



OPEN ACCESS

EDITED BY

Yongcheng Li,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Yongjun Zheng,
University of Hertfordshire, United Kingdom
Fangzhou Xu,
Qilu University of Technology, China

*CORRESPONDENCE

Bangyan Zhou
✉ zby@ahu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 04 September 2023

ACCEPTED 09 October 2023

PUBLISHED 28 November 2023

CITATION

Zhang L, Xiao D, Guo X, Li F, Liang W and Zhou B (2023) Cross-subject emotion EEG signal recognition based on source microstate analysis. *Front. Neurosci.* 17:1288580. doi: 10.3389/fnins.2023.1288580

COPYRIGHT

© 2023 Zhang, Xiao, Guo, Li, Liang and Zhou. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cross-subject emotion EEG signal recognition based on source microstate analysis

Lei Zhang^{1†}, Di Xiao^{1†}, Xiaojing Guo², Fan Li³, Wen Liang⁴ and Bangyan Zhou^{1*}

¹Anhui Province Key Laboratory of Multimodal Cognitive Computation, School of Computer Science and Technology, Anhui University, Hefei, China, ²School of Electronic Information Engineering, Anhui University, Hefei, China, ³Key Laboratory of Civil Aviation Flight Technology and Flight Safety, Civil Aviation Flight University of China, Guanghan, China, ⁴Google Research, California, CA, United States

Electroencephalogram (EEG) signals are very weak and have low spatial resolution, which has led to less satisfactory accuracy in cross-subject EEG-based emotion classification studies. Microstate analyses of EEG sources can be performed to determine the important spatiotemporal characteristics of EEG signals. Such analyses can be used to cluster rapidly changing EEG signals into multiple brain prototype topographies, fully utilizing the spatial information contained in the EEG signals and providing a neural representation for emotional dynamics. To better utilize the spatial information of brain signals, source localization analysis on the EEG signals was first conducted. Then, a microstate analysis on the source-reconstructed EEG signals is conducted to extract the microstate features of the data. We conducted source microstate analysis on the participant data from the odor-video physiological signal database (OVPD-II) dataset. The experimental results show that the source microstate feature topologies of different participants under the same emotion exhibited a high degree of correlation, which was proven by the analysis of microstate feature topographic maps and the comparison of two-dimensional feature visualization maps of the differential entropy (DE) and power spectral density (PSD). The microstate features represent more abstract emotional information and are more robust. The extracted microstate features were then used with the style transfer mapping method to transfer the feature data from the source domain to the target domain and were then used in support vector machines (SVMs) and convolutional neural networks (CNNs) for emotion recognition. The experimental results show that the cross-subject classification accuracies of the microstate features in SVMs were $84.90 \pm 8.24\%$ and $87.43 \pm 7.54\%$, which were 7.19 and 6.95% higher than those obtained with the PSD and 0.51 and 1.79% higher than those obtained with the DE features. In CNN, the average cross-subject classification accuracies of the microstate features were 86.44 and 91.49%, which were 7.71 and 19.41% higher than those obtained with the PSD and 2.7 and 11.76% higher than those obtained with the DE features.

KEYWORDS

EEG, source microstate, style transfer mapping, cross-subject, emotion recognition

Introduction

Emotional theory was proposed as early as the nineteenth century (Darwin, 1872; James, 1894), and it received enthusiastic discussion and research at that time. It also provided more inspiration and guidance for people's current research on emotional classification (Scherer, 2005). As a key factor reflecting human behavior, emotions have a direct internal

connection with the way people communicate with each other. By understanding others' current emotional states, people can choose the best way to communicate to improve their communication process. The expression of emotions is a response that people make to external stimuli, and it is fixed in different areas of the brain according to the unique reactions triggered by different stimuli (Rached and Perkusich, 2013). In particular, emotional states can be mainly classified into two categories. One is the discrete model, where the emotion is thought of as discrete states such as happiness and sadness. The other is the dimensional model, which considers the emotion states to be continuous, such as the valence-arousal metric (Wu et al., 2023).

An essential goal of EEG-based emotion recognition studies is to decode EEG signals and extract more discriminating features for classifying different emotional states more effectively. Based on this goal, a vast number of studies have been conducted. Duan et al. (2013) proposed the use of differential entropy (DE) features on symmetric electrodes for the emotional recognition of subjects' electroencephalogram (EEG) signals, achieving good classification results. Padhmashree and Bhattacharyya (2022) used multivariate variational mode decomposition and multivariate modulated oscillation methods to study the emotional recognition of arousal emotions, dominant emotions, and valence emotions based on instantaneous amplitudes and instantaneous frequencies. Jie et al. (2014) used sample entropy features to distinguish positive and negative emotions in a high arousal state and subjects' emotions in different arousal states. In addition, both traditional machine learning models, such as support vector machines (Wu et al., 2023), and deep learning models, such as deep belief networks (DBNs) (Zheng and Lu, 2015) and long-short term memory (LSTM) (Wu et al., 2022), have achieved great performance for EEG-based emotion recognition. In these studies, satisfactory classification results have been achieved, indicating that EEG signals have good application prospects in the field of emotional recognition. However, the differential entropy, amplitude, and frequency features used in the above studies only contain the temporal information of EEG signals, and the spatial information of multi-channel EEG signals are not provided.

The microstate depicts the topographical map of multiple channel arrays of scalp potential, which can simultaneously result in the signals being recorded in all regions of the cerebral cortex and the topological information of the brain being fully used. In the 1980s, Lehmann et al. (1987) found that the time series of spontaneous EEG signals in the alpha frequency band maintained stability between 80 and 120 ms and that the time series of the scalp potential map suddenly changed to a new state and maintained stability again in this state. This change in the scalp potential field of the brain can reflect the momentary state of the overall activity of the underlying brain network well, and the different topographical map configurations of the brain reflect the changes in the overall and coordinated activity of the brain. These stable periods can reflect the basic steps of human brain information processing.

Li et al. (2021a) used microstate analysis and geoelectrophysiological source imaging methods to study severe depression and found that microstate statistical features have good classification accuracy in identifying severe depression. Strik et al. (1997) studied the microstate of resting electroencephalograms

in Alzheimer's disease and found that the duration of continuous microstates was shortened, providing a direction for the early diagnosis of Alzheimer's disease. Chen et al. (2021) used microstate analysis for emotion recognition research and, based on microstate statistical features, achieved ideal results in valence and arousal accuracy in the recognition of emotions from EEG signals. Microstate features can effectively improve the emotional recognition performance of EEG signals. Shen et al. (2020) used the microstate method of EEG to potentially characterize emotional experiences and determined that the microstate statistical features achieved better classification results in arousal and negativity in the DEAP dataset.

