An Investigation of Olfactory-Enhanced Video on EEG-Based Emotion Recognition

Minchao Wu, Wei Teng, Cunhang Fan, *Member, IEEE*, Shengbing Pei, Ping Li, and Zhao Lv

***Abstract*— Collecting emotional physiological signals is significant in building affective Human-Computer Interactions (HCI). However, how to evoke subjects' emotions efficiently in EEG-related emotional experiments is still a challenge. In this work, we developed a novel experimental paradigm that allows odors dynamically participate in different stages of video-evoked emotions, to investigate the efficiency of olfactory-enhanced videos in inducing subjects' emotions; According to the period that the odors participated in, the stimuli were divided into four patterns, i.e., the olfactory-enhanced video in early/later stimulus periods (OVEP/OVLP), and the traditional videos in early/later stimulus periods (TVEP/TVLP). The differential entropy (DE) feature and four classifiers were employed to test the efficiency of emotion recognition. The best average accuracies of the OVEP, OVLP, TVEP, and TVLP were 50.54%, 51.49%, 40.22%, and 57.55%, respectively. The experimental results indicated that the OVEP significantly outperformed the TVEP on classification performance, while there was no significant difference between the OVLP and TVLP. Besides, olfactory-enhanced videos achieved higher efficiency in evoking negative emotions than traditional videos. Moreover, we found that the neural patterns in response to emotions under different stimulus methods were stable, and for Fp1, FP2, and F7, there existed significant differences in whether adopt the odors.**

***Index Terms*— Electroencephalogram (EEG), emotion recognition, human–computer interface (HCI), olfactory-enhanced video, neural pattern.**

Manuscript received 27 November 2022; revised 7 February 2023; accepted 5 March 2023. Date of publication 8 March 2023; date of current version 13 March 2023. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61972437, in part by the Excellent Youth Foundation of Anhui Scientific Committee under Grant 2208085J05, in part by the National Key Research and Development Program of China under Grant 2021ZD0201502, in part by the Special Fund for Key Program of Science and Technology of Anhui Province under Grant 202203a07020008, in part by the Natural Science Foundation of Anhui Province under Grant 2108085MF207, in part by the Open Research Projects of Zhejiang Laboratory under Grant 2021KH0AB06, in part by the Open Projects Program of National Laboratory of Pattern Recognition under Grant 202200014, in part by the Open Fund Project of Key Laboratory of Civil Aviation Flight Technology and Flight Safety under Grant FZ2022KF15, and in part by Cloud Ginger XR-1. *(Corresponding author: Zhao Lv.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Anhui University.

The authors are with the Anhui Province Key Laboratory of Multimodal Cognitive Computation and the School of Computer Science and Technology, Anhui University, Hefei 230601, China (e-mail: wu_min_chao@hotmail.com; tengwei1998@outlook.com; cunhang.fan@ahu.edu.cn; shengbingpei@ahu.edu.cn; pingli0112@gmail.com; kjlz@ahu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3253866

I. Introduction

COMPREHENDING humans' emotions and making corresponding responses are humans' basic capabilities in daily communication and decision making. However, the two capabilities are essential to whether a machine is more intelligent [1]. In recent years, the development of Human-Computer Interactions (HCI) system with emotional autonomous perception has become an significant research hot spot in the field of artificial intelligence (AI) and HCI [2].

As an important tool for building an affective HCI system, electroencephalogram (EEG) is outstanding for analyzing brain activities and understanding mental states [3]. Besides, compared with some non-physiological signals, such as facial expressions and body languages, EEG can reveal more real emotional states, that is to say, humans can easily disguise their true emotions by controlling their facial expressions or body languages, but they hardly control their brain activities. Therefore, EEG-based emotion recognition has attracted great mention, from the theoretical research of emotional physiology to the engineering application of affective Brain-Computer Interactions (aBCI) system [4].

Evoking emotions and collecting related EEG signals are keys for studying EEG-based emotion recognition. However, defining emotional states is always a fuzzy problem, for humans' emotions may be influenced by many different exterior factors, such as time, culture, language, space, etc [5]. What's more, the individual's emotional states are also determined by their physiological and cognition states [6]. Hence, when performing the experiment of collecting emotional EEG signals, how to induce participants' emotions more sufficiently, and how to activate the level of participants' physiological states more effectively? The stimulation of multi-sensory seems to be potential when developing an emotional EEG experiment, for the multi-sensory integration, namely interactive synergy among the senses, can enhance the physiological salience of a stimulus [7]. Besides, some previous works had also indicated that the multisensory integration could improve the efficiency of emotion recognition task [8], [9], [10].

In the present work, we simultaneously adopted odor and video as stimulus materials for the purpose of fully stimulating the subjects' multi-sensory, thus inducing their emotions more efficiently. The main contributions of this work can be summarized as follows: (1) We have designed a novel emotional EEG experiment paradigm, and developed a multi-modal multi-sensory emotional EEG dataset (multi-modal: EEG and EOG, multi-sensory: vision, audition, and olfaction) which is publicly available.[1] (2) The odors are employed to participate in the different stage of a video-stimulated emotional experiment, to investigate the evoking efficiency of olfactory-enhanced video, and the results illustrated that odors could significantly improve the efficiency for evoking emotions in the early stage of traditional video rather than the later stage. (3) The experimental results demonstrate that the olfactory-enhanced video could achieve higher classify accuracies on recognizing negative emotions, and preliminary reveal the capability of odors on emotion regulation on a video-stimulated emotional experiment.

The layout of this paper is organized into following sections: In Section II, a brief introduction of related works is presented. The detailed information of the experiment is described in Section III. Methods and experimental results are presented in Section IV and V, respectively. The discussion can be available among Section VI. Finally, Section VII presents the conclusions of the current work.

II. Related Works

Emotions are usually an individual's subjective response arisen by either an internal or external event with positive or negative meaning [11]. In emotion recognition reaserch, emotional states can be roughly categorized as affective states (such as positive, neutral, and negative), dimensional states (such as high/low valence and arousal), and discrete states (such as happiness, sadness, calmness, and etc) [12]. However, quite a number of EEG-based emotional experiments have been developed which utilized kinds of stimulation materials with different emotional polarities to stimulate participants' different sensors, and employed these emotional categorization principle. Lin et al. [13] adopted the discrete states as emotional principle and the music as stimulus material to analyze the emotional EEG. Do Bos [14] employed the pictures, musics, or a combination of the both to stimulate participants' emotions, and categorized those emotions as dimensional states. Besides, video clips are the most popular stimulation materials among researchers, for the video clips have the ability to stimulate subjects' visual and auditory senses simultaneously. Both Koelstra et al. [15] and Soleymani et al. [16] adopted the dimensional principle and built the emotional dataset, i.e., DEAP and MAHNOB-HCI, respectively. In addition, Zheng and Lu [17] also built a widely used dataset, SEED, which classified emotions as positive, neutral, and negative. However, emotion-evoked experiments involve not only stimulating participants' vision and audition, but also collecting emotion-related signals by stimulating participants' other senses, namely the taste sense [18], the tactile sense [19], and the olfactory sense [20], these experiments have indicated that individuals' emotion also could be evoked successfully by stimulating their taste, tactile, or olfactory senses.

Nevertheless, the generation of individual emotions in real life is often due to the participation of multiple stimuli and multiple senses. Stimulation of a single sense or two senses may not be sufficient to enough activate participants' physiological arousal. Therefore, the strategy that combines different kinds of stimulation materials for the purpose of enhancing the individuals' emotions is developed. Both Raheel et al. [21] and Li et al. [22] adopted the strategy of the tactile enhanced video to simultaneously stimulate subjects' visual, auditory, and tactile senses, to evoke subjects' dimensional or discrete emotions, respectively. Ranasinghe et al. [23] developed an olfaction enhanced game that stimulate both subjects' visual, auditory, and olfactory senses to enhance their discrete emotions. Raheel et al. [24] and Xue et al. [25] employed the pattern of olfactory-enhanced video to evoke subjects' affective states. All the above-mentioned multi-stimuli experiments have demonstrated the higher efficiency of the collaboration of multi-senses on recognizing emotions.

One of the goals for these EEG based emotional experiments is to investigate the brain activities and neural patterns in response to different emotional states. Various works have focused on answering this question. The temporal and parietal lobes were also indicated as the brain areas related to processing the emotional response and change [13], [17], [26]. Zotev et al. [27] reported that the healthy subjects' positive emotion could achieve self-regulation by simultaneously regulating their BOLD fMRI activation in the left amygdala and frontal EEG power asymmetry in the high-beta band. Zheng et al. [17] indicated that the beta and gamma bands of EEG signals were the critical frequency bands on EEG-based emotion recognition tasks through a systematic comparison. Moreover, various machine learning and deep learning methods for emotion recognition also support this conclusion [28], [29], [30]. In addition, Zheng et al. [31] further reported that the stable neural patterns for different emotional states over time do exist. Reheel et al. [24] compared the difference of brain activities between traditional video and olfactory-enhanced video by recording the EEG signals of AF7, AF8, F7, and F8.

