

Video-triggered EEG-emotion public databases and current methods: A survey

Wanrou Hu^{1,2}, Gan Huang^{1,2}, Linling Li^{1,2}, Li Zhang^{1,2}, Zhiguo Zhang^{1,2,3}, Zhen Liang^{1,2} (✉)

¹ School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen 518060, Guangdong, China

² Guangdong Provincial Key Laboratory of Biomedical Measurements and Ultrasound Imaging, Shenzhen 518060, Guangdong, China

³ Peng Cheng Laboratory, Shenzhen 518055, Guangdong, China

ARTICLE INFO

Received: 18 Aug, 2020

Revised: 6 Sep, 2020

Accepted: 22 Sep, 2020

© The authors 2020. This article is published with open access at journals.sagepub.com/home/BSA



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

KEYWORDS

emotion recognition, EEG signals, video-triggered, emotion database

ABSTRACT

Emotions, formed in the process of perceiving external environment, directly affect human daily life, such as social interaction, work efficiency, physical wellness, and mental health. In recent decades, emotion recognition has become a promising research direction with significant application values. Taking the advantages of electroencephalogram (EEG) signals (i.e., high time resolution) and video-based external emotion evoking (i.e., rich media information), video-triggered emotion recognition with EEG signals has been proven as a useful tool to conduct emotion-related studies in a laboratory environment, which provides constructive technical supports for establishing real-time emotion interaction systems. In this paper, we will focus on video-triggered EEG-based emotion recognition and present a systematical introduction of the current available video-triggered EEG-based emotion databases with the corresponding analysis methods. First, current video-triggered EEG databases for emotion recognition (e.g., DEAP, MAHNOB-HCI, SEED series databases) will be presented with full details. Then, the commonly used EEG feature extraction, feature selection, and modeling methods in video-triggered EEG-based emotion recognition will be systematically summarized and a brief review of current situation about video-triggered EEG-based emotion studies will be provided. Finally, the limitations and possible prospects of the existing video-triggered EEG-emotion databases will be fully discussed.

1 Introduction

Emotions are the psychological and mental states formed in the process of perceiving and

interacting with the external stimuli like environmental changes. Emotions reflect the underlying motivation and consciousness of human behaviors and have a direct impact on the

Address correspondence to Zhen Liang, janezliang@szu.edu.cn

establishment and maintenance of interpersonal relationships, cognition, decision-making, work efficiency, and other interactive activities [1]. Previous studies found that many diseases, such as depression, autism, gaming disorder, Alzheimer's, and coronary artery disease are closely accompanied by cognitive and emotional disorders [2–4]. Disrupted by uncoordinated fast pace of life and high pressure of social competition, increasing number of people have difficulty regulating their affective balance. Being under negative emotions for a long period will eventually affect our physical and mental health, resulting in a sharp decline in the quality of life and happiness index. Therefore, the study of affective computing is of great significance and real-time emotion recognition has become one of the research hotspots.

With rich media information on both visual and auditory stimuli, video-based emotion triggering approach offers a favorable technical support for realizing real-time emotion evoking and recognition in a laboratory environment. Similarly, due to the benefits from the outstanding characteristics of its high time resolution and fast transmission speed, electroencephalogram (EEG) signals have been proven to be a useful tool for human emotion decoding in the field of affective computing. Due to the numerous benefits of EEG signals and video-based emotion triggering, our paper will mainly focus on video-triggered emotion recognition using EEG signals. We will first introduce the background of emotion recognition in Section 2. Then, more details about the current available video-triggered EEG-emotion databases will be introduced in Section 3. Next, the commonly used methods in EEG signals analysis and emotion recognition modeling will be presented in Section 4 and Section 5, respectively. In Section 6, current limitations and future trends of video-triggered EEG-based emotion studies will be fully discussed.

2 Basic information about emotion recognition

2.1 Emotion model

In the existing researches, emotions are usually characterized by two types of models: discrete and dimensional emotion models.

In the discrete emotion model, researchers assume that all types of emotions can be well described by a specific subset of basic emotional states. For example, Plutchik et al. [5] claimed that there were eight basic emotions, namely anger, fear, sadness, disgust, surprise, happiness, trust, and expectation as shown in Fig. 1(A). Ekman et al. [8] also claimed that there were only six basic emotions, namely anger, disgust, fear, happiness, sadness, and surprise. However, the defined basic emotions may fail to reflect the complexity and diversity of emotional states due to its limited numbers. Also, it was found that discrete emotion model has many limitations in practical applications when quantifying the emotional type and intensity [7, 9]. On the other hand, dimensional emotion model provides a more effective way to quantify and characterize the type and intensity of emotions from multiple dimensions (e.g., valence, arousal, and dominance). Here, valence refers to the emotional type of pleasantness or unpleasantness. Arousal measures the intensity of evoked emotions, where excitement is characterized as a high arousal value while boredom with a low one. Dominance reflects the controlling or submissive nature of emotion. In current studies, the most frequently used dimensional model is the Russell's valence-arousal (VA) dimensional model [10] shown in Fig. 1(B). In this model, fear could be characterized as an emotion of negative valence and high arousal; joy could be described as an emotion with positive

