in unconstrained environments. The new method demonstrated its complementarity in three-dimensional facial emotion recognition by combining manual features and deep learning features. The experimental results showed that the new method was significantly superior to traditional methods [11]. A P M et al. proposed another new emotion recognition test result to address issues such as emotional intensity affecting patient heterogeneity and lack of facial recognition ability. The research results indicated that controlling the emotional recognition ability of patients could improve their emotional and reduce their emotional decoding disorders [12]. The specific contents of relevant work are shown in Table 1.

Table 1: Specific contents of relevant work

Researchers	Data Modalities	Algorithms	Datasets	Performance
Gao H et al [5]	EEG	Graph convolutional network with connectivity uncertainty	DEAP	Better than previous methods in EEG emotion recognition task
Bardak F K et al [6]	EEG	Hybrid classifier based on adaptive neuro-fuzzy inference systems	DEAP Dataset, Feeling Emotions Dataset	The accuracy rate on DEAP dataset is 73.49%, and the accuracy rate on feeling emotions data set is 95.97%
Ju X et al [7]	EEG	Temporal-difference	SEED Dataset,	Best performance on

		minimizing neural network	SEED-IV Dataset, DEAP Dataset, DREAMER Dataset	SEED, SeED-IV, DEAP, and DREAMER datasets
Hajarolasvadi N et al [8]	Video	Principal component analysis with key feature frame extraction	Not specified	The performance of facial emotion recognition task is better than that of traditional methods by 8% and 4%
Lin YH et al [9]	Social network videos	Partial least squares method	Not specified	Improve the effectiveness and accuracy of video emotion recognition
Liu C et al [10]	Social media data	Convolutional neural network	Not specified	The data platform used by users can reflect their real psychological status
Hariri W et al [11]	3D and 2D depth images	Fusion of handcrafted features and deep learning features	Not specified	It is obviously superior to the traditional method in three-dimensional facial emotion recognition
APM et al [12]	Clinical patient data	Comparative test method of facial emotion recognition	Not specified	Reduce the disorder of emotion decoding and improve the ability of emotion recognition

To sum up, although significant progress has been made in existing emotion recognition research, there are still some limitations. Most existing studies focus on pathological emotion-assisted diagnosis based on facial or text cues, paying insufficient attention to emotion recognition driven by physiological electrical signals, and lack systematic strategies for improving accuracy. Therefore, this study optimizes the hyperparameters of the SVM through the FA to break through the convergence bottleneck, introduces RF to enhance the ensemble learning ability to alleviate overfitting, and combines time-domain and wavelet transform to achieve multi-dimensional physiological signal feature extraction. Finally, an improved model based on emotional signal recognition is constructed to fill the gap in existing research on the anti-interference recognition and accuracy improvement of physiological and emotional signals.

2 Intelligent recognition of physiological and emotional signals based on machine learning

This section mainly focuses on the design of intelligent emotion signal recognition method currently used in research. The SVM algorithm, FA, and RF algorithm are improved to build a new model based on emotion signal recognition, thereby enhancing the effectiveness of emotion signal recognition.

2.1 Emotion recognition based on machine learning

In machine learning algorithm models, commonly used SVM can linearly project nonlinear mapping functions into higher dimensional spaces, thereby building better classification optimal planes. However, the traditional SVM model has problems such as slow convergence and being prone to fall into local optimum during the process of hyperparameter optimization. Therefore, FA is added to the traditional SVM algorithm to improve it. FA has advantages such as strong global search ability, fast convergence speed, and avoidance of premature convergence. When combined with the SVM for improvement, it can enhance the efficiency of traditional models in finding the optimal solution. The improved algorithm flow is shown in Figure 1.

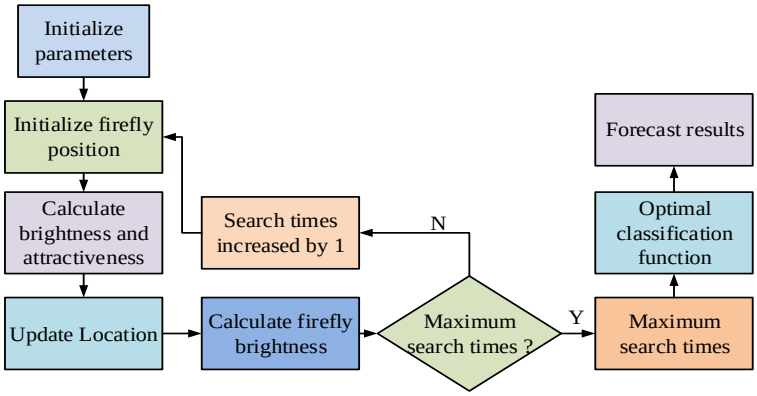

Figure 1: Improved algorithm process

From Figure 1, when improving the SVM, the parameters in the FA are first initialized. Secondly, the position of fireflies is initialized, and the brightness and attractiveness between the current fireflies are calculated. The spatial position of the firefly is updated, and then the brightness of the firefly is calculated to determine whether the current FA has reached the maximum number of searches. If it has been reached, the optimal parameters are calculated. It is considered as the initial input value of the SVM. If not, the search count is increased by 1. The attractiveness of fireflies is recalculated until the current search reaches its maximum. The SVM is trained to find the current optimal classification function. The obtained model data is trained, and the data to be tested is predicted and analyzed. Finally, the optimal training and testing data of the model are output. The position of fireflies in the dimension space is shown in equation (1) [13-14].

$$I = I_o e^{-\gamma r_{i,j}} \quad (1)$$

In equation (1), I_o represents the initial fluorescence position of fireflies. I represents the updated firefly position. γ represents the absorption coefficient of firefly fluorescence. $r_{i,j}$ represents the distance between the coordinates i and j of the firefly. The distance calculation is shown in equation (2).