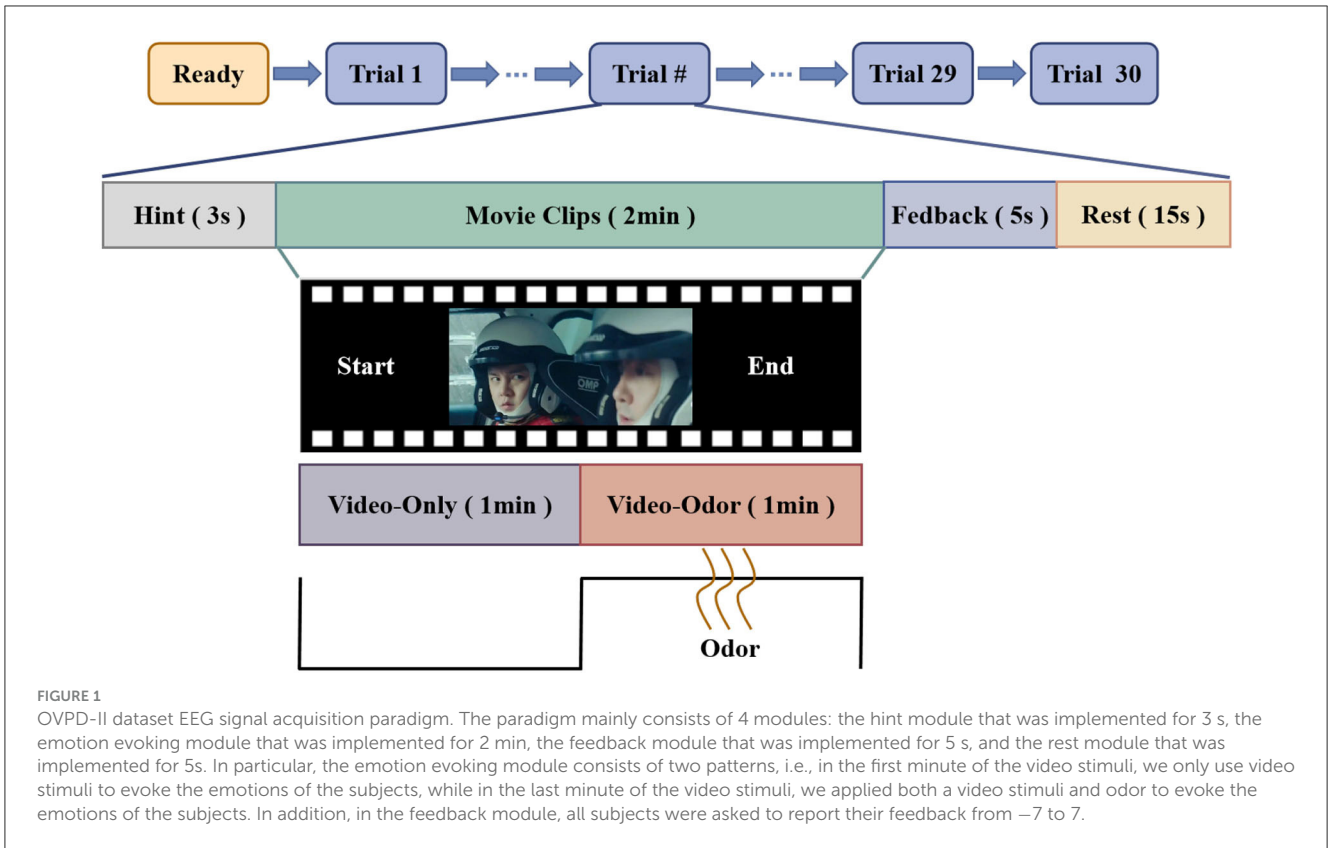
Currently, most studies on EEG-based emotion recognition are conducted within subjects. Due to differences in the education level, living environment, genetic factors, and the non-stationarity of EEG signals (Graimann et al., 2010), the performance of EEG-based emotion recognition tasks significantly decreases when applied across subjects. Transfer learning methods proposed by researchers in 1995 (Pan and Yang, 2010) can improve the accuracy of EEG signal-based emotion classification across subjects by reducing the differences between EEG signals from different subjects and weakening the interference of signals during data collection. Transfer learning is mainly divided into sample-based (Li et al., 2022; Zhao et al., 2023), model-based (Wang and Li, 2022; Zhao et al., 2022), and feature-based (Li et al., 2021b; Khalil et al., 2022) methods. Feature-based transfer learning methods minimize the distance between the source and target domains (Dai et al., 2007), enabling the model trained on the source domain data to perform well in the target domain, thereby improving the accuracy of cross-subject emotion classification.

During the experiment, source localization analysis was first conducted on the OVPD-II preprocessed data, and source microstate analysis was then performed on the traced data after the global field power was calculated. Four statistical features of the source microstate prototype coverage, duration, occurrence rate, and microstate sequence transition probability were calculated for the participants. In this study, source microstate features were applied to cross-subject EEG-based emotion recognition research for the first time. The calculated source microstate statistical features were put into the style transfer mapping method for feature transfer, and the transferred feature data were put into a support vector machine (SVM) and convolutional neural network (CNN) for cross-subject emotion classification.

Materials and methods

Introduction to the dataset

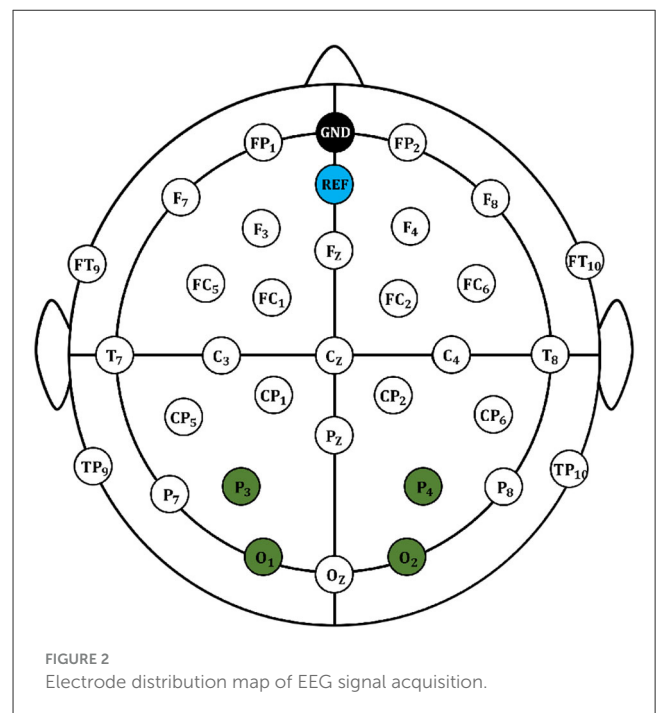
As an important human sense, olfaction constantly affects people's daily lives. OVPD-II is a physiological signal database based on olfactory odors and video induction, and it was developed from the OVPD dataset (Xue et al., 2022) and other datasets from Anhui University. It includes two sub-datasets: video stimulation data and video with odor stimulation data. Thirty representative Chinese film clips were selected as video stimuli from 70 films for the experiment. At the same time, 10 common odors that can arouse the emotions of the participants and are consistent with the



video were chosen as olfactory stimuli. These odors include four that can evoke positive emotions (rose, orange, lavender, and toilet water), four that can generate negative emotions (a cleaning agent, alcohol, vinegar, and ink), and two neutral odors that have little effect on emotions (odorless pure water and air).

As there are individual differences in the emotional tendencies toward the same odor, each participant underwent two sets of repeated odor tests before the formal experiment started. Finally, the odors that can induce the same emotion as that obtained from the video stimulus are selected based on the emotions induced by different odors on the participants and the emotions aroused by the video. Each experiment starts with a 3-s prompt, and the participants need to watch a 2-min video clip. After watching the video for 1 min, the experimenter placed an odor related to the content of the clip approximately 2 cm below the participant's nostrils for 1 min. To minimize the influence of the previous odor on the emotions, there was a transition time of 1.5 s between two adjacent odor stimuli. After watching the video, there was a feedback time of 5 s and a rest time of 15 s. The experimental paradigm is shown in Figure 1.

Thirteen mentally and physically healthy students aged between 18 and 27, without visual, auditory, or olfactory impairments, participated in the experiment. These students included 7 males and 6 females. Moreover, a 32-channel electrode distribution, in which four electrode channels (O1, O2, the P3, and P4) were used to record the participants' eye signals, were utilized in the experiment, and the remaining 28 channels were used to record the participants' EEG signals, as shown in Figure 2. The EEG data were collected at a frequency of 250 Hz, and after watching the video, the participants needed to evaluate their emotions based on their own feelings. The



evaluation criteria were marked from 3 to 7 for positive emotions, from -2 to 2 for neutral emotions, and from -3 to -7 for negative emotions. Then, the obtained two-dimensional emotional labels were mapped to a discrete emotional model of positive, neutral, and negative emotions.