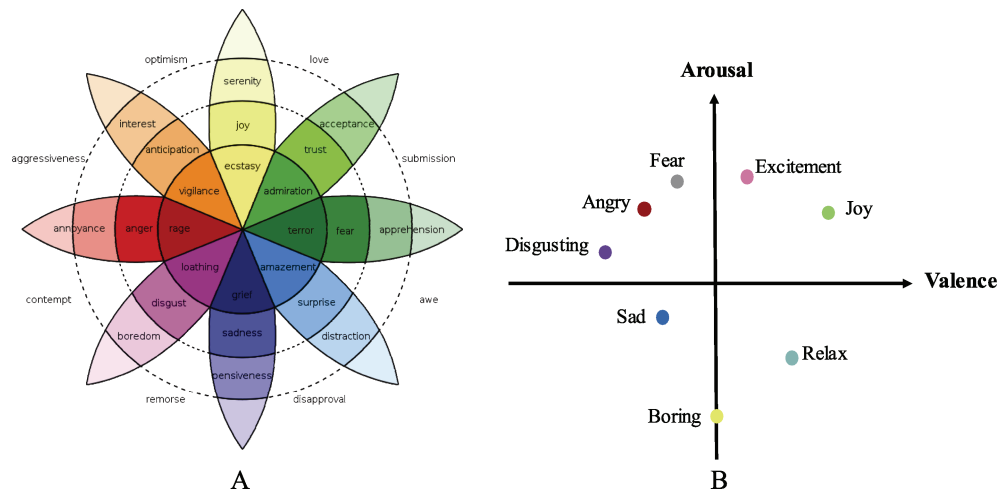


Fig. 1 Emotion models. (A) Plutchik's Wheel of Emotions. Reproduced with permission from Ref. [6], ©Springer Nature. (B) Valence-arousal dimensional model. Reproduced with permission from Ref. [7], ©MDPI.

valence and high arousal. Valence-Arousal-Dominance model is another popular dimensional model obtained by adding dominance dimension into VA model [9, 11]. It is evident that dimensional emotion model provides a quantitative expression of emotions, thereby, presenting the emotion information in a more objective and accurate way and providing benefits to emotion recognition modeling.

2.2 Emotion evoking methods

Various emotion evoking methods have been exploited in emotion recognition studies. A properly selected emotion-evoking method directly affects the quality of the collected data. The selection of appropriate stimulation materials for conducting suitable and effective emotion evoking experiments is one of the key points in emotion recognition. According to the sources of emotion evoking materials, emotion-evoking methods can be roughly divided into internal and external stimuli [12, 13]. Table 1 introduces different evoking methods used in emotion recognition studies with their corresponding experimental designs.

Internal stimuli refer to self-recall of personal experiences or self-imagination under the

guidance of experimental instructions. For example, Kothe et al. [24] carried out a 3–5 min scene imagination emotion-induced experiment. Twelve subjects (25 ± 5 years old) were in a state of deep relaxation with their eyes closed. Under the guidance of verbal instruction, they were instructed to operate scene recall or imagination to induce 12 specified types of negative emotions (e.g., anger, jealousy, depression, fear, sadness, and worry) and positive emotions (e.g., love, happiness, relief, satisfaction, and awe). Based on a sparse feature-selecting classifier, Kothe et al. achieved the recognition accuracy of 71.3% for binary valence classification. Different from other experimental designs that only relied on self-induction, Zhuang et al. [15] incorporated external video stimuli into self-recall experiments to enhance the efficiency of self-induced emotion responses. Specifically, 30 subjects comprising of 20 males and 10 females between 18 and 35 years old recalled specific scenes of each film with their eyes closed after watching emotion-induced videos. Specific emotions were triggered, such as happiness, neutral, sadness, disgust, anger, and fear. Zhuang et al. established a cross-participant emotion classifier based on support vector machines (SVM) and achieved mean accuracy of 54.52% for classification on 6 discrete emotions

Table 1 Experiment protocol of emotion evocation methods.

Trigger	Stimuli	Experimental protocol	Ref.
Internal stimuli	Recall or imagination	Displayed instruction pictures with positive, negative, or neutral hints randomly and recalled relevant events within 8 s for triggering the emotion shown on the hint.	[14]
	Specific scenes recall	Watched a 1–3 min movie clip and recalled specific scenes within 1 min to evoke specific emotions, such as happiness, neutral, sadness, disgust, anger, or fear.	[15]
External stimuli	180 pictures from IAPS	Took a 4 s rest before each experiment, and then displayed 1.5 s emotional pictures randomly. There was a 0.5 s rest between any two experiments.	[16]
	30 pictures from IAPS	Displayed emotional pictures for 5 s in a random sequence and triggered neural, positive, and negative emotions with high arousal. Valence and arousal ratings were evaluated by self-report using Self-Assessment Manikin (SAM) scale.	[17]
	16 original soundtrack music from Oscar movies	Played 16 soundtrack music for 30 s in a random sequence. Participants reported their emotions in terms of valence and arousal and gave the emotion label, such as joy, anger, sadness, and pleasure, to the present soundtrack music.	[18]
	40 movie sound (emotions to be triggered: joy, happiness, tenderness, vitality, sadness, fear, anger, and tension)	Participants listened to randomly played music for 12 s and reported their emotion state by answering 8 random questions from Likert.	[19]
	40 music videos with 1 min duration for each video	Participants watched 1-min music videos and reported their emotion level in terms of valence, arousal, dominance, and liking on SAM.	[20]
	72 movie clips with 2 min duration for each video	Used 72 movie clips to trigger happy, sad, fearful, and neutral emotions, which were presented randomly. Participants reported their emotion state in terms of valence and arousal using a 10-point scale SAM.	[21]
	18 movie clips: emotions to be triggered are fear, anger, disgust, sadness, calm, joy, excitement, entertainment, and surprise	Started with a neutral video as the baseline, and played movie clips randomly. After each movie, participants reported their emotions in terms of valence, arousal, and dominance using a 5-point scale SAM.	[22]
15 Chinese movie clips that reflect positive, negative, and neutral emotions	After watching each 4-min video, participants answered questions from Philippot questionnaire: 1) how they felt after video watching; 2) whether they had seen the video before, and 3) whether the video content was easy to understand.	[23]	

(i.e., happiness, neutral, sadness, disgust, anger, and fear). An increased cross-participant classification accuracy of 87.36% was achieved for binary emotion recognition (i.e., positive emotions: happiness; negative emotions: sadness, disgust, and anger).