Another goal for EEG based emotional research is to accurately recognize individuals' emotional states in daily life to accomplish the HCI task better. Feature extraction and pattern recognition are the two core points for classifying different emotional states. Zheng and Lu [17], [31] developed differential entropy (DE) as emotional features which achieve better performance among machine or deep classifiers. Besides, statistics features (mean, variance, etc), non-linear features (approximate entropy, permutation entropy, etc), and time-frequency features (power specture density (PSD), wavelet energy, etc) are both frequently selected as features to classify emotions [32], [33]. In addition, Aydın et al. [34] and Kılıç and Aydın [35] adopted functional correlation features, such as conventional coherence function (CCF) and phase lag index (PLI), to recognizing different emotional states. However, compared with features, different classifiers would also achieve different recognition

[1] http://iiphci.ahu.edu.cn/toxiujue

performance. Therefore, both traditional machine learning methods such as logistic regression (LR) and support vector machine (SVM), and deep learning models such as deep belief network (DBN) and convolutional neural network (CNN), have achieved great results in EEG-based emotion recognition [17], [24], [25], [28], [29], [30], [31], [32], [33], [34], [35].

These are parts of introductions to experimental and neural pattern studies for EEG-based emotion recognition. However, the majority of these studies are based on analysis of emotional EEG signals evoked by audio-visual-related stimulation materials, whereas only a small minority of studies involved more sensory stimulation. Although Raheel et al. [24] and Xue et al. [25] had performed experiments with olfactory-enhanced videos as stimulus material, and briefly investigated the corresponding neural pattern, the olfactory-enhanced-video stimulation related brain activity still needs further study. Hu et al. [36] indicated that the efficiency of the different period of a complete video clip do exist differences on evoking subjects' emotions, where the later period of a video would reach higher awaking capability. Therefore, when odors participate to the different stage of a video stimulation, whether the efficiency of awaking participants' emotions would change, and whether the brain activity for different pattern would change? In addition, which emotions do olfactory enhancement videos have a stronger ability to induce? To the best of our knowledge, there is no available work for the analysis of these problems.

III. Experiments

A. Subjects

Sixteen healthy participants (8 males and 8 females; age range: 20-29 years old, mean: 23.53, std: 2.11), with normal or corrected-to-normal visual acuity, normal hearing, and normal smelling, participated in the experiment. All participants were native Chinese students from Anhui University with different educational degrees, i.e., undergraduate, master's, and doctorate. And they were informed about the experimental procedure and required to overcome movements as possible. After participating in each trial of an experiment, the participants needed to make a self-assessment for reporting their emotional states.

B. Stimulus Materials

Before performing the emotion experiments, we first selected 60 video clips, where each clip last 2 minutes (about 20 clips each of positive, neutral, or negative emotion labels). The criteria for selecting the video clips were as follows: (a) the videos should not cause any confusion; (b) the videos should elicit a single target emotional state effectively; (c) the videos should not to be too long and contain a relatively complete storyline. We invited 20 volunteers to watch these clips and report their emotions by giving scores from -7 to 7 (negative: -7 to -3, neutral: -2 to 2, positive: 3 to 7). Finally, 30 video clips for the three different emotions had been chosen where each emotion had 10 video clips, where the sources of the selected video clips were shown in Table I.

Besides, we also selected 20 different odors. The criteria for selecting the odors were as follows: (a) the odors should be not harmful to participants' health; (b) each odor should smell single rather than mixed; (c) the odors should elicit a single target emotional state effectively. It is worth pointing out that the odor smell single means that when participants smell the odor, they could accurately judge the source of the smell, rather than worrying about how many flavors the odor contains, such as a mixture of alcohol and rose. Typically, We borrowed the fluid volatility to subject the participants to the corresponding odor stimulus. Similarly, after scoring the odors by 20 volunteers, we finally chose 9 odors for three emotions shown in Table I, where the positive odors contained lavender, rose, orange, and florida water; the negative odors contained industrial alcohol, essential balm, vinegar, and ink; while the neutral odor only contained water. Besides, we also performed the blank controls when subjects watch the videos with the neutral label.

TABLE I
Details of Film Clips and Odors in Our Emotion Experiment

Emotion label	Film clips sources	Odors sources
Positive	My Own Swordsman Home With Kids Kung Fu Hustle Hello Mr. Billionaire	Rose Orange Lavender Floral water
Neutral	Hexi Corridor Soothing Light Music videos	Water Blank control
Negative	After Shock Better Days Back to 1942 Dying to Survive	Alcohol Essential balm Vinegar Ink

[1] The blank control means don't add any odors when subjects watching video clips.
[2] The alcohol typically represents medicinal alcohol.
[3] For each emotional label, to avoid avoid the bias of experimental results caused by some fixed collocation, we employed the strategy for random matching of different videos and odors.

C. Experiment Paradigm And Data Acquisition

In order to investigate the different emotional neural patterns for olfactory-enhanced videos as stimulus materials simultaneously, we designed a new emotional experiment to record affective EEG signals. The protocol was shown in Fig. 1. Each trial contained four parts, i.e., the start was a 3-s hint, the following was a 2-minutes stimulus period, the next was 10-s feedback, and the last was 20-s resting period. Typically, considering the differences in emotional tendencies between individuals when receiving the same video and odor stimuli, subjects were also asked to rate their feeling from -7 to 7 when making feedback, and this is the basis for a subsequent experimental analysis rather than the labels of the stimulus materials. Besides, we easily divided each trial into two periods, i.e., the early stimulation period (the first minute of the video) and the later stimulation period (the second minute of the video). In detail, when performing the k-th trial, subjects were stimulated by the traditional video in the early stimulation period, and then stimulated by the olfactory-enhanced video in the later stimulation period. In contrast, when performing the $k+1$-th trial, subjects were stimulated by the olfactory-enhanced video in the early stimulation period, and then stimulated by the traditional video in the later stimulation

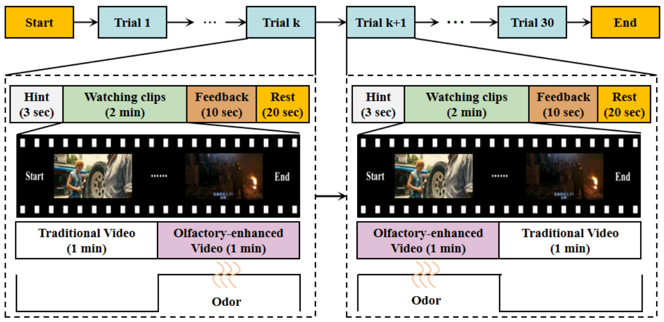
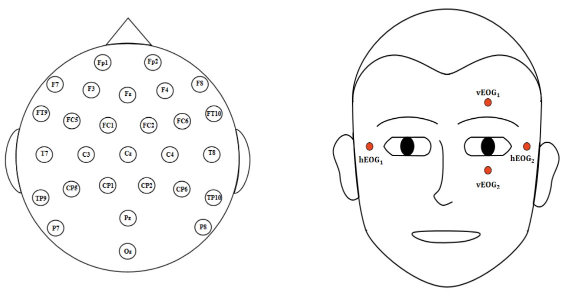


Fig. 1. Paradigm of the EEG experiment with the olfactory-enhanced video as stimulation material.



(a) (b)

Fig. 2. The electrode placements. (a) The EEG cap layout for 28 channels; (b) The EOG setups.

period. For simplicity, we denote the four stimulus pattern as ‘TVEP’, ‘TVLP’, ‘OVEP’, and ‘OVLP’, where the four notations represent the traditional video with early stimulation period, the traditional video with later stimulation period, the olfactory-enhanced video with early stimulation period, and the olfactory-enhanced video with later stimulation period, respectively. In total, there were 30 trials for each experiment.

The EEG data was recorded by Brain Products[2] with a sampling rate of 250 Hz. The layout of EEG electrodes with 28 channels on the cap obeyed the international 10-20 system and was shown in Fig. 2(a). Meanwhile, the EOG data was recorded simultaneously where the electrode placements were shown in Fig. 2(b). In total, the physiological signal with 32 channels (28 channels for EEG, and 4 channels for EOG) for 30 trials had been recorded for each subject, and each trial continued 2 minutes, where the first minute and the second minute of the trial recorded the signals stimulated by the traditional video/ olfactory-enhanced video and olfactory-enhanced video/ traditional video, respectively. The emotional EEG-EOG signals can be publicly available from http://iiphci.ahu.edu.cn/toxiujue.