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

In equation (2), $x_{i,k}$ represents the distance of the firefly from position i to position k. $x_{j,k}$ represents the distance of the firefly from position j to position k. x_i represents the position of the firefly at position i. x_j represents the position of the firefly at position j. The attractiveness of fireflies is shown in equation (3) [15].

$$\beta(r) = \beta_0 e^{-\gamma r_{i,j}^2} \quad (3)$$

In equation (3), $\beta(r)$ represents the attractiveness of fireflies. β_0 represents the initial attractiveness of fireflies. The optimal iteration of the algorithm is shown in equation (4) [16-17].

$$x_i(t+1) = x_i(t) + \beta(x_j(t) + x_i(t)) + \alpha(rand + 1/2) \quad (4)$$

In equation (4), t represents the number of iterations of the model used in the study. α represents the step size factor. $rand$ represents the random factor of the target, with a range of values between 0 - 1. To solve the poor generalization ability and accuracy of the optimal solution in traditional models, the RF is added to the above model. The structure of RF is shown in Figure 2.

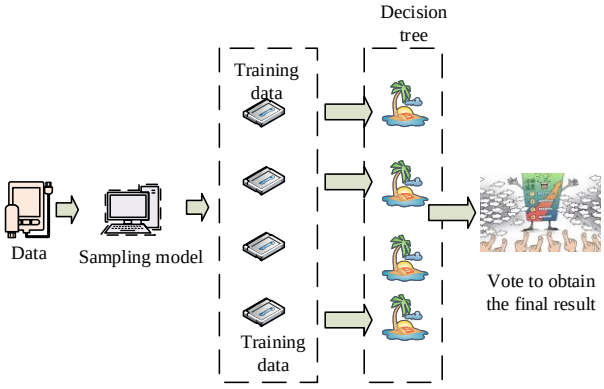

Figure 2: RF structure

From Figure 2, the main principle is to build an independent decision tree model by constructing the dataset of the model mentioned above. Next, the feature data of the dataset is extracted. The optimal solution of the data is searched for in the feature data to randomly extract the decision tree. Finally, the decision tree projection is used to obtain the final optimal classification result. The concept of ensemble learning in RF models can avoid overfitting. The process of the improved RF model is shown in Figure 3.

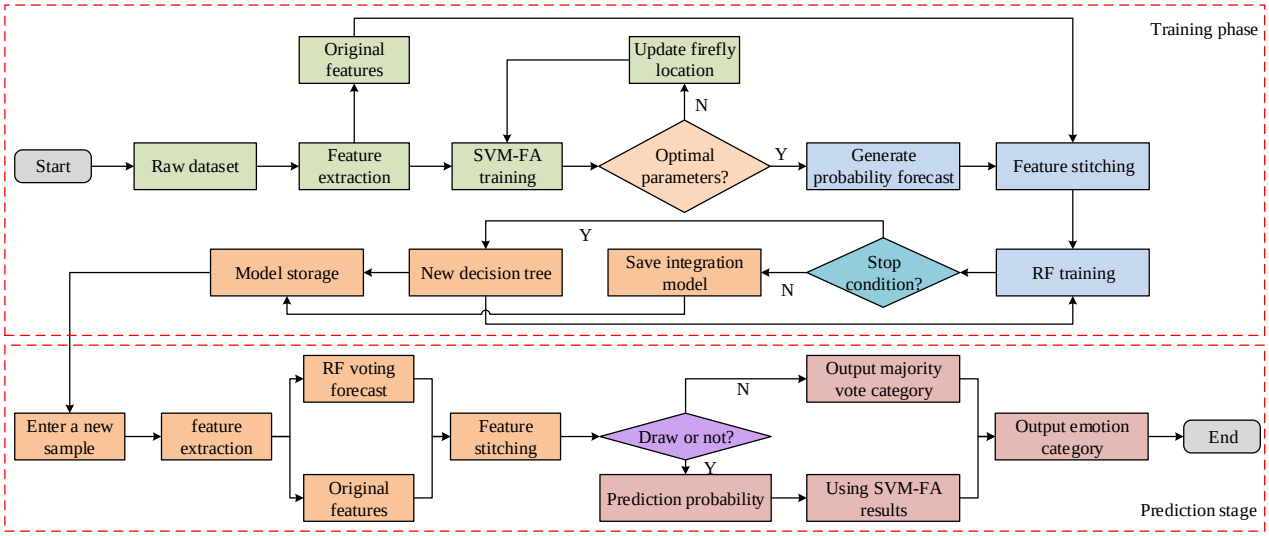

Figure 3: Improved RF algorithm process

From Figure 3, RF and SVM-FA are integrated in series. Firstly, the optimized SVM-FA model is used to train the training set, and the output device outputs the emotion prediction probability of the training samples. The predicted probabilities output by the SVM-FA model are taken as new features and concatenated with the originally extracted physiological signal features to jointly form meta-features. Then, the RF model is trained on a training set composed of meta-features. At this stage, the RF model acts as a secondary learner and makes the final emotion classification decision by integrating the prediction probability of SVM-FA and the original feature information. Compared with the traditional RF model, this method significantly improves the local classification performance and effectively enhances the overall accuracy of emotion recognition. The minimum principal calculation method in the model is shown in equation (5) [18].

$$X_i = 1 - \sum_{i=1}^{m} p(i)^2 \quad (5)$$

In equation (5), $p(i)$ represents the proportion of decision tree examples to the total sample data. X_i represents the minimum principal size obtained.

2.2 Construction of an emotion recognition model

This study aims to solve a multi-category emotion classification problem. Specifically, the model is to classify the input signals into five basic emotional states: happiness, sadness, anger, fear, and neutrality, based on the features extracted from physiological signals. The input of the model is a variety of time-domain and frequency-domain features calculated based on human physiological electrical signals. The output of the model is the probability distribution or final classification label of the input samples belonging to the above five emotion categories. The performance optimization objective of the model is to achieve higher classification accuracy and F1 values on these five classification tasks. Emotional recognition is the judgment of current physiological signals and emotions. It is generally recognized and analyzed through human physiological electrical signals. The general physiological signals of the human body include R-waveform, Q-waveform, S-waveform, P-waveform, and T-waveform signals. These data signals usually need to be detected in different ways in the human body. The processed signal data can be recognized and analyzed by the model. Therefore, the monitoring process for R-waveform signals in the emotional signal cycle of the human body is shown in Figure 4.