However, due to the individual differences in age, culture, habits, growth experience, person-

ality, and emotional perception, there are still many limitations of internal stimuli in emotion evoking. Also, it was found that emotions evoked by internal stimuli were mostly negative or mixed emotions [25], hence, it would be difficult to ensure that a specific emotion was accurately triggered.

External emotion stimuli mainly utilize

materials like pictures, music (or sounds), and videos for emotion evoking. Since pictures or music can only provide participants with visual or auditory stimuli, the intensity of evoked emotion is limited. On the other hand, videos provide much richer stimulation information with both visual and auditory stimuli, which could be more suitable for triggering specific emotions in a more efficient, accurate, and vivid approach in a laboratory environment [23, 26]. Ellard et al. [27] verified the effectiveness of emotion triggering experiments induced by pictures, music, and movie clips in a laboratory environment. With the ANOVA results of different self-evaluation indicators like Self-Assessment Manikin (SAM) scale, Positive and Negative Affect Schedule (PANAS), and Personal Relevance Scale, Ellard et al. found that video stimuli resulted in better performance of emotion induction in a laboratory environment. Thus, compared with other evoking materials such as pictures and music (or sounds), video-based external stimuli provide useful technical supports for effective emotion-inducing in the laboratory and help participants to have a stronger sense of substitution and rapidly respond to the triggering materials.

2.3 Emotional information acquisition

Both computer technology and emotion recognition algorithms could be used to develop an emotional human–computer interaction system for real-time emotion recognition and regulation. The decoding of emotion information in such human–computer interaction system is mainly based on subjective self-report from participants (i.e., self-perception of the emotional state), external expressions (namely tone, volume, facial expression, gesture, etc.) [28–30], and internal expression (i.e., spontaneous physiological signals) [31]. Here, subjective self-report refers to individual descriptions and ratings. However, it

is difficult to guarantee the validity and accuracy of subjective self-report as participants may fail to accurately describe their subjective feelings and there may exist individual differences in perceptions of emotions as well [28, 32]. There are similar consciously controlled, restrained, or disguised problems in external behavioral expressions. Due to their low validity and authenticity, external behavioral expressions provide conscious and indirect emotional information [33] and show obvious individual differences influenced by gender, education level, and cultural differences [8]. Different from the weak generalization of subjective experience and external behavioral data, spontaneous physiological signals have relatively higher consistency across cultures and countries [34], which cannot be disguised or restrained consciously. They could be considered as the reflection of realistic emotional information. Therefore, emotion recognition based on spontaneous physiological signals has become an important direction in affective computing [35].

Spontaneous physiological signals can be used to evaluate the activities of the central and autonomic nervous systems and to provide more objective and effective detection of emotional states from the perspective of internal physiology [36, 37]. For example, electrocardiogram (ECG) reflects the activity of the myocardial autonomic nervous system in different emotional states [38]. Galvanic skin response (GSR) measures arousal of emotion by measuring the changes of skin electrical property along with the activity of autonomic nervous systems [39]. Electromyography (EMG) is used to measure the degree of muscle tension in different emotional states by recording bioelectric changes collected on the skin surface [12]. Respiratory signals (RSP) contain a wealth of emotional information and it has been shown that breathing rate and respiratory depth change with different emotions

[40]. However, due to the problems of low classification accuracy, lack of reasonable quantitative standards and indirect relationship between emotional states and peripheral physiological signals, it is very difficult to accurately quantify emotions using the spontaneous physiological signals mentioned above [41, 42].

Among the spontaneous physiological signals, EEG signals provide a more direct, comprehensive, and objective approach for emotional recognition from a neurophysiological perspective by measuring the spontaneous physiological activities of cerebral cortex under different emotion states [43–45]. In comparison with other neuroimaging technologies, such as functional magnetic resonance imaging (fMRI), magnetoencephalogram (MEG), and positron emission tomography (PET), EEG signals have better time resolution, fast data collection and transmission, which could be easily applied to real-time emotion recognition system [43, 44, 46]. Moreover, EEG acquisition is low cost, non-invasive, and easy to use; there are already different types of wireless portable EEG recording devices developed for various practical and clinical applications. Thus, EEG signals are recommended as a useful tool in emotion

recognition and online human–computer interaction system [47].

2.4 Emotion-related neurophysiological mechanism

Division of brain regions is shown in Fig. 2(A) where each brain region has its dedicated function [49, 50]. Specifically, frontal cortex takes part in conscious thinking activities, language expression, and emotional control; it is the functional area of mental activities. Temporal lobe is responsible for the sense of smell and hearing; it also handles complex stimuli, such as facial recognition, scene recognition, and memory. Participating in the process of attention, body language, and skill learning, the parietal lobe integrates sensory information from various functional systems and plays an important role in the manipulation of objects. The occipital lobe is the visual function area for integration and processing of visual information; it also participates in the process of shape, color, and movement perception. During emotion characterization, particular neural coordination pattern occurred under different emotional states [51, 52]. Schmidt et al. [53] found that left frontal brain region was active under the stimulation of cheerful music, while sad and scary music activated