IV. Methods

A. Preprocessing

According to the feedback of subjects after each trial, only the trial when the subject reported a same emotion state for both the two stimulus stages of the trial was chosen for further analysis. The raw EEG signals were first notch filtered by 50 Hz to remove the power-line noise, followed by a band-pass filter between 1 and 50 Hz. Then, the average re-reference was performed on the EEG signals. Finally, the independent component analysis (ICA) was ran on the preprocessed EEG signals to remove the artifacts. Specially, the ADJUST-plugin[3] toolbox was employed to automatically classified the independent components (ICs). All above-mentioned preprocessing steps were ran in EEGLAB toolbox.[4]

B. Feature Extraction

A more concrete band pass filter was performed for EEG signals which had been preprocessed to extract a specific frequency band EEG signals, i.e., the delta band (1-3 Hz), the theta band (4-7 Hz), the alpha band (8-12 Hz), the theta band (13-30 Hz), and the gamma band (31-50 Hz). After that, a moving window with 1 second was used to partition the preprocessed EEG of each trial into multiple EEG segments, and each trial were obtained 120 data segments (60×2, 60 data segments for the traditional video as stimulus material, and 60 data segments for the olfactory-enhanced video as stimulus material). For further analysis, we extracted the DE feature from the EEG segments with five different frequency bands, for the reason that the DE feature had been reported to achieve better performance than other features, such as power spectral density [31].

C. Feature Smoothing

Performing the feature smoothing allows us to reduce the influence of emotion-unrelated EEG, and take advantage of the time dependency when emotion changes [31], [37]. Therefore, feature smoothing may be a crucial step on building an EEG-based emotion recognition system. Zheng and Lu [17], [31] developed the linear dynamic systems (LDS) to smooth EEG feature, and indicated that the LDS outperformed the method of moving average. Nevertheless, LDS may spend more time on smoothing the features. Val-Calvo et al. [37] compared the Savitzky-Golay (SG) filtering with LDS, and demonstrated that both the two smoothing method had outstanding property on improving the classification accuracy, while the SG was significantly faster than LDS on smoothing features. Therefore, in the current work, the SG smoothing method is employed to make the DE features more robust and discriminative on performing EEG-related emotion classification.

D. Classifiers

The experimental environment was built on a Windows 10 PC with Core (TM) i7–8700 CPU and 16 GB memory. The computing environment was pytorch 1.8.0.

The DE features were further fed to four different classcifiers, i.e., logistic regression (LR), support vector machine (SVM), k nearest neighbors (KNN), and deep believe network (DBN), to achieve the EEG-based emotion recognition.

[2]https://www.brainproducts.com/

[3]https://www.nitrc.org/projects/adjust/

[4]https://sccn.ucsd.edu/eeglab/

LR, SVM, and KNN are traditional classifiers. The LR classifier is a generalized linear model based on statistics knowledge and aims at find an optimal hyperplane to achieve the classification task. Similar to LR, the SVM classifier also aims to find an optimal hyperplane, while the SVM is a model based on geometric which only consider the support vectors. Specially, the regulation parameter C of LR, and the optimal regulation parameter C and the kernel function hyperparameter $gamma$ of SVM were chosen from the parameter pool $P_{value} = \{0.01k, 0.1k, k|k = 1, 2, \ldots, 9\}$. Specially, the grid search method was employed to search the optimal parameters, such as the hyperparameter $gamma$. Besides, the $L2$ penalty term and the 'rbf' kernal were also employed for LR and SVM, respectively.The KNN classifier is an Euclidean distance-based model that calculates the distance between the target sample and all the training samples, and classifies the target sample by the training samples corresponding to the first k minimum distances. The number of nearest neighbors k is set to 5.

Compared to the traditional neural network models, DBN is a probabilistic generative model which establishes a joint distribution between the observed data and the labels. In this work, a classical structure of DBN classifier with 2 layers of restricted Boltzmann machine (RBM) and a layer of back propagation (BP) was employed. The optimal numbers of neurons in the first and second hidden layers with step of 50 in the ranges of [50, 200] and [50, 200], respectively. The adam optimizer with learning rate of 0.001 and weight decay parameter of 0.001 was adopted. And the batch size and maximum epochs were 16 and 500, respectively.

Besides, due to the individual's difference on cognition or emotional tendency, the emotional data may unbalanced for different emotional states after emotional stimulation. Hence, we employed the strategy of balancing the class weight to avoid the problem for SVM and LR. And for DBN, we employed the strategy of resampling.

V. Experiments Results

A. Classification Performance

In this subsection, we compared the effectiveness of the four different emotion-evoked patterns, i.e., the TVEP pattern, the TVLP pattern, the OVEP pattern, and the OVLP pattern. The DE features were respectively feed to different classifiers, i.e., LR, SVM, KNN, and DBN, to build the EEG-based emotion recognition systems and achieve the comparison. The leave-one-trial-out (LOTO) cross validation strategy was applied to validate the efficiency of emotion recognition. More concretely, assume that there are n trials of one subject, we select n-1 trials of EEG data as the training data, and the rest one trial of EEG data as the test data. The reason we employed the LOTO strategy was mainly based on that the EEG with a specific task is sensitive to change caused by the differences in cognitive states and environmental variables [39]. That is to say, in k-th trial of an experiment, the pattern of EEG signal may keep stable, while in the next trial, the pattern of EEG signal will change. Therefore, the traditional K-fold cross validation strategy hardly consider the characteristics of EEG signals over time, environment, cognitive states and other factors, while the LOTO strategy does relatively better in this aspect.

To evaluate the performance of different simulated patterns and classifiers, the overall accuracy [40] was employed which is defined as

$$Acc = \frac{\sum_{n=1}^{N}\sum_{c=1}^{C} m_{nc}}{\sum_{n=1}^{N}\sum_{c=1}^{C} M_{nc}} \tag{1}$$

where C is the number of classes, N is the number of trial for an experiment, m_{nc} is the number of correctly predicted samples of the $c-$th class in the $n-$th trial, and M_{nc} denotes the total samples of the $c-$th class in the $n-$th trial. The percent theoretical chance level [41] of classification is given by $\frac{100}{C}$ (i.e., in this work, the chance level = 33.33%.)

The average accuracies (standard deviations) of four classifiers from different frequency bands over all 16 subjects were shown in Table II. As shown in Table II, the beta and gamma bands achieved higher accuracies among the five frequency bands regardless of classifiers or stimulus patterns for the emotion recognition task. Besides, regardless of classifiers, the best accuracies of the beta and gamma bands were achieved with the TVLP pattern, followed by the OVEP and OVLP patterns, and last the TVEP pattern. Compared with the OVEP, OVLP, and TVEP patterns, the TVLP pattern achieved 7.25%, 6.06%, and 19.81% higher accuracies for LR in the gamma band, and 6.47%, 8.54%, and 17.11% higher accuracies for LR in the beta band, respectively. For SVM, KNN, and DBN, a similar difference can be calculated between the TVLP pattern and other three stimulus patterns. In addition, the OVEP and OVLP performed very similarly in terms of the classification accuracy for all classifiers, where the difference between the two patterns was always around 1% in the beta and gamma bands. While, both the OVEP and OVLP patterns achieved around 10% higher accuracy in the gamma band, and around 8% higher accuracy in the beta band than the TVEP pattern, respectively.

To determine if the differences of the accuracies between the four stimulus patterns for the five frequency bands was statistically significant, the paired-sample t-test by MATLAB function $ttest$ was employed. However, before employing the t-test, the Lilliefors test was also employed to verify that the data come from a normal distribution by MATLAB function $lillietest$. We calculated the t-tests for these four pairs of stimulus patterns, i.e., TVEP-TVLP, TVEP-OVEP, OVEP-OVLP, and TVLP-OVLP, where the results were shown in Table III. As shown in Table III, there existed a significant difference between the TVEP and TVLP patterns for LR, SVM, KNN, and DBN, in the beta and gamma bands. Besides, there also existed a significant difference between the TVEP and OVEP patterns in the beta band for LR and KNN, as well as in the gamma band for LR, SVN, and KNN. Although there was no significant difference between the TVEP and OVEP patterns for DBN in the beta and gamma bands when $\alpha = 0.05$, the p-values were still close to 0.05 with a difference of around 0.01. However, the p-values of the OVEP and OVLP patterns were always higher than 0.05 for the four classifiers in the beta and gamma bands, which means that there was hardly no significant difference between the two

TABLE II
THE AVERAGE ACCURACIES AND STANDARD DEVIATIONS OVER 16 SUBJECTS FOR DIFFERENT CLASSIFIERS WITH PATTERNS OF TVEP, TVLP, OVEP, AND OVLP