Data preprocessing

During the experiment, the OVPD-II dataset was recorded shown in Figure 3A, and preprocessed before the microstate features and EEG source localization were extracted. First, the four electrode channels used for collecting eye signals were removed, and only the data from the 28 channels used for collecting EEG signals were used. Then, bad electrodes were manually interpolated. The EEGLAB toolbox in MATLAB was used for preprocessing (Delorme and Makeig, 2004). First, a notch filter with a frequency of 49–51 Hz was used to remove power line interference. Then, a bandpass filter was used to retain the EEG data in the frequency range of 0.05–47 Hz. The data after filtering were re-referenced, and independent component analysis was used to remove eye movement artifacts from the EEG signals. Finally, the denoised EEG data were divided into five frequency bands: delta, theta, alpha, beta, and gamma, as shown in Figure 3B.

Source microstate analysis

As shown in Figure 3C, the source microstate analysis was employed to extract the microstate information. The EEG source analysis method measures the potential at different locations on the brain to simulate the electric currents generated by the neuronal activity in the brain. It calculates the optimal current source in the brain that best fits these EEG data, and the activity during the fitting period can be approximated well by a current dipole (Helmholtz, 1853). The cortical surface of the brain is divided into a very detailed grid as the potential location of sources in EEG source analysis methods. The brain source signal is calculated by considering the direction and magnitude of the source of all grid points perpendicular to the cortical surface (Michel et al., 2004). In this experiment, a distributed source localization method was used (Jun et al., 2019). It is assumed that there are m sensors and n samples in the EEG signal. Moreover, $A \in R^{m \times n}$ represents the measured scalp signal, d represents the number of dipoles in the model, $B \in R^{m \times d}$ represents the lead field matrix in the model, $\theta \in R^{d \times n}$ represents the dynamic sources of d dipoles and n EEG samples, and ω_1 represents the random fluctuations in the sensor space. The specific algorithm for the inverse problem of distributed source localization is as follows:

$$A = B\theta + \omega_1 \quad (1)$$

Standardized low-resolution brain electromagnetic tomography (sLORETA) uses the collected EEG data to achieve EEG source localization analysis and create a standardized low-resolution brain source model for further study of the brain's processing of signal stimulation (Scharmüller et al., 2012). Assuming X represents the Laplacian operator in discrete space, $B \in R^{m \times d}$ represents the lead field matrix in the model, W represents the weighting matrix, and d represents the number of dipoles in the model, the specific calculation method of W is:

$$W = X \times \text{diag}(\|B_1\|, \|B_2\|, \dots, \|B_d\|) \quad (2)$$

The sLORETA method further considers the postprocessing of the standardized current density power. Assuming $\theta \in R^{d \times n}$

represents the dynamic source of d dipoles and n EEG samples, D_θ represents the variance of the estimated current density; then, the specific algorithm for $\theta \in R^{d \times n}$ and D_θ is:

$$\theta'_i = \theta_i^T \{[D_\theta]_{ii}\}^{-1} \theta_i \quad (3)$$

$$D_\theta = W^{-1}A^T(AW^{-1}A^T)^+ A \quad (4)$$

In microstate analysis, the electric field topography results at the local maximum of the global field power (GFP) (Koenig et al., 2002) curve are considered the discrete states of the EEG signals (Brunet et al., 2011). Assuming C is the number of electrodes in the EEG signal, 28 channels are used in the OVPD-II dataset to record the EEG signals of participants. $V(i, t)$ represents the instantaneous potential of the i -th electrode at time t , and $\bar{V}(t)$ represents the mean value of the potential of all electrodes at time t . The calculation method of global field power is as follows:

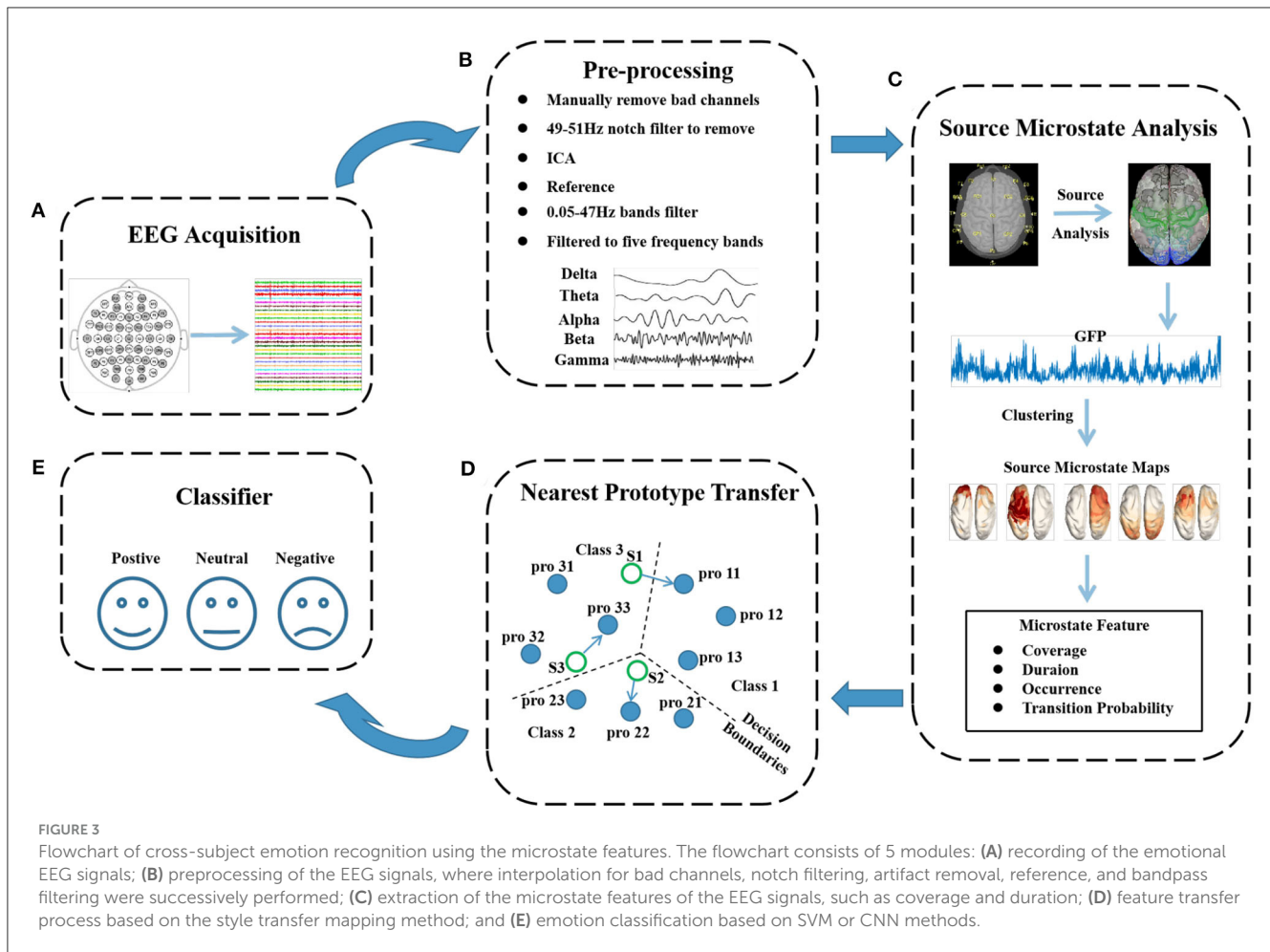
$$GFP = \sqrt{\frac{\sum_i^k (V_i(t) - \bar{V}(t))^2}{C}} \quad (5)$$