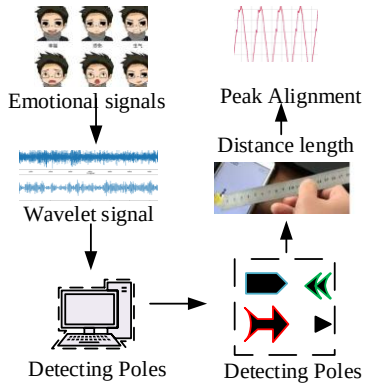

Figure 4: Monitoring process of R-waveform signal

From Figure 4, the signal processing of the R-waveform requires first decomposing the wavelet signals of different layers. Model detection can obtain the extreme points of waveform signals, then find the zero coordinates of the current model, and determine the peak position of the current R waveform. According to the current waveform period, the peak coordinates of the current waveform are calibrated. The peak calculation

method for Q-waveform and S-waveform is shown in Figure 5.

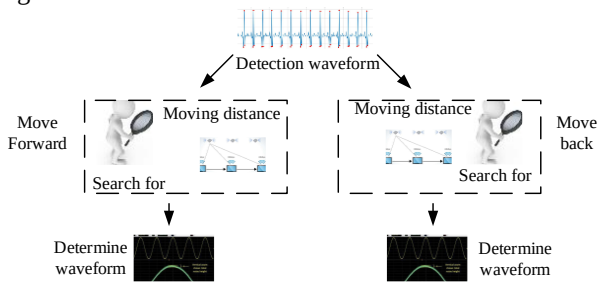

Figure 5: Peak calculation method for Q-waveform and S-waveform

From Figure 5, when detecting the waveform, the first step is to correctly detect the R-waveform. The current standard point of the R-waveform is searched forward and placed in the matrix with an interval of 0.1s. Then, the backward search is placed in the matrix with an interval of 0.1s. The forward search point determines the peak of the Q-waveform, and the backward search point determines the peak of the S wave. At this point, the waveform signal obtained through the matrix is the peak position of points Q and S. After determining the peak point, 0 is used as the threshold of the current matrix to determine the peak point and obtain the waveform peak of points Q and S. The judgment of P-waveform and T-waveform signals are shown in Figure 6.

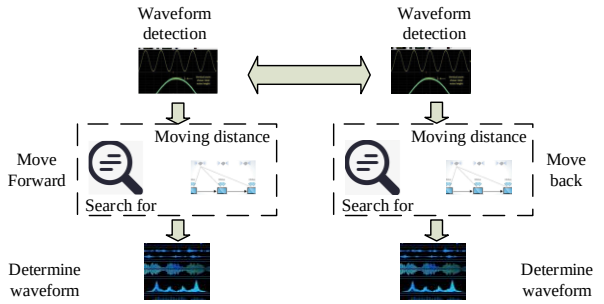

Figure 6: Analysis of P-waveform and T-waveform signals

From Figure 6, the P-waveform and T-waveform are determined to use the Q-waveform as the reference point for their position. The current standard point of the Q-waveform is searched forward and placed in the matrix with an interval of 0.2s. Then, the backward search is placed in the matrix with an interval of 0.2s. The forward search point determines the peak of the P-waveform. The backward search point determines the peak of the T-waveform. Similar to the above method, after determining the waveform position, the current peak position is determined through threshold setting. After obtaining the peak position information of the emotion waveform, the emotion signal of the current signal waveform is obtained through feature extraction. Then, the emotional waveform recognition is achieved through signal data processing. Therefore, data feature extraction is the most important part in the model construction. For this purpose, the study takes time-domain feature extraction for data extraction. The waveform average calculation method for time-domain features is shown in equation (6).

$$\bar{X} = \frac{1}{N}\sum_{m=1}^{N} X_m \quad (6)$$

In equation (6), $\bar{X}$ represents the average value of the emotional signal waveform. N represents the number of points sampled for the current feature extraction. X_m represents the feature amplitude of the m-th point in the current feature extraction. The mean in emotional signals can reflect the average state of current emotions. The mean is shown in equation (7).

$$\overline{RR} = \frac{1}{N}\sum_{m=1}^{N} RR_m \quad (7)$$

In equation (7), RR_m represents the interval size at the m-th position. N is the length size of the data. $\overline{RR}$ represents the average size of emotional signals. The standard deviation of the mean is shown in equation (8) [19].

$$SDNN = \sqrt{\frac{1}{N}\sum_{m=1}^{N}(RR_m - \mu)^2} \quad (8)$$

In equation (8), m represents the sampling point position corresponding to the peak. μ represents the average size of RR at position m. The standard deviation for the two intervals at this time is shown in equation (9).

$$SDSD = \sqrt{\frac{1}{N}\sum_{m=1}^{N}[((RR_m - RR_{m-1}) - \mu)^2]} \quad (9)$$

In equation (9), $RR_m - RR_{m-1}$ represents the time insertion change between two adjacent RR. The main method for extracting time-frequency signals from emotional signals is to extract features from wavelet transform signals. Daubechies 4 (db4) wavelet is adopted as the mother wavelet. This wavelet has tight support, nearly optimal symmetry, and good time-frequency localization characteristics [20]. Five layers of wavelet decomposition are established, and the corresponding frequency bands for each layer are calculated based on the sampling rate of the dataset. When sampling, for the preprocessed signal, the peak value of the R-waveform is taken as the center, and the signal is sampled 1.5 seconds before and after to ensure the inclusion of the complete wave group period. The signal boundary is processed by the symmetrical extension method, and the extension

length is twice the support length of the mother wavelet. Only 3 to 5 layers of wavelet coefficients are retained for feature extraction to reduce the impact of noise on feature quality. Equation (10) is the wavelet signal sub signal extraction method.

$$s_j^n(t) = \sum_k D_k^{j,n} \Psi_{j,k}(t) \quad (10)$$

In equation (10), $D_k^{j,n}$ represents the spatial signal coefficients in wavelet transform. $\Psi_{j,k}(t)$ represents a wavelet function. The subspace size of wavelet transform is U_j^{m-1}. $s_j^n(t)$ represents the sub-signal of wavelet transform. The energy variation of wavelet signal nodes is shown in equation (11) [21].

$$E_l = \sum_k \left| D_k^{j,n} \right|^2 \quad (11)$$

In equation (11), E_l represents the node energy size of the wavelet. The total energy of the obtained signal is shown in equation (12).

$$E_{all} = \sum_{i=1}^{2^j} E_l \quad (12)$$

In equation (12), E_{all} represents the total energy of the feature extraction signal in the wavelet transform. The probability change of wavelet distribution is shown in equation (13) [22].