Stimulus Pattern	Classifier	Delta	Theta	Alpha	Beta	Gamma
TVEP	LR	35.52(7.58)	37.51(7.68)	39.15(6.82)	39.32(9.00)	37.74(10.64)
	SVM	37.47(7.64)	38.78(6.98)	39.85(7.36)	41.19(9.07)	40.03(10.29)
	KNN	36.93(10.25)	37.16(10.26)	37.75(8.57)	39.77(8.57)	39.23(9.57)
	DBN	37.36(7.71)	37.72(6.81)	40.14(7.07)	39.74(7.78)	40.22(9.99)
OVLP	LR	40.38(8.38)	40.55(8.76)	40.18(9.13)	47.89(12.88)	51.49(11.93)
	SVM	40.65(9.39)	39.74(11.24)	43.16(10.07)	49.11(12.01)	51.13(13.38)
	KNN	37.44(11.29)	38.61(10.28)	38.68(9.65)	46.57(12.73)	47.18(13.94)
	DBN	39.41(7.95)	39.11(8.37)	39.30(9.03)	45.69(11.93)	48.68(13.04)
OVEP	LR	40.38(8.22)	41.85(8.52)	43.27(9.96)	49.96(17.38)	50.30(18.05)
	SVM	40.24(9.03)	41.01(8.95)	43.70(8.96)	50.04(17.81)	50.54(18.21)
	KNN	40.54(10.05)	40.50(9.15)	40.63(9.24)	47.55(15.71)	48.87(17.46)
	DBN	38.45(8.80)	39.76(8.70)	41.07(8.52)	47.67(16.60)	48.14(17.31)
TVLP	LR	41.26(6.79)	41.02(7.74)	44.66(10.72)	56.43(19.55)	57.55(19.34)
	SVM	40.49(6.76)	40.65(7.93)	43.11(10.04)	54.88(18.59)	55.24(18.56)
	KNN	40.54(9.11)	40.98(9.43)	42.64(10.16)	52.50(16.72)	53.64(17.73)
	DBN	39.61(7.20)	39.40(7.62)	42.49(11.34)	55.30(17.77)	57.17(18.26)

TABLE III
PAIRED-SAMPLE t-TEST RESULTS ON THE ACCURACIES OVER 16 SUBJECTS FROM THE FIVE FREQUENCY BANDS ($\alpha = 0.05$)

Frequency	TVEP - TVLP				OVEP - OVLP				TVEP - OVEP				TVLP - OVLP			
	LR	SVM	KNN	DBN	LR	SVM	KNN	DBN	LR	SVM	KNN	DBN	LR	SVM	KNN	DBN
Delta	**0.0282**	0.2183	0.2397	0.3577	0.9982	0.8728	0.3299	0.7085	0.0561	0.3104	0.2557	0.6831	0.7466	0.9560	0.3436	0.9370
Theta	0.1545	0.4475	0.2083	0.5039	0.6029	0.7053	0.5525	0.7970	0.0819	0.4334	0.2902	0.4524	0.8479	0.7782	0.4509	0.9071
Alpha	0.0710	0.2405	0.0739	0.4333	0.3324	0.8578	0.4332	0.4432	0.1120	0.1237	0.2125	0.6210	0.1994	0.9859	0.1722	0.3237
Beta	**0.0040**	**0.0108**	**0.0091**	**0.0022**	0.6845	0.8590	0.8315	0.7055	**0.0330**	0.0566	**0.0437**	0.0616	0.0919	0.2600	0.1704	**0.0447**
Gamma	**0.0001**	**0.0051**	**0.0039**	**0.0041**	0.8318	0.9133	0.7429	0.9251	**0.0192**	**0.0386**	**0.0432**	0.0532	0.2765	0.4071	0.1863	0.1006

TABLE IV
THE AVERAGE ACCURACIES AND STANDARD DEVIATIONS OVER 16 SUBJECTS FOR DIFFERENT CLASSIFIERS WITH THE PATTERNS OF TRADITIONAL VIDEO AND OLFACTORY-ENHANCED VIDEO

Stimulus Pattern	Classifier	Delta	Theta	Alpha	Beta	Gamma
Traditional video	LR	38.65(5.72)	39.15(5.55)	42.08(4.67)	49.44(10.35)	48.83(10.72)
	SVM	40.35(6.04)	38.85(6.25)	42.75(4.79)	50.83(9.98)	49.73(11.11)
	KNN	39.45(7.51)	37.93(7.40)	39.91(6.84)	46.68(9.10)	46.24(10.99)
	DBN	38.46(5.35)	38.01(5.63)	41.05(4.91)	48.69(10.08)	47.31(11.10)
Olfactory-enhanced video	LR	38.95(6.43)	40.57(5.89)	42.82(7.80)	52.19(11.75)	53.78(12.46)
	SVM	40.07(7.01)	40.12(6.61)	42.55(7.76)	52.39(10.96)	53.30(12.91)
	KNN	38.40(8.59)	39.20(7.46)	40.43(8.12)	48.38(9.64)	48.06(11.73)
	DBN	39.19(5.35)	38.72(6.00)	40.58(7.28)	49.18(10.90)	49.38(12.02)

patterns. For the TVLP and OVLP patterns, we can make a similar conclusion.

According to the classification accuracies shown in Table II and the paired t-test results shown in Table III, it can be concluded that enhancing emotion induction through odor is mainly achieved with a higher efficiency in the early stage of the traditional video stimulation, while the enhancement effect in the later stage of the stimulation is not significant.

Furthermore, we also combined all trials of TVEP with TVLP, or all trials of OVEP with OVLP, to compare the efficiency of traditional video with olfactory-enhanced video on evoking participants' emotions. The classify accuracies were shown in Table IV. It can be found that, for each classifier, the stimulus pattern of olfactory-enhanced video always achieved higher accuracy.

B. Confusion Matrices

To further study the recognition performance of different classifiers in the three emotional states, the confusion matrices over all 16 subjects under the patterns of TVEP, TVLP, OVEP, OVLP, traditional video, and olfactory-enhanced video, had been calculated. Fig. 3 showed the results of the gamma band, for the gamma band is most tightly associated with evoking emotions and achieved higher performance in emotion recognition task.

As we could found that, the recognition capabilities of the different stimulus patterns had obvious differences in the three emotional states. For positive emotional state, the TVLP pattern achieved higher accuracies for four classifiers, followed by the OVEP pattern, the OVLP pattern, and last the TVEP pattern; For negative emotional state, the OVLP achieved higher accuracies for four classifiers, followed by the TVLP pattern, the OVEP pattern, and last the TVEP pattern; For neutral emotional state, the TVLP pattern achieved higher accuracies for LR and SVM, followed by the OVLP pattern, the OVEP pattern, and last the TVEP pattern. While for KNN, the OVLP pattern achieved higher accuracies, followed by the TVLP pattern. Finally, for DBN,

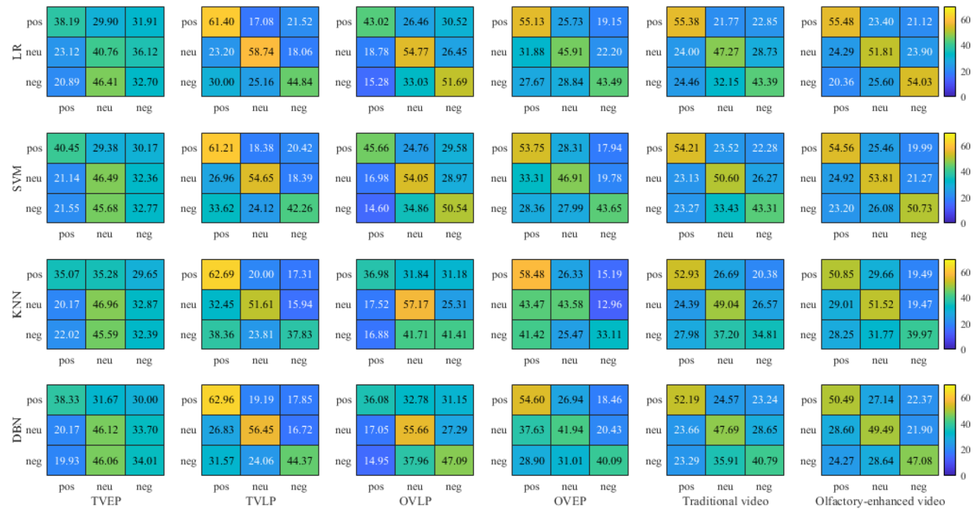


Fig. 3. Confusion matrices for different classifiers with the TVEP, TVLP, OVEP, OVLP, traditional video, and olfactory-enhanced video patterns. For each subfigure, from top to bottom: the LR, the SVM, the KNN, and the DBN classifiers; from left to right: the TVEP, TVLP, OVLP, OVEP, traditional video, and olfactory-enhanced video patterns. The label, 'pos', 'neu', and 'neg' represent the positive, neutral, and negative emotional states, respectively.

the OVLP and TVLP pattern achieved higher accuracies, respectively.

What's more, the OVEP pattern can significantly improve the classification accuracies for both positive and negative emotional states compared with the TVEP pattern. Besides, the OVLP pattern can also slightly improve the classification accuracies for negative emotional state compared with the TVLP pattern. In addition, it also can be observed that the OVEP pattern performed better on recognizing the positive emotional state than the OVLP pattern.