The K-means clustering algorithm (Arthur and Vassilvitskii, 2007) uses the peak values of the GFP of the EEG signals as the original topography maps for clustering. Then, several scalp topography maps are randomly selected from all original topography maps as template maps for microstate classes. The spatial correlation between each template topography map and the original topography map of the EEG data is calculated. Based on the calculated spatial correlation sequence, the global explained variance (GEV) of each microstate class topography map is calculated to measure the explained variance of the selected microstate topography maps for the entire EEG data (Khanna et al., 2015). Assuming μ_1, μ_2 to μ_n represent the mapping of n clustering templates, where μ_j represents the j -th clustering template, $x^{(i)}$ is the original topographic map of EEG data, and $l^{(i)}$ is the label of the nearest microstate template to the original topographic map. The calculation method is as follows:

$$l^{(i)} := \text{argmin}_j \|X^i - \mu_j\|^2 \quad (6)$$

$$\mu_j := \frac{\sum_{i=1}^n \{l^i = j\} X^i}{\sum_{i=1}^n \{l^{(i)=j}\}} \quad (7)$$

For the selection of the optimal number of microstate clusters, the GEV and cross-validation (CV) criteria (Pascual-Marqui et al., 1995) are used in this study to select an appropriate number of microstate clusters. The GEV is used to calculate the correlation between the EEG signals and the microstate prototypes at a certain time, quantifying the degree of match between the EEG data and the assigned microstates. The higher the GEV value is, the better the fit between the microstates and the data. The CV criterion is used to measure the residual noise between the EEG signals and the microstate topography maps. The smaller the value is, the more the clustering of the microstate numbers can explain the EEG signals



(Murray et al., 2008). Assuming x_i is the EEG signal at time i , p_{li} represents the label of the microstate topographic map of the $i - th$ EEG sample, GFP_i represents the standard deviation of all electrodes in the $i - th$ EEG sample, where C represents the number of EEG channels, K represents the number of microstates, and $\hat{\sigma}^2$ represents the variance of the residual noise. The specific calculation methods of the GEV and CV criterion are as follows:

$$GEV_i = \frac{(Corr(X_i, p_{li}) \bullet GFP_i)^2}{\sum_i^{N} GFP_i^2} \tag{8}$$

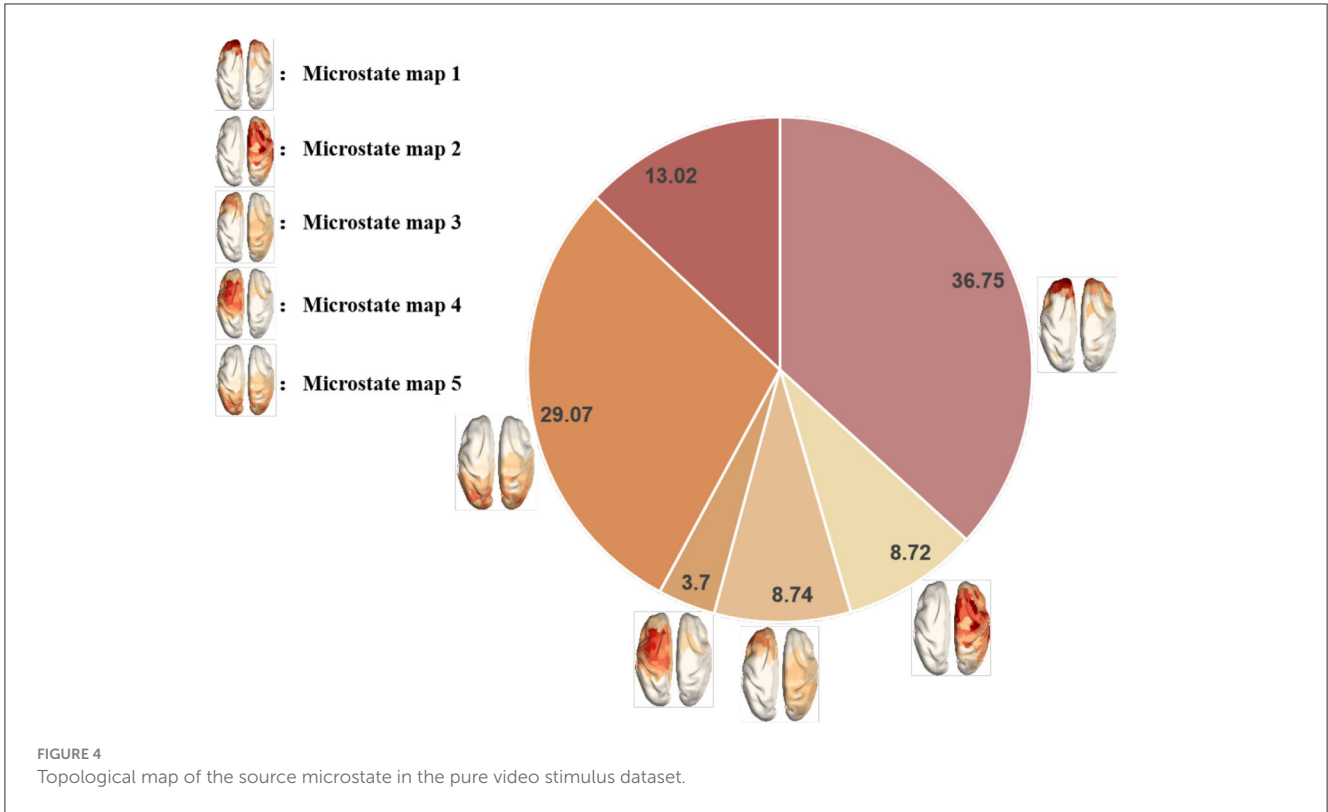
$$CV = \hat{\sigma}^2 \bullet \left(\frac{C - 1}{C - K - 1} \right)^2, \hat{\sigma}^2 = \frac{\sum_i^N X_i^T X_i - (a_i^T X_i)^2}{N(C - 1)} \tag{9}$$

After an appropriate source microstate prototype is selected, the source microstate topographic map is inputted back to the original data, and the statistical features of the source microstate, including the time coverage, duration, occurrence, and transition probabilities between the source microstates, are calculated. The time coverage of the source microstate template refers to the proportion of the source microstate topographic map active in the entire analyzed EEG data. The duration refers to the average

length of each source microstate topographic map in milliseconds of EEG data sequences. The occurrence of the source microstate represents the average frequency of occurrence of a source microstate class per second. It is calculated by dividing the number of segments belonging to a source microstate class by the total duration of the analysis data in seconds. The transition probabilities refer to the transition probability between any two given source microstate templates.