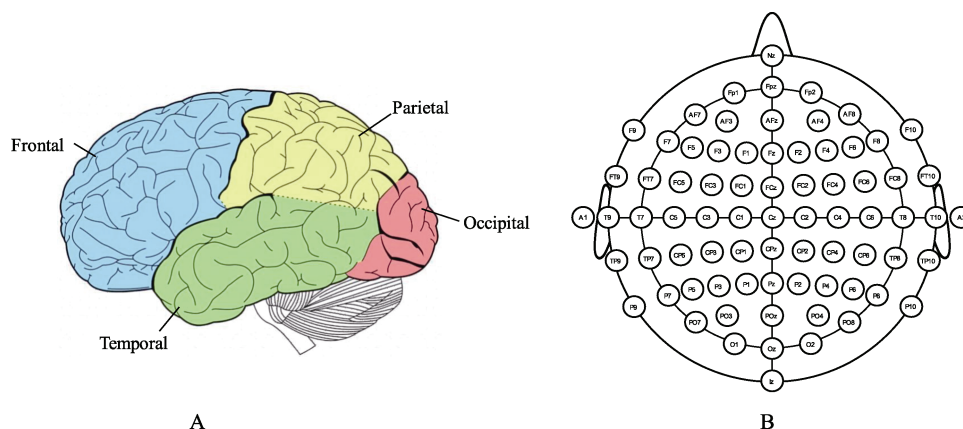


Fig. 2 Brain neurophysiological mechanisms. (A) Brain regions. Reproduced with permission from Ref. [48], ©Springer Nature. (B) 64-electrode system using international 10–20 standard. Reproduced with permission from http://www.mariusthart.net/downloads/eeg_electrodes_10-20.svg, ©Marius 't Hart - <http://www.beteredingen.nl>.

the activities in the right frontal brain region. Schutter et al. [54] found that when participants watched facial expression pictures with angry emotions, the activities in the right parietal brain region increased. Therefore, by detecting the activities in different brain regions, the corresponding physiological response of brain activities in different emotional states can be well characterized.

EEG is a non-invasive monitoring technique for electrophysiological activity of the brain. Using multiple electrodes placed on surface of the scalp [as shown in Fig. 2(B)], the frequency and amplitude of spontaneous electrical activity of cerebral cortex are recorded, and the changes in brain electrophysiology are detected. EEG signals contain rich brain electrophysiological information which are significantly important for the direct and accurate interpretation of emotion from the perspective of neurophysiology [55, 56]. The oscillation characteristics contained in the EEG signals and the rhythms of brain activities in specific frequency bands play a guiding role for better understanding the brain activities such as emotion induction, thinking, cognition, and memory formation [57–60].

- **Delta (1–4 Hz):** Related to subconscious activities, it mainly occurs during deep sleep (i.e., slow-wave sleep) without dreams. Delta waves can be used for sleep staging. Related studies have found that cognitive decline in patients with Alzheimer's disease or epilepsy may be related to abnormal slow-wave sleep [61, 62].

- **Theta (4–7 Hz):** This mainly occurs during subconscious activities (such as sleep, dreaming, meditation) [63]. Similar to the Delta wave, Theta wave can also be used for sleep staging, mainly in the rapid eye movement (REM) sleep period of the healthy [64]. Also, it contributes to the processes of learning, cognition, and memory formation [65].

- **Alpha (8–13 Hz):** This is mainly distributed in the parietal and occipital lobe when people are in relaxed mental state with clear consciousness [66]. There is an opposite relationship between power in Alpha band and the intensity of brain activities, thus, an increase in the activation of the brain activities results in the decrease in Alpha power [47]. Studies have shown that the increased activation of Alpha wave in the right frontal lobe is related to the appearance of negative emotions, such as fear and disgust, while the activation of Alpha band in the left frontal lobe is related to the induction of positive emotions, such as joy and happiness. Researchers have demonstrated that emotions could be effectively recognized using the asymmetry features of the prefrontal Alpha band [67, 68].

- **Beta (13–30 Hz):** This is related to the active mental state that can be used to evaluate cognitive activities and emotional states [58]. When participant is in a state of concentration and thinking, the activities of Beta wave in frontal and occipital cortex will be activated [69, 70].

- **Gamma (> 30 Hz):** This is associated with active brain states and cognitive activities, such as memory, attention, and perception, and occurs in the process of information integration. The asymmetry of Gamma band in temporal lobe and parietal lobe has also been found to be useful in identifying emotional states. When in a negative emotion, the powers of Gamma band in the left temporal lobe are activated, whereas the powers of Gamma band in the right temporal lobe are activated when in a positive emotion [71–73].

3 Video-triggered emotion recognition with EEG signals

3.1 Available video-triggered EEG-based databases for emotion recognition

With the rapid development of emotion recogni-

tion, a series of standardized emotion-triggering databases with pictures and music materials have been established with the corresponding emotion labels provided by psychologists. They include the International Affective Picture System (IAPS) [74], Nencki Affective Picture System (NAPS) [75] and International Affective Digital Sound System (IADS) [76]. However, current studies with the video-based emotion-evoking approach are still limited without a generally accepted standard. In this paper, a survey on the available video-triggered EEG-based databases for emotion recognition will be conducted to provide instructive guidance on the establishment of video-based emotion databases as well as choosing a proper dataset for future emotion recognition researches.

In order to promote the progress of emotion recognition, a number of video-triggered EEG-emotion databases have been developed. As benefits of these publicly available databases, the emotion classification performances using different algorithms or classified models could be verified in a more standard approach. Here, a brief summary of available video-triggered EEG-emotion databases is presented in Table 2, including details of the number of participants, video duration, self-assessment method, and triggering emotion types. Currently, the most commonly used video-triggered EEG-emotion databases are Database for Emotion Analysis using Physiological Signals (DEAP), MAHNOB-HCI, and SJTU Emotion EEG Dataset (SEED) series EEG databases (i.e., SEED, SEED-IV, SEED-VIG).

- **DEAP database:** This is a publicly available EEG-based and peripheral physiological signals-based emotion recognition database established by Professor Ioannis Patras's research team. This database simultaneously recorded EEG signals and peripheral physiological signals (e.g., ECG, EMG, GSR) when 32 participants were watching

40 one-minute music videos. SAM scale was used to collect self-assessment of emotion in valence, arousal, dominance, and liking dimensions [20].