Furthermore, comparing the pattern of traditional video with that of olfactory-enhanced video, the pattern of the olfactory-enhanced video improved with about 7% and 3% higher accuracies on classifying the negative and neutral emotions than of the traditional video, respectively. While for recognizing the positive emotion, these two patterns performed similar.

C. Neural Patterns

To obtain the neural patterns associated with different emotions and different stimulus patterns, we projected the DE features to the scalp. Fig. 4 displayed the neural patterns of three emotions by calculating the average DE features over all subjects and then projecting them to the scalp. Fig. 4(a), Fig. 4(b), Fig. 4(c), and Fig. 4(d) displayed the neural patterns of the TVEP, TVLP, OVEP, and OVLP patterns, respectively. As shown in Fig. 4, an obvious difference could be observed between the five frequency bands for the different emotional states regardless of the stimulus patterns. That is to say, the beta and gamma bands had a stronger activation in the bilateral temporal cortex (T7, T8, FT9, FT10, TP9, and TP10) than the other three low frequency bands. Besides, we can also find that the neural patterns also existed differences between the positive, neutral, and negative emotional states in the beta and gamma bands. Compared with the neutral emotional state, the positive and negative emotional states activated more in the bilateral temporal cortex for both the beta and gamma bands.

What's more, as shown in Fig. 4, for the neutral emotion state, we observed that the TVLP (shown in Fig. 4(b)), OVEP (shown in Fig. 4(c)), OVLP pattern (shown in Fig. 4(d)) had a weaker activation response in the left temporal cortex (T7, FT9, and TP9) than the TVEP pattern (shown in Fig. 4(a)), while the activation responses of the TVLP, OVEP, and OVLP patterns performed similarly; For the positive emotional state, in general, both the four stimulus patterns performed similar activation responses in bilateral temporal cortex. Besides, it is obviously that the right temporal cortex (T8, FT10, and TP10) activated more than the left (T7, FT9, and TP9). In addition, for the negative emotional state, an obvious difference can be found from the four stimulus patters. That is to say, the TVLP pattern had a stronger activation response in the bilateral temporal cortex than the TVEP and OVLP patterns, while the OVEP pattern had a slightly stronger activation response than the OVLP pattern. What's more, contrary to the neutral emotion state, the activation response of the OVEP and OVLP patterns in the bilateral temporal cortex were stronger than the TVEP pattern, while the TVLP and OVEP patterns performed similarly on neural activation response in the bilateral temporal cortex.

Furthermore, comparing the traditional video pattern with the olfactory-enhanced video pattern, it can be drawn that there existed a significant difference in the frontal lobe and prefrontal lobe (F3, F4, F7, F8, Fp1, and Fp2). To indicate the significance of the difference for the DE features extracted from the EEG signal recorded by each electrode, we applied the non-parametric statistical test, the Kruskal-Wallis test (K-W test), on the DE feature of each frequency band of each electrode. The reason we adopted the K-W test was that the DE feature of each electrode was not completely subject to a normal distribution after performing the Lilliefors test.

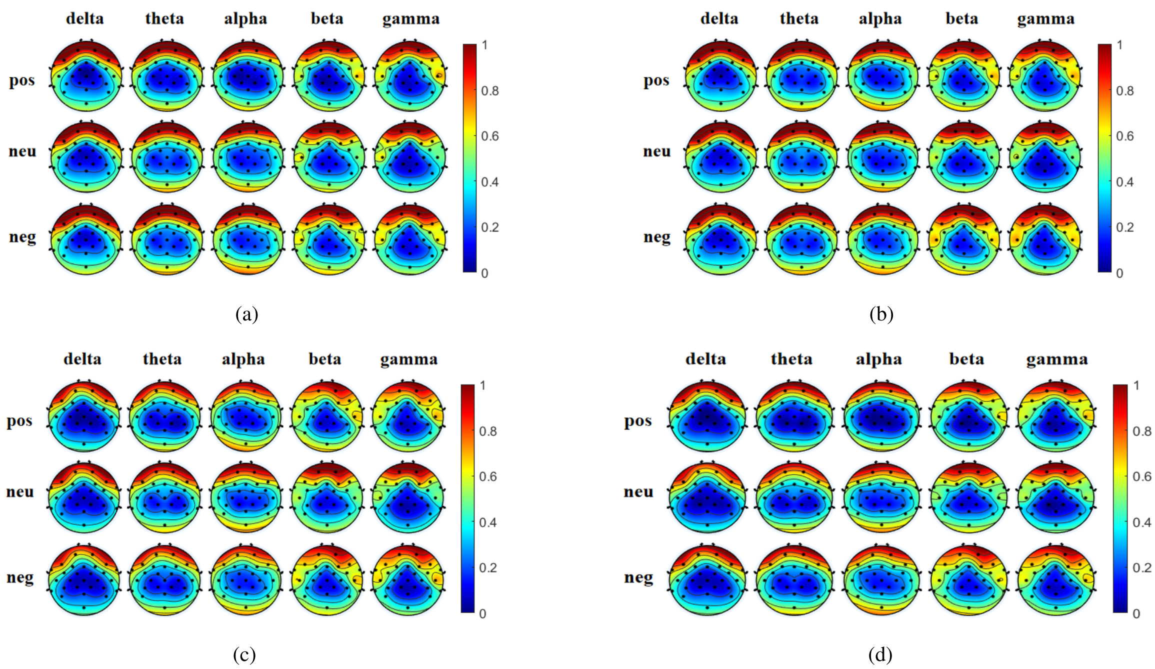


Fig. 4. Topographical maps of the DE features in the five canonical frequency bands. For each subfigure, from top to bottom: positive, neutral, and negative emotions; from left to right: delta, theta, alpha, beta, and gamma bands. For each frequency band, we normalized the DE features to range form 0 to 1 across the three emotional states and four stimulus patterns. (a) TVEP; (b) TVLP; (c) OVEP; (d) OVLP.

TABLE V

P-Values Against Traditional and Olfactory-Enhanced Video With Different Stimulus Period on All Bands of Fp1, Fp2, F3, F4, F7, and F8 Channels ($\alpha = 0.05$)

Channel	Frequency	TVEP-OVLP	TVEP-OVEP	TVLP-OVLP	TVLP-OVEP	TVEP-TVLP	OVEP-OVLP
Fp1	Delta	**0.0001**	**0.0001**	**0.0001**	**0.0001**	0.9724	0.9311
	Theta	**0.0001**	**0.0001**	**0.0001**	**0.0010**	0.5970	0.4136
	Alpha	**0.0342**	**0.0071**	**0.0156**	**0.0032**	0.7528	0.5186
	Beta	0.0624	0.1019	**0.0241**	**0.0426**	0.6043	0.9029
	Gamma	**0.0300**	0.0688	**0.0120**	**0.0301**	0.6528	0.8012
Fp2	Delta	**0.0001**	**0.0001**	**0.0001**	**0.0001**	0.9269	0.5887
	Theta	**0.0001**	**0.0001**	**0.0001**	**0.0001**	0.8173	0.4291
	Alpha	**0.0047**	**0.0498**	**0.0006**	**0.0104**	0.4850	0.4013
	Beta	**0.0378**	0.4571	**0.0072**	0.1719	0.4705	0.2651
	Gamma	**0.0221**	0.2609	**0.0042**	0.0893	0.5231	0.2959
F3	Delta	**0.0016**	**0.0221**	**0.0029**	**0.0319**	0.8892	0.5119
	Theta	**0.0306**	0.1761	**0.0306**	0.1642	0.8940	0.4976
	Alpha	0.5957	0.8473	0.5006	0.9817	0.8429	0.5827
	Beta	0.7842	0.9074	0.7303	0.9760	0.9503	0.7397
	Gamma	0.4857	0.5554	0.4935	0.5404	0.9917	0.9476
F4	Delta	**0.0042**	0.0632	**0.0019**	**0.0329**	0.7462	0.3691
	Theta	0.0548	0.3763	**0.0275**	0.2250	0.7306	0.3065
	Alpha	0.3731	0.0612	0.6296	0.1551	0.6760	0.3687
	Beta	0.3931	**0.0244**	0.9321	0.1464	0.3403	0.2610
	Gamma	0.7772	0.1301	0.7070	0.3931	0.3747	0.3214
F7	Delta	**0.0029**	**0.0349**	**0.0010**	**0.0140**	0.6428	0.5042
	Theta	**0.0008**	**0.0124**	**0.0002**	**0.0039**	0.5786	0.4959
	Alpha	**0.0138**	**0.0449**	**0.0082**	**0.0278**	0.7959	0.6913
	Beta	**0.0110**	**0.0267**	**0.0091**	**0.0215**	0.8667	0.8517
	Gamma	**0.0084**	**0.0120**	**0.0115**	**0.0151**	0.9614	0.9807
F8	Delta	0.1959	0.7869	**0.0263**	0.2409	0.2868	0.3742
	Theta	0.1002	0.5653	**0.0083**	0.1217	0.2647	0.3412
	Alpha	0.1633	0.6320	**0.0250**	0.1877	0.3794	0.3893
	Beta	0.2825	0.7166	0.0667	0.2612	0.3970	0.5103
	Gamma	0.4676	0.8582	0.1957	0.4621	0.5288	0.6213