$$p_l = \frac{E_l}{E_{all}} \quad (13)$$

In equation (13), p_l represents the energy distribution probability of the wavelet transform. The entropy change of wavelet is shown in equation (14).

$$WEP = -\sum p_l \ln[p_l] \quad (14)$$

In equation (14), WEP represents the entropy value of the wavelet. The index correlation dimension and approximate entropy value are used for emotional signal analysis to analyze different emotional signals. The analysis principle of the entire data feature extraction is to set its adjacent position point x. The distance between the initial position and adjacent positions is L. The distance evolution of time variation is used to specify the maximum value, forget the current time node, retain position information, and search for new data positions at adjacent position points. The position point satisfies the initial position angle. The distance distance and angle calculation between the two position points are repeated. The distance obtained is the optimal position of the feature data. Through this data feature extraction method, emotional signal data can be extracted. The feature extraction signal can be input into an improved algorithm model. Then, the model can recognize and predict the current emotional signal. The analysis and detection of emotional signals are completed.

3 Analysis of physiological emotion signal recognition results based on the improved algorithm

3.1 Experimental setup

To test the emotional signal detection and recognition performance of the model, this study adopts the well-known DEAP as the experimental data source (https://www.tue.nl/en/research/research-labs/signals-systems-control-lab/datasets/). It is multi-channel data collected through experiments by Koelstra et al. from Queen Mary University of London in the UK, University of Twente in the Netherlands, University of Geneva in Switzerland, and Swiss Federal Institute of Technology for studying human emotional states, which is publicly and freely accessible. The collection process has been approved by the Ethics Review Committee of Queen Mary University of London. The database contains 32 healthy subjects. Each subject needs to watch 40 3-minute emotion-inducing videos. Multi-modal physiological signals are collected simultaneously, and the sampling rate is uniformly set at 360Hz. The db4 wavelet basis 5-layer decomposition denoising and 50Hz notch filtering are adopted to suppress interference. The waveform extraction and calibration in the ECG signal are completed based on the threshold method and extreme point location. After eliminating 42 abnormal samples, 1,280 valid samples are retained, and the valid samples are mapped to five types of emotions: happiness, sadness, anger, fear, and neutral. In the experiment, the stratified sampling method is adopted to divide the data set into the training set and the test set in an 8:2 ratio. The training set includes 1,024 samples, and the test set includes 256 samples. The hyperparameters of the emotion recognition model are shown in Table 2.

Table 2: Model hyperparameter settings

Algorithm Module	Parameter	Value	Setting Basis
FA	Population Size	50	A size < 40 is prone to local optima, while a size > 60 increases computational cost by 50%. 50 achieves a balance between global search and efficiency
	Maximum Number of Iterations	100	After 80 iterations, the validation set accuracy fluctuation corresponding to the optimal SVM hyperparameters is < 0.3%. Setting it to 100 avoids redundant computation
	Initial Attractiveness	1.0	Refer to similar physiological signal optimization studies to ensure a sufficiently wide search range for fireflies in the initial stage
	Fluorescence Absorption Coefficient	0.01	A value that is too small leads to low search efficiency, while a value that is too large weakens local search ability. 0.01 achieves a balance between global and local search
	Step Factor	0.5	Ensures a moderate step size in each iteration, avoiding skipping the optimal solution due to an excessively large step size or slow convergence due to an excessively small step size
SVM	Kernel Type	RBF	The RBF kernel can handle high-dimensional feature space mapping
	Penalty Paramete	10	Within the range of 1–100, when C = 10, the test set accuracy is the highest without obvious overfitting
	Kernel Parameter	0.1	Within the range of 0.01–1, when γ = 0.1, the classification discrimination of high-dimensional physiological signals is optimal
RF	Number of Decision Trees	200	When the number of trees < 150, the model is underfitted. When the number > 200, the accuracy improves but the computational cost increases
	Feature Subset Size	Sqrt	The total number of features is 12 dimensions, so the subset size is set to 3 to avoid increased inter-tree correlation caused by feature redundancy
	Maximum Tree Depth	None	The depth is not limited. Only the minimum samples split is used to control complexity, ensuring the capture of local features of emotional signals
	Minimum Samples Split	10	When the value < 8, the tree structure is simple and cannot distinguish between sad and fearful emotional signals. When the value ≥ 10, the classification accuracy is stable
	Minimum Samples per Leaf	5	Ensures that leaf nodes have sufficient samples to support decisions, adapting to individual differences in physiological signals

To optimize the model's hyperparameters and alleviate overfitting, this study adopts 5-fold cross-validation within the training set. The training set is randomly divided into 5 non-overlapping subsets, with 4 subsets as the training subset and 1 subset as the validation subset each time. The validation of all subsets is completed in 5 cycles. The average accuracy rate of five verifications is taken as the evaluation index for hyperparameter selection. After screening out the optimal combination of hyperparameters, the model is trained with the complete training set. Finally, the generalization performance of the model is evaluated on an independent test set. The experiment is conducted through Python 3.9 for simulation analysis under the operating system equipped with Intel Core i7-12700K CPU, NVIDIA RTX 3090 GPU, 32GB of running memory and Windows 11.

3.2 Analysis of the effect of physiological and emotional signal recognition model

Firstly, the signal processing effect of the proposed model is presented, and the original signal, interference signal, and the signal processed by the model are compared. The test results are shown in Figure 7.