Style transfer mapping

As shown in Figure 3D, the style transfer mapping is introduced to map the target domain EEG features into source domain to narrow the difference of data distribution between different domains. In transfer learning, we use the training set data as the source domain and the test set data as the target domain. The style transfer mapping method maps the data from the source domain to the target domain through affine mapping, reducing the distance between the source and target domains (Zhang and Liu, 2012). This causes the classification model to be more familiar with the target domain data, leading to better classification results. It is assumed that the target domain data are represented as $X = \{x_i \in R^M | i = 1, 2, 3, \dots, N\}$ and that the source domain data



are represented as $Y = \{y_i \in R^D | i = 1, 2, 3, \dots, N\}$, where M represents the feature dimension and N represents the number of data points in the target domain. Assuming x_i is mapped to y_i under the condition that the confidence level is set to $f_i \in [0, 1]$, concept drift may occur after mapping. The style transfer mapping transforms y_i back to x_i through the inverse function $Ay_i + b$ and learns the transfer matrix x_i by minimizing the weighted square error. The specific calculation method is as follows:

$$\min_{A \in R^{M \times M}, b \in R^M} \sum_{i=1}^N f_i \|Ay_i - x_i\|_2^2 + \beta \|A - I\|_F^2 \quad (10)$$

β is used to balance the degree of data transfer and non-transfer to avoid excessive transfer. A value of β that is too small may result in excessive transfer, while a larger value may not be conducive to transfer. The calculation method of β is:

$$\beta = \tilde{\beta} Tra(f_i y_i y_i^T) \quad (11)$$

After cross-validation, the value of $\tilde{\beta}$ can be chosen from the range of [1,3]. During the style transfer mapping, the nearest prototype model (Bezdek and Kuncheva, 2001) is used for destination mapping. The source domain data are clustered using the K-means clustering algorithm, and the cluster centers serve as the prototypes required for the experiment. The specific calculation method is as follows:

$$p_{ij} = \epsilon R^M, j = i, \dots, n_i, i = 1, \dots, M \quad (12)$$

Experimental settings

In this experiment, the source microstates of the preprocessed OVPD-II dataset were first analyzed. We first calculated the GFP of the five frequency bands of the EEG data for 13 subjects in the dataset. Then, we used the K-means clustering method to cluster the GFP for source microstate clustering and selected an appropriate number of source microstate clusters based on the GEV and CV criteria. Through calculation, we finally selected 5 source microstate topologies that can explain approximately 85% of the EEG data information in the OVPD-II dataset, as shown in Figure 4, which is the source microstate topology of the OVPD-II dataset in the alpha band. We fitted the EEG data based on the selected source microstate prototypes and finally smoothed the fitted data over time. Then, we calculated the statistical features of the 5 source microstates, including the time coverage, duration, occurrence, and transition probabilities. The dimensions of the coverage range, duration, and occurrence frequency are consistent with the number of microstate categories, all of which are row vectors of size 1×5 . The dimension of the transition probability is a data matrix of size 5×5 . Therefore, the feature dimension of a single frequency band sample is $5 \times (5 + 3) = 40$ dimensions. The EEG data features of the 5 frequency bands are fused, and the feature dimension after frequency band fusion is 200 dimensions.

The OVPD-II dataset contains EEG data from 13 participants. We employed a multi-source domain style transfer mapping method (Li et al., 2019), using 12 subjects in the dataset as the source domain and the remaining subject as the target domain. The 12 subjects were trained using SVM and CNN (shown in Figure 3E), and the trained model was tested using verification

TABLE 1 Setting of the hyperparameters in the convolutional neural networks.

Hyper parameter	Values
Learning rate	0.005
Batch size	16
Learning decay coefficient	0.8
Regularization	0.001
Epoch	100

data. The best performing model was selected, and the source domain data used to train the model were identified. The target domain data and source domain data were migrated using the style transfer mapping method, and the migrated data were classified using a classifier. The experimental confidence level f_i value was determined to be 0.8, and β was determined to be 0.2 through cross-validation analysis.

For the parameter setting of a classification model, choosing the linear kernel function for the SVM classifier can achieve the highest classification accuracy. Moreover, the grid search method from the parameter pool, $P_{value} = \{0.01k, 0.1k, k | k = 1, 2, \dots, 9\}$, was adopted to choose the optimal regulation parameter C and the kernel function hyperparameter γ . For the CNN experiment, a single-layer CNN was used, with the hyperparameters listed in Table 1. The experimental environment was built on a Windows 10 PC with Core (TM) i7-8700 CPU and 16 GB memory, and the computing environment was MATLAB 2019b.

Results and discussion

Analysis of the source microstate features

In the experiment, non-overlapping 1-s sliding windows were used to extract the source microstate features from the subjects in the OVPD-II dataset, and the topological maps of the source microstates were analyzed. By calculating the GFP values of each subject under each emotion, the source microstate prototypes that could explain approximately 70% of the EEG information were selected as representative topological maps for each subject. We calculated the common active areas of the representative microstate topological maps of 13 subjects under three emotions and drew a typical source microstate prototype that could represent the dataset. The Pearson correlation coefficients between the 13 subjects and the source microstate prototype were calculated, as shown in Figure 5, where “V” represents the video stimuli and “OV” represents odor with video stimuli. The results of the experiment show that the topological maps of the source microstates extracted from different subjects under the same emotion had high correlation. We believe that this is because the microstates reflect the overall coordinated changes in brain activity. Regarding the clustering of the source microstates of all 13 subjects in the dataset, the selected 5 topological maps could explain 85% of the EEG data, and the differences in the active areas of different subjects were ignored in the selected microstate

templates under the same emotion. Therefore, the statistical features of the source microstates from different subjects have high robustness.

Under video stimulation, the active regions of the positive emotion source microstate topography were mainly concentrated in the frontal and parietal lobes, those of the neutral emotion were concentrated in the frontal, parietal, and occipital lobes, and those of the negative emotion were concentrated in the frontal and parietal lobes. After adding odor stimulation that could induce the same emotion in the subjects, the activity level of the source microstate in the frontal lobe increased for positive and negative emotions, while the odor stimulation for the neutral emotion was colorless and tasteless (pure water and air), and the active region of the source microstate topography did not show a significant enhancement. This indicates that adding odor stimulation can induce stronger emotional responses in the subjects, and this result is consistent with the experimental conclusion (Xue et al., 2022).

Non-transfer cross-subject emotion recognition

Moreover, the DE and power spectral density (PSD) features of non-overlapping 1-s sliding windows are calculated in this experiment, where the DE features are calculated as:

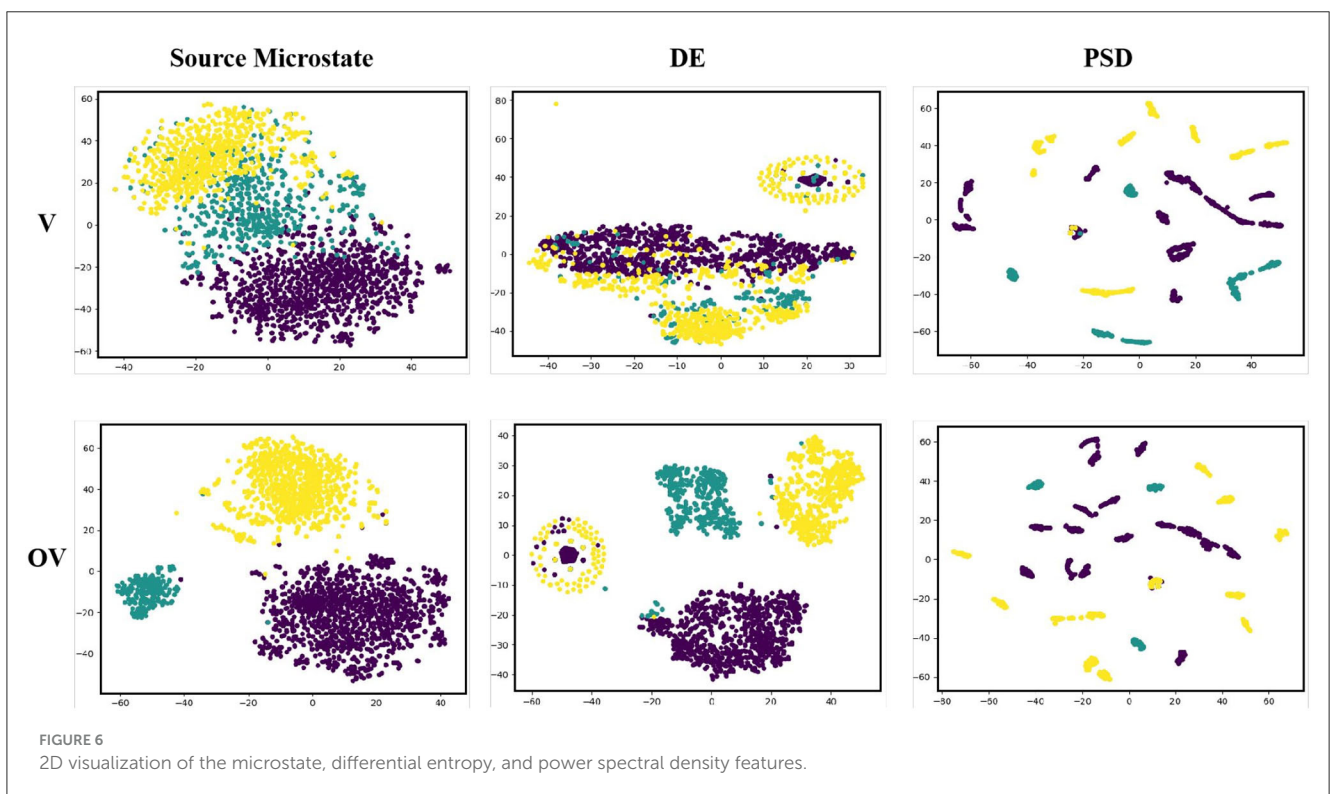
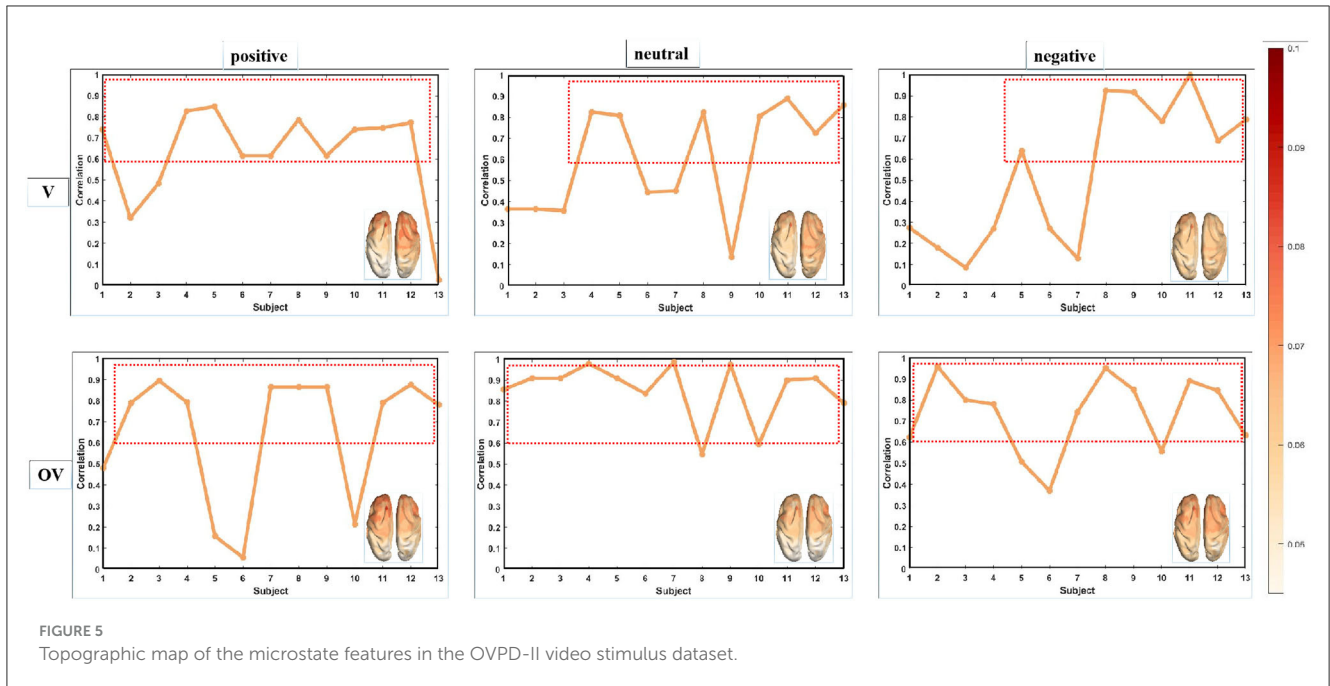
$$h(x) = - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx \quad (13)$$

$$= \frac{1}{2} \log(2\pi e\sigma^2)$$

The time series X follows a Gaussian distribution $N(\mu, \sigma^2)$.

The t-distributed stochastic neighbor embedding (t-SNE) function was used to reduce the dimensionality of the data features and draw a two-dimensional visualization feature map. Figure 6 shows the feature visualization map of Subject HH, where the purple dots represent positive emotions, the yellow dots represent neutral emotions, and the green dots represent negative emotions. The results of the experiment show that the two-dimensional visualization feature map of the source microstate was different from those of DE and PSD. The clustering distribution boundaries of the features obtained were more obvious. In addition, it was observed that there was more obvious clustering of the positive and neutral emotion samples in DE and clustering of the positive, neutral, and negative emotion samples in PSD, which may be due to the difference in the induced materials. This phenomenon did not occur in the source microstate features we extracted, and we believe that this feature can capture more abstract emotional information compared to DE and PSD.

Before performing data transfer, we separately input the extracted three types of features into SVM and CNN for emotion classification. Table 2 shows the cross-subject classification results of the source microstate features, DE and PSD before style transfer mapping under pure video stimulation and video-odor stimulation. In SVM, the average classification accuracy of the source microstate was the best, with 60.95 and 64.79%, respectively, which were 4.06



and 6.14% higher than the accuracy of DE and 3.84 and 5.83% higher than that of PSD. In CNN, the average classification accuracy of the source microstate was 65.09 and 66.39%, respectively, which was 9.59 and 11.73% higher than that of DE and 2.99 and 4.08% higher than that of PSD. The experimental results show that cross-subject emotion recognition based on source microstate features achieved good classification results in both SVM and CNN.

Cross-subject emotion recognition based on style transfer mapping

We used the extracted source microstate features, DE and PSD to perform nearest prototype feature transfer mapping. The transferred features were separately input into SVM and CNN for cross-subject emotion classification under video stimuli. The recognition accuracy of the three features is shown in Table 3. The

experimental results show that the source microstate features had the highest recognition accuracies in cross-subject classification in both SVM and CNN, with average recognition accuracies of 84.90 and 86.44% for 13 subjects, which were 1.28 and 2.7% higher than those obtained with the DE, and 7.19 and 7.13% higher than those obtained with the PSD.

Figure 7 shows the cross-subject recognition accuracy of the three features under odor with video stimuli. In SVM and CNN, the recognition accuracies of the source microstate were 87.43 and 91.49%, respectively, which were 1.79 and 18.77% higher than those obtained with the DE and 6.95 and 19.82% higher than those obtained with the PSD. The experimental results show that the source microstate achieved ideal classification results in both cross-subject emotion recognition under video and odor with video stimuli. Moreover, the cross-subject recognition accuracy based on source microstate features under odor plus video stimuli was 2.53 and 5.05% higher than that under video stimuli in SVM and CNN, respectively, indicating that the addition of odor stimuli induced stronger emotions in the subjects.

TABLE 2 Cross-subject classification accuracy without transfer learning (%).

Feature	V		OV	
	SVM	CNN	SVM	CNN
Source microstate	60.95	65.09	64.79	66.39
DE	56.89	55.50	58.65	54.66
PSD	57.11	62.10	58.96	62.31

The bold values mean the highest of each column.