The results of K-W test for six pairs of stimulus patterns, i.e., TVEP-OVLP, TVEP-OVEP, TVLP-OVLP, TVLP-OVEP, TVEP-TVLP, and OVEP-OVLP, in terms of p-values against five frequency bands of six channels were shown in Table V. As we can observe from Table V, as long as the kind of stimulus materials was consistent, namely both

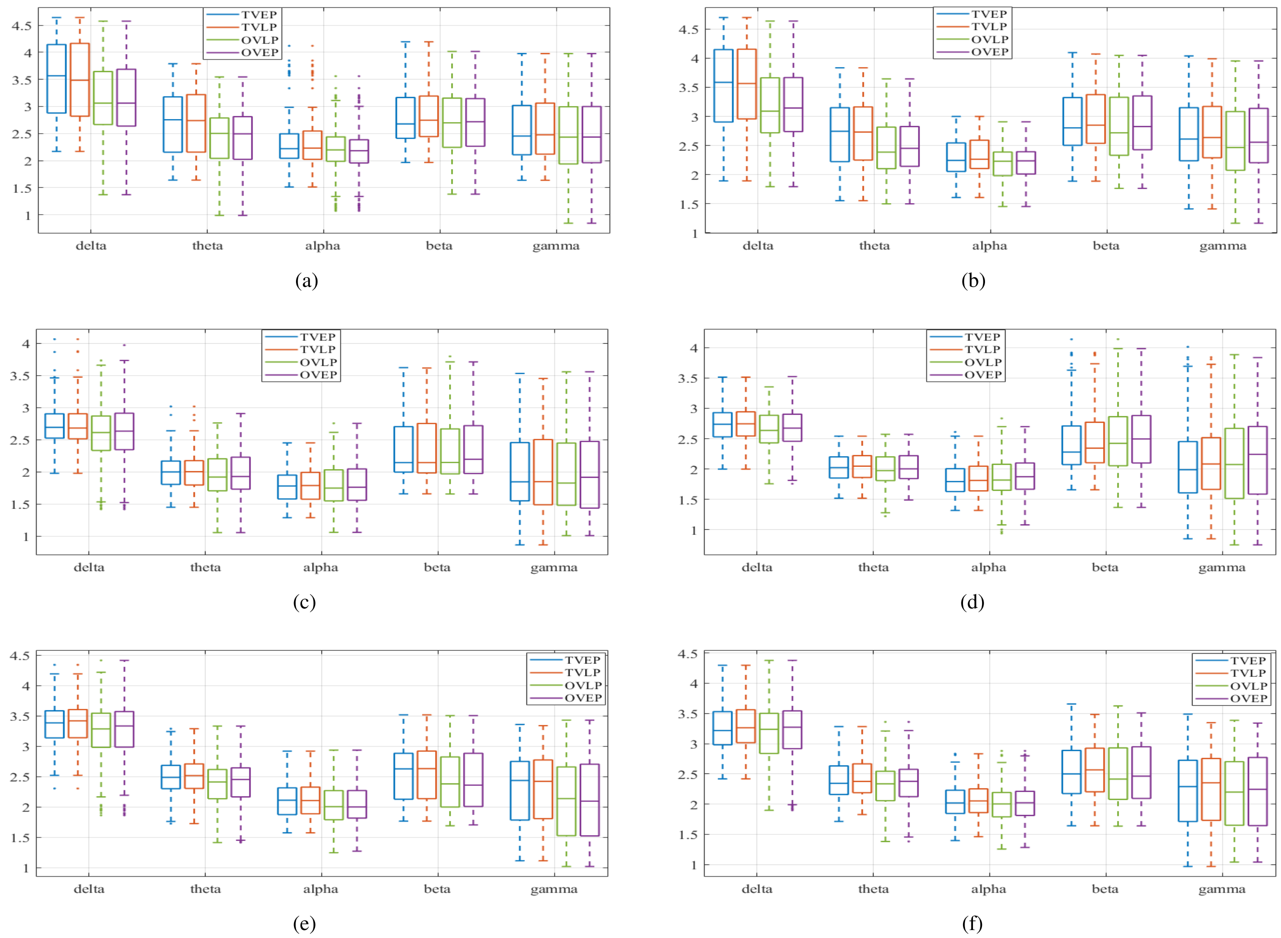


Fig. 5. Box plot of different EEG frequency bands against the four different stimulus patterns for EEG channels (a) Fp1; (b) Fp2; (c) F3; (d) F4; (e) F7; (f) F8.

traditional video or both olfactory-enhanced video, there were no significant difference (p-value $>$ 0.05) in response to DE features of each frequency band of each channel. Besides, the delta, theta, and alpha bands on Fp1 and Fp2, were found significant differences (p-value $\leq$ 0.05) between the traditional video and olfactory-enhanced video regardless of the stimulus period. Similar to the Fp1 and Fp2 channels, the delta band on the F3 and F4 channels also performed significant differences between the traditional video and olfactory-enhanced video, while the theta and alpha bands on the F3 and F4 channels didn't showed the significance. In particular, we observed that for the F7 channel, not only the delta, theta, and alpha bands, but also the beta and gamma bands, performed the significant differences between the two different kinds of stimulus materials. While for the F8 channel, there almost didn't exist significant differences, except for the delta, theta, and alpha bands with the pair of TVLP-OVLP.

Box plots for the DE feature of five frequency bands on channels, the Fp1, Fp2, F3, F4, F7, and F8 channels for the TVEP, TVLP, OVEP, and OVLP patterns were shown in Fig. 5. As shown in Fig. 5, for the Fp1 and Fp2 channels, the box plots (Fig. 5(a), Fig. 5(b)) showed a significant difference in the delta, theta, and alpha bands among the DE features between the traditional video and olfactory-enhanced video. In addition, for the F7 channel, the box plots (Fig. 5(e)) showed significant differences in both the five frequency bands. Obviously, the results displayed in the Fig. 5 were consistent with those of the K-W test.

VI. Discussion

A. Efficiency of the Proposed Experimental Paradigm

The stimulus pattern of the olfactory-enhanced video had been validated to outperform the traditional video, Table VI displayed some of the major studies about EEG-based emotion recognition in response to the traditional video, odor, or olfactory-enhanced video. As displayed in Table VI, the literature [24], [25], [42] adopted the stimulus pattern of olfactory-enhanced video with different experiment paradigms, where the literature [42] reported that the odors increase subjects' Quality of Experience (QoE) levels for olfactory-enhanced video by the approach of a combination of experiments and questionnaires, while both the literature [24], [25] verified the effectiveness of the olfactory-enhanced video through the machine learning method. However, all the above-mentioned three works were with a fixed stimulus order, which neglected the difference in the effectiveness of evoking subjects' emotions during the different periods of a complete video. In general, a stimulus video contains a relatively complete plot, thus with a gradual understanding of

TABLE VI
COMPARISON OF THE PROPOSED EXPERIMENT PARADIGM FOR EEG-BASED EMOTION RECOGNITION TASK WITH WORKS AVAILABLE IN THE LITERATURE

Literature	Emotions	Stimuli	Experimental paradigm	Subjects	Cross validation	Results
[42]	Pleasant, Unpleasant	TV/OV[1]	k-th trial: TV(30s) $\rightarrow^2$ OV(30s) $\rightarrow$ OV(30s) $\rightarrow$ TV(30s)	1000	-	-
[24]	Pleasant, Unpleasant	TV/OV	k-th trial: TV(45s) $k+1$-th trial: OV(45s))	20	10-fold	TV: acc=68.7% OV: acc=75%
[25]	Positive, Neutral, Negative	TV/OV	k-th trial: TV(20s) $\rightarrow$ OV(20s) $\rightarrow$ TV(20s) $\rightarrow$ OV(20s)	10	10-fold	TV: acc=96.07% OV: acc=99.54%
ours	Positive, Neutral, Negative	TV/OV	k-th trial: TV(60s) $\rightarrow$ OV(60s) k+1-th trial: OV(60s) $\rightarrow$ TV(60s)	16	LOTO	TVEP: acc=40.22% TVLP: acc=57.55% OVEP: acc= 51.49% OVLP: acc= 50.54%

[1] The abbreviations, TV, TO and OV, denote the traditional video, traditional odor, and olfactory-enhanced video, respectively.
[2] $\rightarrow$ denotes that the next stimulus pattern, for instance the OV(30s) pattern, is tightly followed with the previous pattern, i.e., the TV(30s).

the video content, subjects' QoE levels during the later period of the video will be obviously stronger than in the early period.