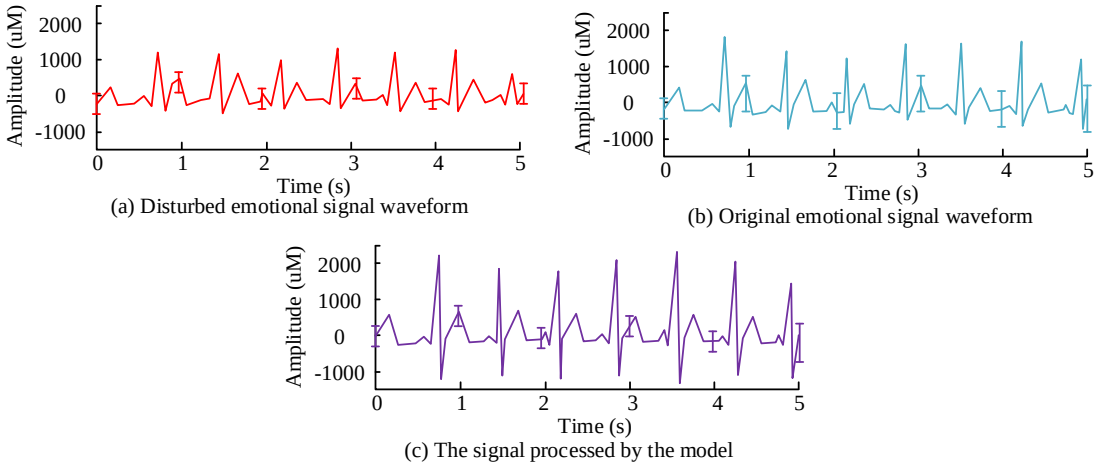

Figure 7: Changes in emotional waveforms after adding the algorithm

From Figure 7 (a), the amplitude of the emotional signal waveform after interference was relatively small, ranging from -1,000uM to 1,000uM. The overall trend of signal waveform changes was more complex, with no good signal changes. From Figure 7 (b), the overall variation range of the original signal waveform ranged from -1,000uM to 2,000uM. The overall fluctuation signal showed more regular fluctuations. From Figure 7 (c), when the model was added for signal anti-interference changes, the overall signal variation ranged from -1,000uM to 2,000uM. The signal waveform presented more fluctuating signal changes. After adding the algorithm, the overall signal change trend is closer to the real fluctuation change, indicating that this model has excellent denoising and recognition capabilities. To test the recognition performance on different emotional expressions, the model is used to test emotional signals. The results are shown in Figure 8.

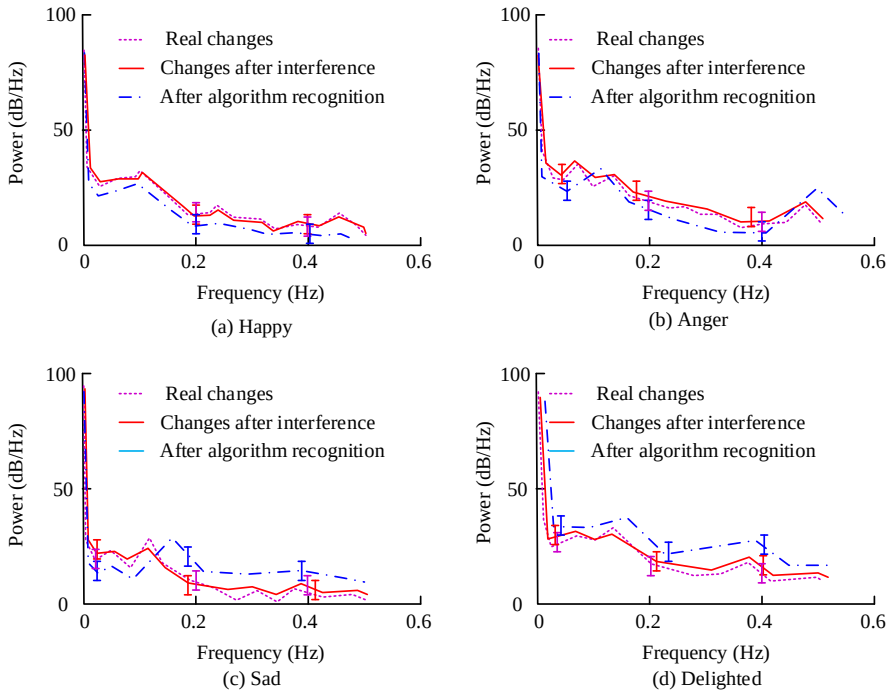

Figure 8: Recognition of four emotional waveforms

From Figure 8, among the four types of emotional changes, the power of emotions showed a sudden decrease followed by a slow decrease as the frequency of emotions increased. Compared with before and after adding the model for recognition, the emotional change in the model was closer to the real curve. It indicates that the model can improve the accurate recognition of emotions. The recognition effect is basically consistent with the real curve. However, in sad emotions, there was a noticeable deviation in the signal. This may be due to slight fluctuations during recognition, resulting in a deviation in the overall signal variation. To test the recognition effect of the algorithm on different emotions, the data information in the dataset is input to train. The

emotion signal waveform is shown in Figure 9.

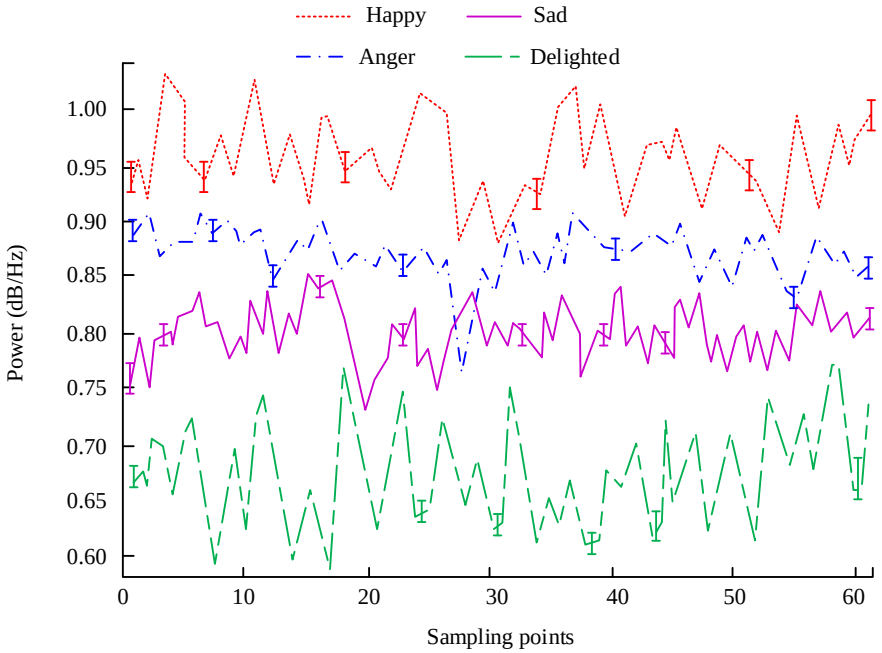

Figure 9: Four different emotion algorithms recognize waveform changes

From Figure 9, after extracting data from emotional signals through the model, the emotional signal recognition were achieved. The happy emotion was relatively high, resulting in a significant power variation in the waveform signal obtained. The delighted emotion signal was relatively soothing, resulting in a smaller power of the emotion signal obtained. From this fluctuation signal image, the current emotion in the waveform could be analyzed. For example, when the signal power was higher than 0.90 dB/Hz, it could be determined as a strong emotional state. When it was lower than 0.70 dB/Hz, it corresponded to a calm emotional state. It had the best analytical effects on judging the emotions of the organism itself. To test the accuracy of feature extraction signals in the current model, different signals are subjected to feature extraction. One-way Analysis of Variance (ANOVA) is used to test the differences in each feature and overall accuracy among different samples, with a significance level of 0.05. If $p<0.05$, it indicates that the ANOVA results show significant differences. The results are shown in Table 3.