- **MAHNOB-HCI database:** Soleymani et al. collected a total of 27 participants' EEG signals and other peripheral physiological signals when they were watching 20 emotion-inducing movie clips. Self-assessment information of arousal, valence, dominance, predictability, and emotion labels were recorded by SAM scale [77].

- **SEED series database:** This public EEG series dataset was established by the team led by Professor Lu Baoliang from Shanghai Jiao Tong University, which contains 3 sub-databases: SEED, SEED-IV, and SEED-VIG.

(1) SEED database: Emotion-related EEG signals from 15 participants were recorded. 15 Chinese movie clips were used for positive, neutral, and negative emotions triggering [82].

(2) SEED-IV database: This database contains simultaneously recorded EEG signals and eye gaze data for emotion decoding. This database used a 6-electrode wearable wireless Emotion Meter hardware to record the EEG signals from 15 subjects when watching movie clips. The evoked emotions include happiness, neutrality, sadness, and fear. The simultaneously recorded eye gaze data can be used to interpret the emotion state from external searching behavior. Moreover, combining EEG and eye movement signals with different modality fusion strategies will be of great value for emotion recognition [83].

(3) SEED-VIG database: This experiment recruited 23 participants and asked them to monotonously repeat straight road driving tasks in a simulated driving system, which could easily induct fatigue and change their vigilance [84]. Then, the vigilance level of each participants was evaluated based on the simultaneously collected EEG signals, EOG signals,

Table 2 Details of the existing video-triggered databases for EEG-based emotion recognition.

Database	Subjects (F/M)	No. of videos (duration)	Self-assessment	Evoked emotions	EEG equipment (No. of electrodes)	Recorded signals
DEAP ¹ [71]	32 (16/16)	40 (1 min)	SAM [9-point scale]: valence, arousal, dominance, and liking	LALV, HALV, LAHV, and HAHV	Biosemi ActiveTwo system (32)	EEG (sampling rate: 512 Hz, downsampled to 256 Hz), EOG, EMG, GSR, BVP, temperature, respiration, and face video from 22 participants
SEED ² [23]	15 (8/7)	15 (~ 4 min)	Philippot questionnaire: emotion state, familiarity, understandable level	positive, neutral, and negative emotions	ESI NeuroScan (62)	EEG (sampling rate: 1000 Hz, downsampled to 200 Hz), EOG, and face video
SEED-IV ² [72]	15 (8/7)	72 (~ 2 min)	PANAS [10-point scale]: valence and arousal	happiness, sadness, fear, and neutral	EmotionMeter hardware (6 electrodes: FT7, FT8, T7, T8, TP7, and TP8)	EEG and Eye Gaze Data (SMI ETG eye-tracking glasses)
DREAMER [73]	23 (9/14)	18 (~ 199 s)	SAM [5-point scale]: valence, arousal, and dominance	amusement, surprise, excitement, happiness, calmness, anger, disgust, fear, and sadness	Emotiv EPOC (16 electrodes)	EEG (sampling rate: 128 Hz), ECG (SHIMMER™ wireless sensor)
MAHNOB-HCI ³ [77]	27 (16/11)	20 (~ 81.4s)	SAM [9-point scale]: emotion label/tag, arousal, valence, dominance, and predictability	amusement, joy, neutral, sadness, fear, and disgust	Biosemi ActiveTwo system (32)	EEG (sampling rate: 256 Hz), ECG, GSR, skin temperature, respiration amplitude, Eye Gaze Data (Tobii X120), and facial expressions
HR-EEG4EMO ⁴ [78]	27 (5/22)	13 (40 s–6 min)	Valence: negative (-1), neutral (0), positive (1)	tenderness, amusement, anger, sadness, disgust, fear, and neutral	EGI system (257)	EEG (sampling rate: 1000 Hz, downsampled to 250 Hz), ECG, GSR, SpO ₂ , respiration, and pulse rate
RCLS ⁵ [79]	14 (8/6)	15	—	happy, sad, and neutral	ESI NeuroScan (64)	EEG (sampling rate: 1000 Hz) and EOG
MPED [80]	23 (13/10)	28 (122–295 s)	PANAS [5-point scale] and SAM [9-point scale]: arousal and valence DES [9-point scale]: 10 basic emotions	joy, funny, anger, fear, sadness, disgust, and neutral	ESI NeuroScan (62)	EEG (sampling rate: 1000 Hz), ECG, GSR, RSP
CAS-THU ⁶ [81]	30 (0/30)	16 (60–139 s)	PANAS [5-point scale]: positive, negative SAM [9-point scale]: arousal and valence DES [9-point scale]	amusement, tenderness, joy, neutral, sadness, disgust, anger, and fear	Emotiv EPOC (14 electrodes)	EEG (sampling rate: 128 Hz)

(Continued)

1. [http://www.eecs.qmul.ac.uk/mmv/data sets/deap/](http://www.eecs.qmul.ac.uk/mmv/data%20sets/deap/)
2. <http://bcmi.sjtu.edu.cn/~seed/index.html>
3. <http://mahnob-db.eu>
4. <http://www.technicolor.com/en/innovation/scientificcommunity/scientific-data-sharing/eeg4emo-dataset>
5. <http://aip.seu.edu.cn>
6. <http://enginpsych.psych.ac.cn/indexen.php>

LALV, low arousal low valence; HALV, high arousal low valence; LAHV, low arousal high valence; HAHV, high arousal high valence; EEG, electroencephalogram; ECG, electrocardiogram; EOG, electrooculogram; GSR, galvanic skin response; EMG, electromyogram; RSP, respiration; SpO₂, pulse oxygen saturation; BVP, blood volume pressure.

and eye gaze data.