In the current work, we developed a new experimental paradigm that could dynamically enable the odors to evoke subjects' emotions synchronously during different periods when they were stimulated by video clips. To the best of our knowledge, there is no other experimental paradigm dynamically adds odor to different stages of video-evoked emotion trials. According to the t-test, no matter which stimulation stage was chosen (early or later period of a trial), the olfactory-enhanced videos always significantly outperformed the early stage of traditional videos, while it was not significantly different from the later period of traditional video (shown in Table III). These results suggest that the olfactory enhancement of emotion primarily takes effects during the early period of a stimulus trial, with no significant enhancement at the later period. Moreover, as shown in Table IV, the stimulus of olfactory-enhanced videos generally outperformed that of traditional video, where the result is consistent with that of [25] and [24].

B. Efficiency for Classifying Different Emotions

In general, most people typically prefer emotions which feel pleasant, and avoid those feel unpleasant [43], [44], [45]. Moreover, a resent sampling survey over 246 U.S. adults had reported that participants empathize with positive emotions 3 times as frequently as with negative emotions [46]. Indeed, these conclusions are consistent with the results of the EEG-based emotion recognition. Actually, most EEG-based experimental paradigms usually adopt traditional videos to evoke participants' emotions, which means that the participants need to understand the video contents and then empathize with the emotions conveyed by the videos. Therefore, it is certainly that the affective classify models usually achieved higher accuracies on recognizing positive emotions, followed by the neutral emotions and last the negative emotions [17], [28]. Therefore, improving the classify accuracy in response to the negative emotions is significant for building an affective HCI system.

Comparing with the traditional videos, the olfactory-enhanced videos provide more potential possibility. That is to say, the pattern of olfactory-enhanced videos offers more discriminant information to improve the classify accuracy for different emotion states, especially the negative emotions, in an emotional recognition task. The average classify accuracies in response to the three emotional states over the four classifiers was displayed in Fig. 6. Obviously, our experiment further approved the conclusion that the traditional video would more easily evoke subjects' positive emotions than negative emotions. Furthermore, as shown in Fig. 6, the olfactory-enhanced video stimuli always achieved higher accuracies in recognizing negative emotions than the traditional video stimuli during the same stimulus stage of a trial. However, for positive emotions, there existed a relatively big difference for the olfactory-enhanced video stimuli in different stimulus stages, namely the OVEP pattern achieved higher accuracies than the OVLP pattern, nevertheless, the reason cause this problem is ambiguous. We think that one of the possible reasons is that the odors we chose didn't exactly match the video content. In other words, the positive video contents were mainly about comedy for comedy could more easily make an individual happy or excited, while the main role of positive odors is to be relaxing or pleasant. More concretely, in the early stage of a stimuli trial, because of the lack of understanding about the video content, subjects' positive emotions were mainly evoked by positive odors; While in the later stage of a stimuli trial, with the deepened understanding of the video content, subjects' were more easily be happy or excited, but due to the simultaneous stimulation of the odor, the positive emotions may be more smoothing or even distracted. Finally, for the neutral emotions, the results were consistent with our expectation, since during the different stages of the stimulus trial, the two kinds of stimulus patterns, namely traditional video, and olfactory-enhanced video, performed similarly.

C. Neural Pattern for Different Stimulus Patterns

The beta and gamma bands had been reported to show higher activation responses in previous studies, and play an important role in EEG-based emotion recognition [17], [28]. In our study, we observed that compared with the delta, theta, and alpha bands, the beta and gamma bands were more active in the bilateral temporal cortex regardless of the stimulus patterns, which further demonstrated that the high frequency EEG signals reflect more emotional information [47], [48].

The EEG pattern in response to emotions had been indicated to be provided with stability over time [31]. In our experiment,

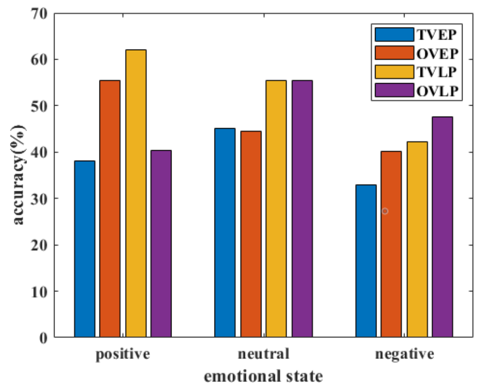


Fig. 6. The average accuracies of recognizing three emotions over the four classifiers for the four stimulus patterns in the gamma band.

we found that during the stimulation by same kind of stimulus mode (i.e. traditional video or olfactory-enhanced video), the topographical maps of DE features shown in Fig. 4 were similar over the different stage of a trial. Besides, during the same stage of a trial, the topographical maps also showed similarity over the different kinds of stimulus mode.

Moreover, we observed that for all the five frequency bands, the activities of pre-frontal lobe (Fp1, Fp2) for traditional video were significantly stronger than the olfactory-enhanced video in general according to the Table V. When stimulating by odors, subjects' olfactory attention will significantly increase, which may further increase their visual attention because of the multi-sensory integration among the cross-modal stimuli [7]. In addition, a study has also demonstrated that odors can affect visual processing by attracting attention more faster under the cases where the odor matches the video content [49]. Therefore, in our experiment, under the stimulation of olfactory-enhanced videos, the frequency of subjects' eye activity, such as eye movement and blink, may relatively decrease, and thus the activities of the pre-frontal lobe were weaker than the stimulation of traditional videos, for the pre-frontal lobe is the brain region most strongly affected by EOG artifacts [50]. Furthermore, we also found that there existed a significant difference for the DE feature extracted from F7 on whether adding the odors as stimulation, where the result is similar to [24], in which the MUSE EEG headband (AF7, AF8, TP7, TP8) was employed to record the EEG signals, and the channel AF8 was found to show significant difference between traditional videos and olfactory-enhanced videos in the delta, theta, and alpha bands.

D. Limitations and Future Works

Our experimental results illustrated the efficiency of the designed emotional BCI paradigm. Nonetheless, the current study still has some limitations. Firstly, The age range of the participants was concentrated between 20 and 30 years old, with at least a bachelor's degree. In the future, we will extend the proposed paradigm to subjects with a wider age range and other educational background, to study the universality of olfactory-enhanced videos on evoking individuals' target emotions. In addition, our experimental results preliminary showed the possibility of odors for regulating the emotions on the emotional stimulation experiment with traditional videos as stimulus materials, while sufficient evidence is lacking to confirm this possibility. Hence, we will further adopt odors with different stimulus intensity by controlling the strength of odors, and study the effectiveness of odor participation in the problem of emotion regulation. Moreover, the respective evoked efficiency of odors and videos are not clear when both the two kinds of materials are involved in emotional evoking experiments. Therefore, a more detailed and reasonable experiment paradigm needs to be developed to investigate the primary and secondary aspects of odor and video in the efficiency of inducing emotions. Finnally, the reaserch about brain network is also an important work to study the interaction of different brain areas when subjects are stimulated by traditional video or olfactory-enhanced video.

VII. Conclusion

Stimulating individuals' multi-sensory will evoke their emotions more efficiently. In this paper, we developed a novel EEG emotional experiment paradigm which allow odors dynamically evoke participants' emotion during the early or later term of a video stimulation, and we built an EEG-EOG emotional dataset which is publicly available. From the experimental results, we have found that regardless of the period, olfactory-enhanced videos were significantly better than the early term of traditional videos in statistics, while slightly worse than the later term of traditional video in classifier performance. In addition, compared with the traditional video stimulus, the olfactory-enhanced video stimulus had achieved higher performance in recognizing negative emotions, which partly made up for the weak ability of the traditional video stimulus to identify negative emotions. Furthermore, we found that topographical maps with different stimulus patterns didn't show obviously differences, which indicated that, to a certain extent, the neural pattern in response to emotion changes is stable when stimulating individuals' different senses. Besides, we also observed that the olfactory-enhanced video stimuli and traditional video stimuli differ significantly on the Fp1, Fp2, and F7 channels in almost all five frequency bands.

Acknowledgment

The authors declare that they do not have any financial or personal relationships with other people or organizations that could have inappropriately influenced this study.

References

[1] R. W. Picard, "Affective computing: Challenges," *Int. J. Hum.-Comput. Stud.*, vol. 59, nos. 1–2, pp. 55–64, 2003.

[2] M. Ben and Z. Lachiri, "Emotion classification in arousal valence model using MAHNOB-HCI database," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 3, pp. 1–6, 2017.

[3] S. Jang, S.-E. Moon, and J.-S. Lee, "EEG-based emotional video classification via learning connectivity structure," *IEEE Trans. Affect. Comput.*, early access, Nov. 9, 2021, doi: 10.1109/TAFFC.2021.3126263.

[4] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from EEG-based brain–computer interface using machine learning: A comprehensive review," *Neural Comput. Appl.*, vol. 34, pp. 12527–12557, May 2022.

[5] J. Kim and E. André, “Emotion recognition based on physiological changes in music listening,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[6] S. Schachter and J. Singer, “Cognitive, social, and physiological determinants of emotional state,” *Psychol. Rev.*, vol. 69, no. 5, p. 379, 1962.