Table 3: Recognition accuracy rates of different signal features

Sample	PH	QRS	PRQ	QT	QTC	ST	ALL
Sample 1	60.20±2.1%	54.30±2.5%	58.40±1.9%	76.50±1.8%	75.60±2.2%	81.50±1.6%	91.50±1.2%
Sample 2	61.10±1.8%	62.40±2.0%	64.50±2.3%	56.30±2.4%	45.60±2.8%	64.50±2.1%	92.60±1.1%
Sample 3	58.60±2.3%	59.00±2.2%	54.60±2.0%	90.60±1.5%	82.40±1.7%	58.60±2.4%	91.50±1.3%
Sample 4	51.60±2.5%	58.20±2.1%	61.30±1.8%	93.40±1.2%	71.60±2.0%	78.40±1.9%	92.60±1.0%
Sample 5	59.40±2.0%	57.50±2.3%	58.70±2.1%	68.40±1.6%	76.50±1.8%	72.30±2.2%	94.60±0.9%
F	3.21	4.58	2.97	8.53	7.12	5.34	4.10
p	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

In Table 3, Wave Height (PH), Complex Duration (QRS), Interval (PRQ), Interval (QT), Corrected QT interval (QTC) and ST-segment length (ST) respectively represent the expression of feature values in feature extraction. Only six feature extraction feature value changes are shown. From Table 2, the overall accuracy across all extracted features (ALL) represents the total accuracy of feature extraction for all current signals, while the preceding signals represent a single change in accuracy. Among the five emotional changes, the accuracy of PH was the highest in sample 2, with the highest accuracy of 61.10%±1.8%%. The accuracy of QT was the highest in sample 4, with the highest accuracy of 93.40%±1.2%. The sample with the highest accuracy value appeared differently in different signals, but the overall accuracy was above 90.00%. This indicates that the algorithm can better extract features from signals, with the highest accuracy of 94.60%±0.9%. Compared with sample 1 and sample 3, the accuracy rate was approximately 3.10% higher. Moreover, the results of

ANOVA showed that the differences among different samples in each feature and overall accuracy were statistically significant (p<0.05). This indicates that there are differences in signal feature extraction among different samples, which may be due to individual differences in the samples. Comparative testing was conducted on different emotions. Different emotions are subjected to comparison tests. The recognition accuracy of SVM, RF, and FA is compared, as shown in Figure 10.

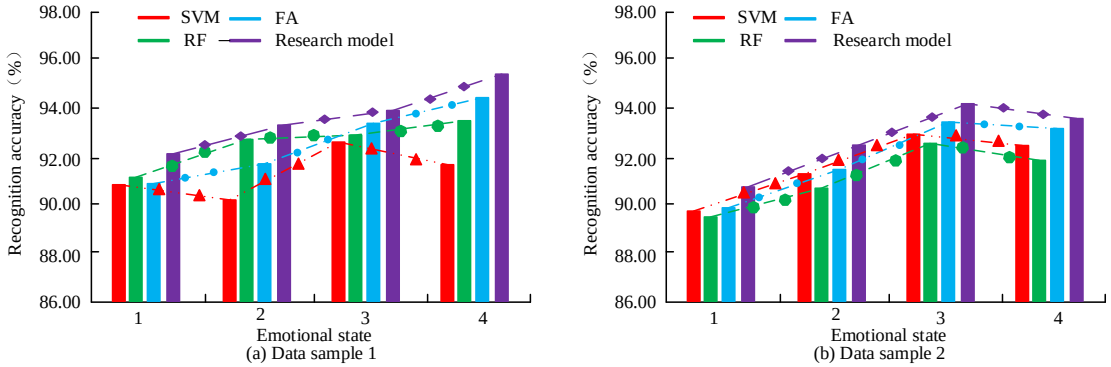

Figure 10: Comparison of sample recognition accuracy of different models

From Figure 10 (a), in the four models, the accuracy of the models used in the study was better. The accuracy of some models decreased when more emotions were added to the recognition. The accuracy of the model used in the study was improved, reaching a maximum of 95.10%, which was 2.89% higher than the lowest of 92.21%. This indicates that the algorithm also has best emotion signal recognition ability after adding new emotions. From Figure 10 (b), in the new sample, the accuracy of different models first increased with the addition of emotions and then decreased. This may be due to differences in emotional signals generated by different individuals. The accuracy recognition rate of the research model was the best, with a maximum value of 95.20%. The minimum value of the RF model was 92.30%, which was about 2.90% higher than the model used in the research. Compared with the RF, the accuracy of the research model was about 2.90% higher. In different sample training recognition, the research algorithm can also maintain better recognition accuracy. To test the effectiveness of different improved models, the F1 values of different models are compared and tested. The results are shown in Figure 11.