TABLE 3 Cross-subject recognition accuracy based on source microstate features, differential entropy, and power spectral density under video stimuli (%).

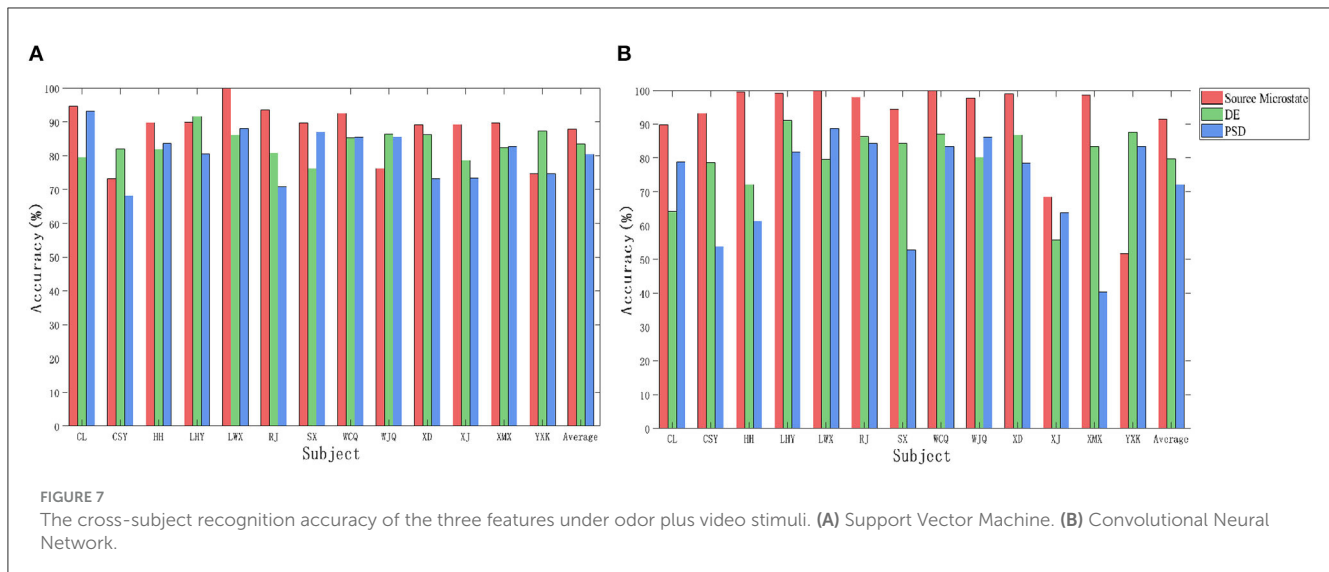
Participant	SVM			CNN		
	Source microstate	DE	PSD	Source microstate	DE	PSD
CL	93.44	88.76	85.07	98.83	92.96	94.37
CSY	84.22	81.65	60.65	88.68	83.31	72.50
HH	82.15	83.78	78.96	77.18	86.57	83.80
LHY	87.25	82.69	81.34	53.85	88.69	83.80
LWX	89.90	80.05	88.81	99.39	88.45	86.62
RJ	89.65	93.08	81.34	63.94	81.925	83.10
SX	88.33	83.48	91.79	98.26	82.96	92.25
WCQ	92.36	84.86	81.25	92.47	84.32	87.04
WJQ	62.53	73.49	60.30	97.18	84.37	88.73
XD	85.08	86.47	84.78	91.94	59.58	38.06
XJ	83.43	76.71	63.48	87.59	85.185	63.889
XXM	89.82	85.86	86.49	97.922	83.59	62.47
YXK	75.60	86.23	65.94	76.48	86.71	94.37
Mean	84.90	83.62	77.71	86.44	83.74	79.31
Std	8.24	5.04	11.10	14.53	7.85	16.19

The bold values mean the highest of each column.

The experimental results show that the combination of the microstate features and style transfer mapping achieved promising results in cross-subject emotion classification. After feature transfer, the cross-subject emotion classification accuracy significantly improved compared to that before transfer. Specifically, the EEG data with video stimuli increased by 23.95 and 21.35% in SVM and CNN, respectively. The cross-subject emotion classification accuracy with video and odor stimuli increased by 22.64 and 25.10% in the two classifiers, respectively. This indicates that the combination of style transfer mapping and source microstate features can be effectively applied to cross-subject emotion recognition research based on EEG signals.

Conclusions and future works

Based on the problem of low recognition accuracy in cross-subject emotion recognition using EEG signals, the robust source microstate and style transfer mapping methods are combined in this study, and they are applied to the odor-video-induced physiological signal database. The source microstate reflects the overall coordinated changes in brain activity. When conducting source microstate analysis based on the dataset, we ignored the differences in the active areas of different subjects under the same emotion, resulting in a high correlation between the source microstate topological maps of different subjects and certain robustness in the statistical features of source microstates from different subjects. Furthermore, the two-dimensional visualization feature map of the source microstate, in contrast to the feature maps of DE and PSD, has more obvious distribution boundaries and can characterize more abstract emotional information than DE



and PSD, showing significant advantages in cross-subject emotion recognition research.

After combining the source microstate feature and style transfer mapping method, the cross-subject emotion recognition accuracy was significantly improved. The recognition accuracies of the source microstate after style transfer in pure video were 84.90 and 86.44% in SVM and CNN, respectively, which were 23.95 and 21.35% higher than the cross-subject recognition accuracies before transfer. The recognition accuracies of the source microstate after style transfer in the odor plus video were 87.43 and 91.49% in SVM and CNN, respectively, which were 22.64 and 25.10% higher than the cross-subject recognition accuracies before transfer. Adding odor stimulation increased the emotion recognition accuracies of the subjects by 2.53 and 5.05% compared to video stimulation, indicating that stronger emotions were induced in the subjects with the addition of odor stimulation.