Due to the benefits of these publicly available emotion databases, video-triggered emotion recognition has received extensive attention. For example, Tripathi et al. [85] compared the performance differences between simple deep neural networks and convolutional neural networks in emotion classification based on the EEG signals from DEAP database. The recognition accuracies of simple deep neural network were 75.78% for valence and 73.13% for arousal. Whereas, the corresponding accuracies of valence and arousal increased to 81.41% and 73.36% respectively when the convolutional neural network was adopted. Liang et al. [86] proposed an EEG-based emotion decoding system using the hypergraph theory and verified the effectiveness of emotion recognition on DEAP database. More, it's an unsupervised learning method for emotion-related EEG features extraction and multi-dimensional emotion recognition. In addition, Chai et al. [87] proposed a new unsupervised learning method, known as

Subspace Alignment Auto-encoder (SAAE), and verified the emotion recognition performance on SEED emotion database by comparing with other classification algorithms, such as SVM, logistic regression (LR), Geodesic Flow Kernel (GFK), Transfer Sparse Coding (TSC), Transfer Component Analysis (TCA), Auto-encoder (AE), and Subspace Alignment without Auto-encoder (SA). In cross-subject emotion classification, SAAE algorithm obtained the best accuracy (i.e., 77.88%) for positive, neutral, and negative emotions recognition. The recognition accuracies of SVM, LR, GFK, TSC, TCA, AE, and SA algorithms were respectively 53.06%, 55.64%, 45.19%, 69.69%, 73.82%, 60.46%, and 74.74%.

3.2 A general video-triggered emotion recognition process

Generally, a video-triggered EEG-based emotion recognition pipeline mainly includes four parts (as shown in Fig. 3).

(1) Emotion evoking and simultaneous EEG signals recording. A proper selection of evoking

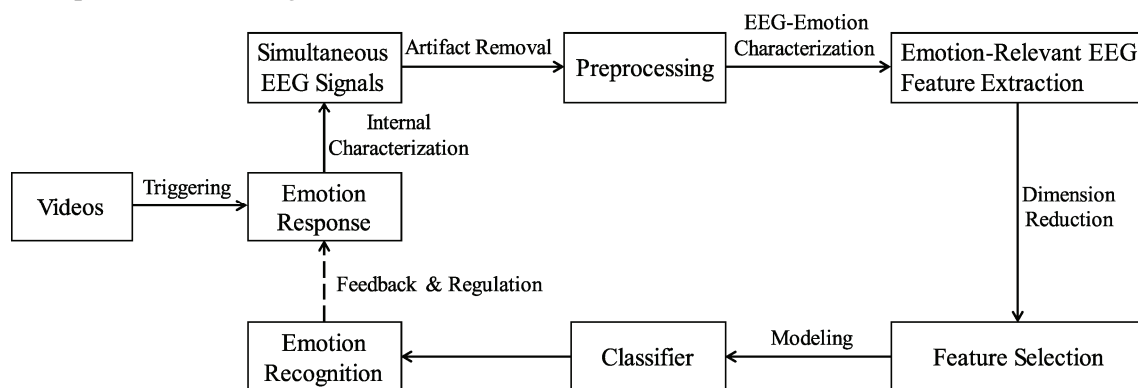


Fig. 3 A standard processing pipeline for video-triggered emotion recognition using EEG signals.

materials will directly affect the quality of evoked emotions and the collected EEG signals. In this paper, we compare the experimental effects with different emotion evoking methods and find that video-based emotion evoking could achieve the best performance; it can effectively trigger accurate and specific emotion responses in the laboratory environment. Following, we focus on video-triggered based emotion recognition and introduce the details of available public video-triggered EEG-emotion databases with full discussions.

(2) EEG signals preprocessing. The amplitude of EEG signal measured from adult scalp is in the range of 10–100 μV [88], which is very susceptible to physiological signals such as ECG, EMG, and EOG. Also, slight body movements, power-line interference, and baseline drift could bring serious artifact into the recorded EEG signals, resulting in bad signal to noise ratio. Therefore, it is necessary to preprocess the collected signals for artifact removal.

(3) Emotion-related EEG feature extraction and selection. In Section 4, we will introduce the commonly used EEG features in video-triggered EEG-based emotion recognition. Generally, EEG features are extracted from time domain, frequency domain, and time-frequency domain along with pair-wise electrodes features and connectivity features. By characterizing EEG information in multiple dimensions, emotion-related brain activities can be accurately captured. Then, EEG feature selection based on supervised or unsupervised methods should be applied to further select highly relevant EEG features for emotion recognition modeling.

(4) Classifier building for emotion recognition. The emotion classifier is established based on the EEG features after feature extraction and selection. In Section 5, we will mainly introduce the commonly used supervised, unsupervised,

and semi-supervised learning classifiers in emotion recognition with public video-triggered EEG-based databases.

Therefore, commonly used feature extraction, feature selection, and modeling methods in video-triggered emotion recognition system with EEG signals will be presented with full details, providing a brief review of the current situation about video-triggered EEG-based emotion studies. Then, the shortcomings and future possible improvements of the existing video-triggered databases for EEG-based emotion recognition will be discussed.

4 EEG feature extraction and selection methods

In this section, more details about typical EEG feature extraction and selection methods used in emotion recognition will be reported (as shown in Table 3 and Fig. 4).

4.1 EEG feature extraction

EEG features are commonly extracted from time domain, frequency domain, and time-frequency domain. By computing the asymmetrical brain activity distribution of pair-wise electrodes, interconnections between brain regions are used to represent the spatial characteristics of brain activity under different emotions. Recently, researchers also found that connectivity features showed promising performance in emotion decoding. A list of the typical EEG signal feature extraction methods in video-triggered emotion recognition research are exhibited in Fig. 4.

4.1.1 Time domain feature extraction

Time-domain EEG features represent the amplitude changes over time. Here, the pre-processed EEG signals are denoted as $x(n)$, $n = 1, 2, \dots, N$, where N is the total number

Table 3 Commonly used EEG feature extraction and selection methods for video-triggered emotion recognition.