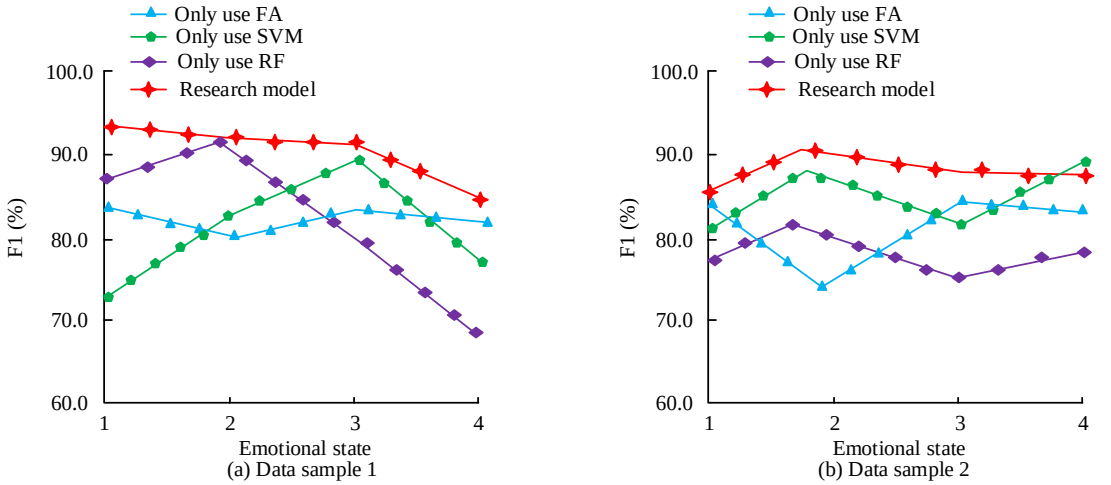

Figure 11: Changes in F1 values with different algorithm improvements

From Figure 11 (a), after adding different models for improvement, the F1 value of the model showed a fluctuating trend when emotions were added to different quantities. However, the F1 value of the model used in the study showed a decreasing trend. The entire curve changed more smoothly, with the lowest value being 84.9%. Compared to other algorithms, the highest value was about 3.6% higher. From Figure 11 (b), the proposed model had higher F1 value. The highest value occurred when the emotion was 2, with a highest value of 90.3%. Compared with the highest value of 88.4% in other improved models, the model used in the study was 1.9% higher. When comparing different improved models, the proposed model has better recognition performance. To comprehensively evaluate the performance of the proposed model in the physiological emotion recognition

task, this study introduces the macro average F1 value, Kappa coefficient, and average AUC-macro. The average F1-macro is calculated separately for each category and then taken as the arithmetic mean, which is used to assess the model's recognition ability in a few categories. The Kappa coefficient is used to measure the consistency between the model's prediction and manual annotation after eliminating random consistency. The average AUC-macro is to take the macro average by category based on the area under the multi-class ROC curve, comprehensively evaluating the model's ability to distinguish various emotional signals. These three index values of the proposed model, the method based on uncertainty graph convolutional network [5], the method based on adaptive neural fuzzy reasoning system [6], the method based on temporal difference minimization neural network [7], and the manual annotation method are calculated, respectively. The results are shown in Table 4.

Table 4: Comparison results of average F1-macro, Kappa coefficients and average AUC-macro of different models

Method	Indicator		
	F1-macro /%	Kappa	AUC-macro
Manual annotation	71.4	0.66	0.78
Reference [5]	72.1	0.67	0.79
Reference [6]	73.8	0.69	0.81
Reference [7]	76.3	0.73	0.84
Research model	77.8	0.75	0.86

From Table 4, the average F1-macro of the model proposed in this study was 77.8%, which was 6.4% higher than that of the manual annotation method, 5.7% higher than that of the method based on uncertainty graph convolutional network, and 4.0% higher than that of the method based on the adaptive neural fuzzy reasoning system. The Kappa coefficient of the proposed model was 0.75, which increased by 0.09, 0.08, 0.06 and 0.02 respectively compared with other methods. The average AUC-macro of the proposed model was 0.86, which was significantly higher than that of several other methods. The above results indicate that the proposed model has good recognition ability, high consistency, and strong discriminative ability for emotional signals, proving its comprehensive performance and stability.

4 Discussion

This study aims to address the low recognition accuracy of physiological emotion signals and their susceptibility to noise interference. An improved machine learning model based on FA is proposed to optimize SVM hyperparameters and integrate it with RF. This machine learning framework realizes multi-dimensional physiological signal feature extraction through time domain and wavelet transform. The results showed that the proposed model had the highest recognition accuracy of 94.60% on the DEAP dataset, which was similar to the findings drawn by Bardak et al [6]. The method based on adaptive neural fuzzy reasoning system proposed by Bardak et al. achieved an accuracy of 73.49% on the same dataset, and the proposed model was significantly better than that of this model. The reason is that the FA has the ability to globally optimize the SVM hyperparameters, which effectively solves the problem that traditional SVM is prone to falling into local optimality in the high-dimensional space of physiological signals. The model achieved an average F1-macro of 77.8%, which was better than that of the method based on temporal difference minimization neural network proposed by Ju et al. [7]. The reason is that the proposed model does not need to rely on prior information about time series. It can fully utilize the time domain and frequency domain features of physiological signals through the classification prediction of FA-SVM and the integrated decision of RF, and can still stably identify samples without obvious time series regularity. The QT interval feature recognition accuracy of the proposed model reached 93.40±1.2%, which was similar to the results obtained by APM et al [12], but the proposed model was better. This is because the proposed model realizes automatic feature extraction by standardizing wavelet parameters, avoiding the subjective bias of manual labeling. In short, the proposed model significantly outperforms existing methods in recognizing physiological emotion signals through algorithm integration optimization and multi-dimensional feature extraction, providing a reference for subsequent research on emotion recognition based on physiological signals.

5 Conclusion

Aiming at the low recognition accuracy of physiological and emotional signals and susceptibility to interference during the recognition process, an improved algorithm based on machine learning is proposed. The research first builds an improved model based on machine learning. Then, the model intelligently recognizes and extracts data features from the current physiological emotional signals. Finally, the performance and feasibility of the model are compared and tested through experiments. Among the five emotional changes, the accuracy of PH was the highest in sample 2, with the highest accuracy of 61.10%. The highest accuracy in QT was 93.40% for sample 3. The overall accuracy was above 90.00%. The

highest accuracy was 94.60%, which was about 3.10% higher than the lowest sample. In the four models, the accuracy of the proposed model was better, with the highest accuracy of 95.10%, which was 2.89% higher than other models. The highest recognition accuracy of the proposed model was 95.20%, which was about 2.90% higher than the RA. Compared with the existing models, the fusion algorithm proposed in this paper demonstrates higher accuracy and stability in the emotion recognition task. Although the research has achieved many results, improvements are still necessary, such as adding more emotions when conducting emotional tests. Therefore, future research will test more emotions. Meanwhile, the data used in the study is relatively small. More and larger data will be analyzed and tested in the future.