In the current work, we focus on the emotion recognition of EEG single modalities, while other modalities, such as facial expressions, speech signals, and eye-tracking signals, also contain rich emotional information. OVPD-II is a physiological signal database for emotion recognition based on odor video, in which 4 channels record eye-tracking signals during emotional induction. In future work, we will attempt to fuse eye-tracking signal features and source microstate features for multi-modal cross-subject emotion recognition research. However, brain connectivity is a common and effective analysis method for EEG signals, and we will study the correlation and difference between the brain connectivity functions and the EEG microstates for emotion recognition tasks. We will also test the effectiveness of the EEG microstate features in other public datasets by combining deep learning methods.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Biomedical Ethics Committee, Anhui University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LZ: Methodology, Software, Writing—review and editing. DX: Data curation, Formal analysis, Software, Writing—original draft. XG: Data curation, Methodology, Supervision, Writing—review and editing. FL: Methodology, Supervision, Writing—review and editing. WL: Investigation, Supervision, Writing—review and editing. BZ: Methodology, Software, Supervision, Validation, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by grants from the National Natural Science Foundation of China (NSFC) (No. 61972437), Excellent Youth Foundation of Anhui Scientific Committee (No. 2208085J05), Open Fund of Key Laboratory of Flight Techniques and Flight Safety, CAAC (No. FZ2022KF15), Special Fund for Key Program of Science and Technology of Anhui Province (No. 202203a07020008), and Cloud Ginger XR-1 platform.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Arthur, D., and Vassilvitskii, S. (2007). "K-means++ the advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 1027–1035.
- Bezdek, J. C., and Kuncheva, L. I. (2001). Nearest prototype classifier designs: an experimental study. *Int. J. Intell. Syst.* 16, 1445–1473. doi: 10.1002/int.1068
- Brunet, D., Murray, M. M., and Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: cartool. *Comput. Intell. Neurosci.* 2011, 1–15. doi: 10.1155/2011/813870
- Chen, J., Li, H., Ma, L., Bo, H., Soong, F., and Shi, Y. (2021). Dual-threshold-based microstate analysis on characterizing temporal dynamics of affective process and emotion recognition from EEG signals. *Front. Neurosci.* 15, 689791. doi: 10.3389/fnins.2021.689791
- Dai, W., Xue, G.-R., Yang, Q., and Yu, Y. (2007). "Co-clustering based classification for out-of-domain documents," in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 210–219.
- Darwin, C. (1872). *The Descent of Man, and Selection in Relation to Sex*, Vol. 2. D. Appleton.
- Delorme, A., and Makeig, S. (2004). EEGLab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Duan, R.-N., Zhu, J.-Y., and Lu, B.-L. (2013). "Differential entropy feature for EEG-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE), 81–84.
- Graimann, B., Allison, B. Z., and Pfurtscheller, G. (2010). *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*. Springer Science & Business Media.
- Helmholtz, H. v. (1853). Ueber einige gesetze der vertheilung elektrischer ströme in körperlichen leitern, mit anwendung auf die thierisch-elektrischen versuche (schluss.). *Annalen der Physik* 165, 353–377.
- James, W. (1894). Discussion: the physical basis of emotion. *Psychol. Rev.* 1, 516.
- Jie, X., Cao, R., and Li, L. (2014). Emotion recognition based on the sample entropy of EEG. *Biomed. Mater. Eng.* 24, 1185–1192. doi: 10.3233/BME-130919
- Jun, Y.-H., Eom, T.-H., Kim, Y.-H., Chung, S.-Y., Lee, I.-G., and Kim, J.-M. (2019). Source localization of epileptiform discharges in childhood absence epilepsy using a distributed source model: a standardized, low-resolution, brain electromagnetic tomography (Sloreta) study. *Neurol. Sci.* 40, 993–1000. doi: 10.1007/s10072-019-03751-4
- Khalil, K., Asgher, U., and Ayaz, Y. (2022). Novel fnirs study on homogeneous symmetric feature-based transfer learning for brain-computer interface. *Sci. Rep.* 12, 3198. doi: 10.1038/s41598-022-06805-4
- Khanna, A., Pascual-Leone, A., Michel, C. M., and Farzan, F. (2015). Microstates in resting-state EEG: current status and future directions. *Neurosci. Biobehav. Rev.* 49, 105–113. doi: 10.1016/j.neubiorev.2014.12.010
- Koenig, T., Prichep, L., Lehmann, D., Sosa, P. V., Braeker, E., Kleinlogel, H., et al. (2002). Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. *Neuroimage* 16, 41–48. doi: 10.1006/nimg.2002.1070
- Lehmann, D., Ozaki, H., and Pal, I. (1987). EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67, 271–288.
- Li, J., Li, N., Shao, X., Chen, J., Hao, Y., Li, X., et al. (2021a). Altered brain dynamics and their ability for major depression detection using EEG microstates analysis. *IEEE Trans. Affect. Comput.* 14, 2116–2126. doi: 10.1109/TAFFC.2021.3139104
- Li, J., Qiu, S., Shen, Y.-Y., Liu, C.-L., and He, H. (2019). Multisource transfer learning for cross-subject EEG emotion recognition. *IEEE Trans. Cybernet.* 50, 3281–3293. doi: 10.1109/TCYB.2019.2904052
- Li, R., Wang, L., Suganthan, P., and Sourina, O. (2022). Sample-based data augmentation based on electroencephalogram intrinsic characteristics. *IEEE J. Biomed. Health Inform.* 26, 4996–5003. doi: 10.1109/JBHI.2022.3185587
- Li, W., Huan, W., Hou, B., Tian, Y., Zhang, Z., and Song, A. (2021b). Can emotion be transferred?—a review on transfer learning for EEG-based emotion recognition. *IEEE Trans. Cogn. Dev. Syst.* 14, 833–846. doi: 10.1109/TCDS.2021.3098842
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and De Peralta, R. G. (2004). EEG source imaging. *Clin. Neurophysiol.* 115, 2195–2222. doi: 10.1016/j.clinph.2004.06.001
- Murray, M. M., Brunet, D., and Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20, 249–264. doi: 10.1007/s10548-008-0054-5
- Padhmashree, V., and Bhattacharyya, A. (2022). Human emotion recognition based on time-frequency analysis of multivariate EEG signal. *Knowledge Based Syst.* 238, 107867. doi: 10.1016/j.knsys.2021.107867
- Pan, S. J., and Yang, Q. (2010). A survey on transfer learning. *IEEE Trans. Knowledge Data Eng.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665.
- Rached, T. S., and Perkusich, A. (2013). "Emotion recognition based on brain-computer interface systems," in *Brain-Computer Interface Systems-Recent Progress and Future Prospects*, ed R. Fazel-Rezai (Books on Demand), 253–270.
- Scharmüller, W., Leutgeb, V., Schäfer, A., and Schienle, A. (2012). Investigating phobic specificity with standardized low resolution brain electromagnetic tomography (Sloreta). *Brain Res.* 1477, 74–82. doi: 10.1016/j.brainres.2012.08.016
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Soc. Sci. Inform.* 44, 695–729. doi: 10.1177/0539018405058216
- Shen, X., Hu, X., Liu, S., Song, S., and Zhang, D. (2020). "Exploring EEG microstates for affective computing: decoding valence and arousal experiences during video watching," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (IEEE), 841–846.
- Strik, W. K., Chiaramonti, R., Muscas, G. C., Paganini, M., Mueller, T. J., Fallgatter, A. J., et al. (1997). Decreased EEG microstate duration and anteriorisation of the brain electrical fields in mild and moderate dementia of the Alzheimer type. *Psychiatry Res. Neuroimaging* 75, 183–191.
- Wang, W., and Li, B. (2022). A novel model based on a 1D-resCNN and transfer learning for processing EEG attenuation. *Comput. Methods Biomech. Biomed. Eng.* 1–14. doi: 10.1080/10255842.2022.2162339
- Wu, M., Hu, S., Wei, B., and Lv, Z. (2022). A novel deep learning model based on the ICA and Riemannian manifold for EEG-based emotion recognition. *J. Neurosci. Methods* 378, 109642. doi: 10.1016/j.jneumeth.2022.109642
- Wu, M., Teng, W., Fan, C., Pei, S., Li, P., and Lv, Z. (2023). An investigation of olfactory-enhanced video on EEG-based emotion recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* 31, 1602–1613. doi: 10.1109/TNSRE.2023.3253866
- Xue, J., Wang, J., Hu, S., Bi, N., and Lv, Z. (2022). OVPD: odor-video elicited physiological signal database for emotion recognition. *IEEE Trans. Instrument. Meas.* 71, 1–12. doi: 10.1109/TIM.2022.3149116
- Zhang, X.-Y., and Liu, C.-L. (2012). Writer adaptation with style transfer mapping. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1773–1787. doi: 10.1109/TPAMI.2012.239
- Zhao, X., Liu, D., Ma, L., Liu, Q., Chen, K., Xie, S., and Ai, Q. (2022). Deep CNN model based on serial-parallel structure optimization for four-class motor imagery EEG classification. *Biomed. Signal Process. Control* 72, 103338. doi: 10.1016/j.bspc.2021.103338
- Zhao, Y., Zeng, H., Zheng, H., Wu, J., Kong, W., and Dai, G. (2023). A bidirectional interaction-based hybrid network architecture for eeg cognitive recognition. *Comput. Methods Biomed. Eng.* 238, 107593. doi: 10.1016/j.cmpb.2023.107593
- Zheng, W.-L., and Lu, B.-L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497