Ref.	Database (stimuli)	EEG features	Estimation of frequency power	Electrodes	Feature selection
Koelstra et al. [20]	DEAP (music videos)	Spectrum power and DASM	Welch's method with windows of 256 samples	32	LDA (supervised)
Soleymani et al. [77]	MAHNOB-HCI (movie clips)	Spectrum power and DASM	—	32	—
Zheng et al. [23]	SEED (Chinese movie clips)	DE, DASM, and RASM of DE and PSD	STFT method with non-overlapping 1s Hanning windows of 256 samples	62	linear dynamic system (unsupervised)
Zheng et al. [21]	SEED-IV (movie clips)	PSD and DE	STFT method with non-overlapping windows of 4 s	8	—
Kroupi et al. [89]	Koelstra et al. database (music videos)	PSD, normalized 1st difference, and NSI	Welch's method with windows of 128 samples	32	—
Liu et al. [90]	DEAP (music videos)	Time domain: mean, standard deviation, 1st difference, 2nd difference, HOC, FD, Hjorth features, NSI, etc. Frequency domain: PSD Time-Frequency domain: energy, root mean square, and entropy Electrodes features: DASM, RASM, magnitude-squared coherence estimate, MI, and PCC	—	32	mRMR (supervised) PCA (unsupervised)
Atkinson et al. [91]	DEAP (music videos)	Statistical features (such as median, standard deviation, kurtosis symmetry), band power, Hjorth features, FD	Band pass frequency filter	32	mRMR (supervised)
Liang et al. [86]	DEAP (music videos)	Time domain: power, mean, standard deviation, 1st difference, 2nd difference, Hjorth features, FD, etc. Frequency domain: PSD Time-Frequency domain: energy and Shannon entropy extracted from the detail coefficients of level 4 to 7	Hamming window with 50% overlap	32	KPCA, unsupervised discriminative feature selection (unsupervised)

DASM, differential asymmetry; LDA, linear discriminant analysis; DE, differential entropy; RASM, rational asymmetry; STFT, short time fourier transform; PSD, power spectral density; NSI, non-stationary index; HOC, higher order crossings; FD, fractal dimension; MI, mutual information; PCC, Pearson correlation coefficient; mRMR, maximum relevance minimum redundancy; PCA, principal components analysis; KPCA, kernel principal component analysis.

of EEG samples. The commonly used time-domain based feature extraction methods are listed as follows.

(1) Event related potentials (ERPs): By recording the changes of brain response potentials over time, ERPs have good time resolu-

tion and provide the time-locked relationship with the stimulation events [92]. ERPs have been used to decode the process of emotion inducing and cognition [92, 93]. Martini et al. [94] found that both P300 and late positive potential increased when participants were stimulated by

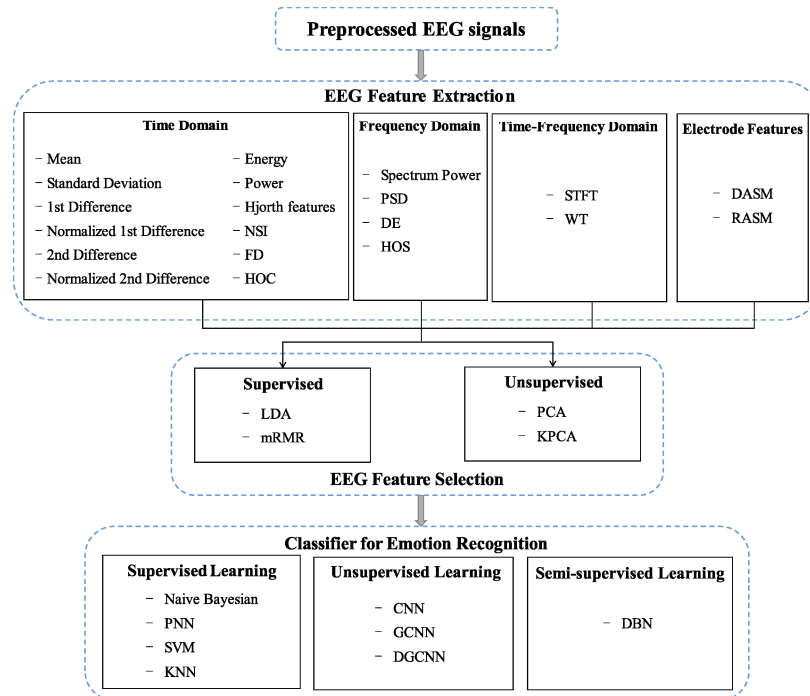


Fig. 4 Commonly used EEG analysis methods in video-triggered emotion recognition.

negative pictures compared to when they were stimulated by neutral pictures. Soroush et al. [95] used pictures from IAPS database for positive and negative emotion evoking. Their experiment found that the P200 and P300 components collected at parietal and occipital lobe could effectively decode emotions in valence dimension.

However, in the video-triggered emotion recognition, it is difficult to determine and control the emotion-inducing time point because of individual differences. Thus, the ERPs feature is not suitable for video-triggered emotion recognition.