Fundings

The research is supported by science and technology planning project of Jilin Province under Grant number 20230203032SF.

References

[1] Sidorova J, Karlsson S, Rosander O, Berthier Marcelo L. Towards Disorder-Independent Automatic Assessment of Emotional Competence in Neurological Patients with a Classical Emotion Recognition System: Application in Foreign Accent Syndrome. IEEE transactions on affective computing, 2021, 12(4):962-973. https://doi.org/10.1109/TAFFC.2019.2908365

[2] Graumann L, Duesenberg M, Metz S,Schulze Lars. Facial emotion recognition in borderline patients is unaffected by acute psychosocial stress. Journal of Psychiatric Research, 2021, 132(3):131-135. https://doi.org/10.1016/j.jpsychires.2020.10.007

[3] Smietanka L, Maka T. Audio Feature Space Analysis for Emotion Recognition from Spoken Sentences. Archives of Acoustics, 2021, 46(2):271-277. https://doi.org/10.24425/aoa.2021.136581

[4] Kelvin Leong, Anna Sung (2023). An Exploratory Study of How Emotion Tone Presented in A Message Influences Artificial Intelligence (AI) Powered Recommendation System. Journal of Technology & Innovation, 3(2): 80-84.

[5] Gao H, Wang X, Chen Z, Wu M, Cai Z, Zhao L, Liu C. Graph convolutional network with connectivity uncertainty for EEG-based emotion recognition. IEEE Journal of Biomedical and Health Informatics, 2024, 28(10): 5917-5928. https://doi.org/10.1109/JBHI.2024.3416944

[6] Bardak F K, Seyman M N, Temurtaş F. Adaptive neuro-fuzzy based hybrid classification model for emotion recognition from EEG signals. Neural Computing and Applications, 2024, 36(16): 9189-9202. https://doi.org/10.1007/s00521-024-09573-6

[7] Ju X, Li M, Tian W, Hu D. EEG-based emotion recognition using a temporal-difference minimizing neural network. Cognitive Neurodynamics, 2024, 18(2): 405-416. https://doi.org/10.1007/s11571-023-10004-w

[8] Hajarolasvadi N, Demirel H. Deep facial emotion recognition in video using eigenframes. IET Image Processing, 2020, 14(14):3536-3546. https://doi.org/10.1049/iet-ipr.2019.1566

[9] Lin Y H, Giang C M. Online communication self-disclosure and intimacy development on Facebook: the perspective of uses and gratifications theory. Online Information Review, 2021, 45(6):1167-1187. https://doi.org/10.1108/OIR-08-2020-0329

[10] Liu C, Tian Q, Chen M. Distinguishing Personality Recognition and Quantification of Emotional Features Based on Users' Information Behavior in social media. Journal of Database Management, 2021, 32(2):1-16. https://doi.org/10.4018/JDM.20210401.oa1

[11] Hariri W, Farah N. Recognition of 3D emotional facial expression based on handcrafted and deep feature combination. Pattern recognition letters, 2021, 148(6):84-91. https://doi.org/10.1016/j.patrec.2021.04.030

[12] A P M, A A P, Séverine Lannoy a b, D'Hondt Fabien. Tackling heterogeneity: Individual variability of emotion decoding deficits in severe alcohol use disorder - ScienceDirect. Journal of Affective Disorders, 2021, 279(2):299-307. https://doi.org/10.1016/j.jad.2020.10.022

[13] Lundine M A, Brothers L L, Trembanis A C. Deep learning for pockmark detection: Implications for quantitative seafloor characterization. Geomorphology, 2023, 421(7):1-20. https://doi.org/10.1016/j.geomorph.2022.108524

[14] Bai Y, Chen W, Ai B,Zhangdui Zhong. Prior Information Aided Deep Learning Method for Grant-Free NOMA in mMTC. IEEE Journal on Selected Areas in Communications, 2022, 40(1):112-126. https://doi.org/10.1109/JSAC.2021.3126071

[15] Bagyaraj S, Tamilselvi R, Gani P B M,Sabarinathan Devanathan. Brain tumour cell segmentation and detection using deep learning networks. IET Image Processing, 2021, 15(10):2363-2371. https://doi.org/10.1049/ipr2.12219

[16] Duan Y. Evaluation of Mental Health Education Using a Northern Goshawk Deep Multi-Structured Convolutional Neural Network for Emotion Recognition. Informatica, 2024, 48(18): 21-34. https://doi.org/10.31449/inf.v48i18.6420

[17] Fang X L, Deshmukh M, Chee M L, Soh Zhi-Da. Deep learning algorithms for automatic detection of pterygium using anterior segment photographs from slit-lamp and hand-held cameras. British Journal of Ophthalmology, 2021, 106(12):1642-1647. https://doi.org/10.1136/bjophthalmol-2021-318866

[18] Fang Y, Wang Y. Cross Modal Sentiment Analysis of Image Text Fusion Based on Bi LSTM and B-CNN.

Informatica, 2024, 48(21): 95-111. https://doi.org/10.31449/inf.v48i21.6767

[19] Cheng J, Yang L, Tong S. Painting style and sentiment recognition using multi-feature fusion and style migration techniques. Informatica, 2024, 48(21): 127-138. https://doi.org/10.31449/inf.v48i21.6891

[20] Sadiqbatcha S, Zhang J, Amrouch H,Sheldon X.-D. Tan.Real-Time Full-Chip Thermal Tracking: A Post-Silicon, Machine Learning Perspective. IEEE Transactions on Computers, 2022, 71(6):1411-1424.

[21] Whitmore B C, Lee J C, Chandar R. Star Cluster Classification in the PHANGS-HST Survey: Comparison between Human and Machine Learning Approaches. Monthly Notices of the Royal Astronomical Society, 2021, 506(4):5294-5317. https://doi.org/10.1093/mnras/stab2087

[22] Dornelas R S, Lima D A. Correlation Filters in Machine Learning Algorithms to Select De-mographic and Individual Features for Autism Spectrum Disorder Diagnosis. Journal of Data Science and Intelligent Systems, 2023, 3(1): 7-9. https://doi.org/10.47852/bonviewJDSIS32021027