[7] B. E. Stein and T. R. Stanford, “Multisensory integration: Current issues from the perspective of the single neuron,” *Nature Rev. Neurosci.*, vol. 9, no. 4, pp. 255–266, Apr. 2008.

[8] E. Mansouri-Benssassi and J. Ye, “Synch-graph: Multisensory emotion recognition through neural synchrony via graph convolutional networks,” in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 2, pp. 1351–1358.

[9] E. M. Benssassi and J. Ye, “Investigating multisensory integration in emotion recognition through bio-inspired computational models,” *IEEE Trans. Affect. Comput.*, early access, Aug. 19, 2021, doi: 10.1109/TAFFC.2021.3106254.

[10] Q. Shi, J. Fan, Z. Wang, and Z. Zhang, “Multimodal channel-wise attention transformer inspired by multisensory integration mechanisms of the brain,” *Pattern Recognit.*, vol. 130, Oct. 2022, Art. no. 108837.

[11] P. Salovey and J. Mayer, “Emotional intelligence,” *Imag., Cogn. Pers.*, vol. 9, no. 3, pp. 185–211, 1990.

[12] S. Aydin, “Deep learning classification of neuro-emotional phase domain complexity levels induced by affective video film clips,” *IEEE J. Biomed. Health Informat.*, vol. 24, no. 6, pp. 1695–1702, Jun. 2020.

[13] Y.-P. Lin et al., “EEG-based emotion recognition in music listening,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1798–1806, Jul. 2010.

[14] D. O. Bos et al., “EEG-based emotion recognition,” *Influence Vis. Auditory Stimuli*, vol. 56, no. 3, pp. 1–17, 2006.

[15] S. Koelstra et al., “DEAP: A database for emotion analysis; Using physiological signals,” *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan./Mar. 2012.

[16] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, “A multimodal database for affect recognition and implicit tagging,” *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 42–55, Jan. 2011.

[17] W.-L. Zheng and B.-L. Lu, “Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks,” *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.

[18] E. Horska, J. Bercik, A. Krasnodebski, R. Matysik-Pejas, and H. Bakayova, “Innovative approaches to examining consumer preferences when choosing wines,” *Agricult. Econ.*, vol. 62, no. 3, pp. 124–133, Mar. 2016.

[19] H. Singh et al., “The brain’s response to pleasant touch: An EEG investigation of tactile caressing,” *Frontiers Hum. Neurosci.*, vol. 8, p. 893, Nov. 2014.

[20] E. Kroupi, J.-M. Vesin, and T. Ebrahimi, “Subject-independent odor pleasantness classification using brain and peripheral signals,” *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 422–434, Oct./Dec. 2016.

[21] A. Raheel, S. M. Anwar, and M. Majid, “Emotion recognition in response to traditional and tactile enhanced multimedia using electroencephalography,” *Multimedia Tools Appl.*, vol. 78, no. 10, pp. 13971–13985, May 2019.

[22] D. Li et al., “EEG-based emotion recognition with haptic vibration by a feature fusion method,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[23] N. Ranasinghe et al., “Tainted: An olfaction-enhanced game narrative for smelling virtual ghosts,” *Int. J. Hum.-Comput. Stud.*, vol. 125, pp. 7–18, May 2019.

[24] A. Raheel, M. Majid, and S. M. Anwar, “A study on the effects of traditional and olfaction enhanced multimedia on pleasantness classification based on brain activity analysis,” *Comput. Biol. Med.*, vol. 114, Nov. 2019, Art. no. 103469.

[25] J. Xue, J. Wang, S. Hu, N. Bi, and Z. Lv, “OVPD: Odor-video elicited physiological signal database for emotion recognition,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[26] G. Prete, P. Croce, F. Zappasodi, L. Tommasi, and P. Capotosto, “Exploring brain activity for positive and negative emotions by means of EEG microstates,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, Mar. 2022.

[27] V. Zotev, R. Phillips, H. Yuan, M. Misaki, and J. Bodurka, “Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback,” *NeuroImage*, vol. 85, pp. 985–995, Jan. 2014.

[28] W. Zheng, “Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis,” *IEEE Trans. Cogn. Develop. Syst.*, vol. 9, no. 3, pp. 281–290, Sep. 2017.

[29] T. Song, W. Zheng, P. Song, and Z. Cui, “EEG emotion recognition using dynamical graph convolutional neural networks,” *IEEE Trans. Affect. Comput.*, vol. 11, no. 3, pp. 532–541, Jul./Sep. 2020.

[30] Y. Peng, F. Qin, W. Kong, Y. Ge, F. Nie, and A. Cichocki, “GFIL: A unified framework for the importance analysis of features, frequency bands, and channels in EEG-based emotion recognition,” *IEEE Trans. Cogn. Develop. Syst.*, vol. 14, no. 3, pp. 935–947, Sep. 2022.

[31] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, “Identifying stable patterns over time for emotion recognition from EEG,” *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 417–429, Jul./Sep. 2017.

[32] R. Jenke, A. Peer, and M. Buss, “Feature extraction and selection for emotion recognition from EEG,” *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul. 2014.

[33] Z. Liang et al., “EEG entropy measures in anesthesia,” *Frontiers Comput. Neurosci.*, vol. 9, pp. 9–16, Feb. 2015.

[34] S. Aydın, S. Demirtaş, M. A. Tunga, and K. Ateş, “Comparison of hemispheric asymmetry measurements for emotional recordings from controls,” *Neural Comput. Appl.*, vol. 30, no. 4, pp. 1341–1351, Aug. 2018.

[35] B. Kılıç and S. Aydin, “Classification of contrasting discrete emotional states indicated by EEG based graph theoretical network measures,” *Neuroinformatics*, vol. 20, pp. 863–877, Mar. 2022.

[36] W. Hu et al., “EEG microstate correlates of emotion dynamics and stimulation content during video watching,” *Cerebral Cortex*, vol. 33, no. 3, pp. 523–542, Jan. 2023.

[37] M. Val-Calvo, J. R. Álvarez-Sánchez, J. M. Ferrández-Vicente, and E. Fernández, “Optimization of real-time EEG artifact removal and emotion estimation for human–robot interaction applications,” *Frontiers Comput. Neurosci.*, vol. 13, p. 80, Nov. 2019.

[38] G. K. Verma and U. S. Tiwary, “Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals,” *NeuroImage*, vol. 102, pp. 162–172, Nov. 2014.

[39] L. K. McEvoy, M. E. Smith, and A. Gevins, “Test–retest reliability of cognitive EEG,” *Clin. Neurophysiol.*, vol. 111, no. 3, pp. 457–463, Mar. 2000.

[40] J. Chen, Z. Chen, Z. Chi, and H. Fu, “Facial expression recognition in video with multiple feature fusion,” *IEEE Trans. Affect. Comput.*, vol. 9, no. 1, pp. 38–50, Jan./Mar. 2016.

[41] E. Combrisson and K. Jerbi, “Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy,” *J. Neurosci. Methods*, vol. 250, pp. 126–136, Jul. 2015.

[42] N. Murray, B. Lee, Y. Qiao, and G.-M. Muntean, “Multiple-scent enhanced multimedia synchronization,” *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 11, no. 1, pp. 1–28, Oct. 2014.

[43] R. J. Larsen, “Toward a science of mood regulation,” *Psychol. Inquiry*, vol. 11, no. 3, pp. 129–141, Jul. 2000.

[44] A. J. Elliot and T. M. Thrash, “Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals,” *J. Pers. Social Psychol.*, vol. 82, no. 5, pp. 804–818, 2002.

[45] M. Tamir, “What do people want to feel and why? Pleasure and utility in emotion regulation,” *Current Directions Psychol. Sci.*, vol. 18, no. 2, pp. 101–105, Apr. 2009.

[46] G. J. Depow, Z. Francis, and M. Inzlicht, “The experience of empathy in everyday life,” *Psychol. Sci.*, vol. 32, no. 8, pp. 1198–1213, Aug. 2021.

[47] W. Huang, W. Wu, M. V. Lucas, H. Huang, Z. Wen, and Y. Li, “Neurofeedback training with an electroencephalogram-based brain-computer interface enhances emotion regulation,” *IEEE Trans. Affect. Comput.*, early access, Dec. 13, 2021, doi: 10.1109/TAFFC.2021.3134183.

[48] W. J. Ray and H. W. Cole, “EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes,” *Science*, vol. 228, no. 4700, pp. 750–752, May 1985.

[49] A. Seigneuric, K. Durand, T. Jiang, J.-Y. Baudouin, and B. Schaal, “The nose tells it to the eyes: Crossmodal associations between olfaction and vision,” *Perception*, vol. 39, no. 11, pp. 1541–1554, Nov. 2010.

[50] T. P. Jung et al., “Removing electroencephalographic artifacts by blind source separation,” *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.