(2) Statistical features:

1st difference:

$$\delta_x = \frac{1}{N-1} \sum_{n=1}^{N-1} |x(n+1) - x(n)| \quad (1)$$

Normalized 1st difference: This is also known as normalized length density, which can measure the self-similarity of EEG signals [89]. Specifically, when the standard deviation is

defined as $\sigma_x = \sqrt{\frac{1}{N} \sum_{n=1}^N [x(n) - \mu_x]^2}$, where

$$\mu_x = \frac{1}{N} \sum_{n=1}^N x(n). \text{ Then,}$$

$$\overline{\delta_x} = \frac{\delta_x}{\sigma_x} \quad (2)$$

2nd difference:

$$\gamma_x = \frac{1}{N-2} \sum_{n=1}^{N-2} |x(n+2) - x(n)| \quad (3)$$

Normalized 2nd difference:

$$\overline{\gamma_x} = \frac{\gamma_x}{\sigma_x} \quad (4)$$

Energy:

$$E_x = \sum_{n=1}^N |x(n)|^2 \quad (5)$$

Power:

$$P_x = \frac{E_x}{N} = \frac{1}{N} \sum_{n=1}^N |x(n)|^2 \quad (6)$$

(3) Hjorth features: Hjorth et al. [96] introduced features of Activity, Mobility, and Complexity to characterize time domain pattern of EEG signals from the aspects of amplitude, time scale, and complexity. Specifically, Activity measures the deviation of EEG time series from its mathematical expectation and reflects the

dispersion degree of EEG signal amplitude. Thus,

$$\text{Activity} = \frac{1}{N} \sum_{n=1}^N [x(n) - \mu_x]^2 \quad (7)$$

Mobility calculates the ratio between the standard deviation of slope and the standard deviation of EEG time series. $x'(n)$ stands for the first derivative of EEG signals, which is numerically equal to the first difference while $\sigma[x'(n)]$ represents the variance of the first derivative. Thus,

$$\text{Mobility} = \frac{\sigma[x'(n)]}{\sigma[x(n)]} \quad (8)$$

Complexity, in the other hand, measures the number of standard slopes that occur within the time required for a standard amplitude to be generated. Hence,

$$\text{Complexity} = \frac{\text{Mobility}[x'(n)]}{\text{Mobility}[x(n)]} \quad (9)$$

(4) Non-stationary index (NSI) evaluates the transformation of the local average over time. NSI reflects the non-stationarity and complexity of EEG signals by measuring the consistency of local average values [89, 97]. The higher value of NSI signifies higher complexity of the given EEG signals. The specific calculation process is as follows. Firstly, divide the EEG signal $x(n)$ into k segments with equal length and calculate the mean of each segment $s(m)$, $m = 1, 2, \dots, k$. Then, NSI is defined as

$$\text{NSI} = \sqrt{\frac{1}{k} \sum_{m=1}^k [s(m) - \mu_s]^2} \quad (10)$$

where $\mu_s = \frac{1}{k} \sum_{m=1}^k s(m)$.

(5) Fractal dimension (FD) quantifies complexity of EEG signals [98, 99] by measuring variations of reconstructed EEG series with fixed time interval, which is expressed as $X_k^m = \{x(m), x(m+k), x(m+2k), \dots, x[m + \text{floor}(\frac{N-m}{k})k]\}$, $m = 1, 2, \dots, k$, where k represents the

reconstructed time interval. The curve length of each reconstructed EEG series X_k^m is calculated as

$$L_m(k) = \frac{1}{k} \left\{ \sum_{i=1}^{\text{floor}(\frac{N-m}{k})} |x(m+ik) - x[m+(i-1)k]| \right\} \frac{N-1}{\text{floor}(\frac{N-m}{k})k} \quad (11)$$

Then, calculate the averaged curve length $\langle L(k) \rangle$ over k subset EEG series. When curve is fractal-like with dimension satisfying $\langle L(k) \rangle \propto k^{-\text{FD}_x}$, fractal dimension FD_x of the given EEG series can be defined as

$$\text{FD}_x = -\frac{\log \langle L(k) \rangle}{\log k} \quad (12)$$

4.1.2 Frequency domain feature extraction

EEG signals record the spontaneous and rhythmic nerve potential activity of the brain. Frequency domain feature extraction can obtain more rhythm information about brain neural activity. In many researches of affective computing and neuroscience, it has been found that changes in physiological and psychological states would cause corresponding changes in EEG data at different frequency bands: Delta (1–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), Beta (13–30 Hz), and Gamma (30–45 Hz).

(1) Power spectral density (PSD) reflects the change of EEG spectral power with frequency. Since Welch's spectral density estimation method uses smoothing window for noise reduction, it has become one of the most commonly used spectrum-based estimation methods in video-triggered EEG-based emotion recognition [73, 100].

Frequency domain features not only reflect the rhythm of brain activities but also correlate with self-assessment of valence, arousal, and

dominance. Liu et al. [11] found that when participants are in a state of high dominance, the power ratio of Beta band to Alpha band collected from frontal lobe increased. Meanwhile, the power of Beta band recorded in the parietal lobe also increased.

(2) Differential entropy (DE): This is the extension of the Shannon entropy of discrete random variables to calculate the entropy of continuous random variables. It is expressed as

$$DE = -\int_X f(x) \log[f(x)] dx \quad (13)$$

Shi et al. [101] found that the EEG signals preprocessed by bandpass filtering obeyed Gaussian distribution $N(\mu, \sigma^2)$. Therefore, the calculation formula of DE can be simplified as

$$DE = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] dx$$

$$= \frac{1}{2} \log(2\pi e \sigma^2) \quad (14)$$

Their study proved that for a fixed-length EEG signal, DE was approximately equal to the logarithmic value of spectral spectrum in a certain frequency band [102]. Generally, the power of low frequency band is much higher than that of high frequency band. DE can balance such difference by taking logarithm into consideration. Hence, DE feature has the ability to distinguish EEG patterns between low frequency band and high frequency band, thus, improving the recognition accuracy [28]. Duan et al. [101] used DE in emotion recognition research for the first time and found that DE features had a better emotion classification accuracy (of 84.22%) than the tradition spectral feature (with accuracy of 76.56%). Zheng et al. [72] compared the performance of DE and PSD features in recognizing happy, sad, scared, and neutral emotions based on SEED-IV database. Using SVM with linear kernel, it was found that the recognition accuracy was much better when

the classifier was modeled by DE feature, with accuracy of 70.58%, than that modeled by PSD feature, with accuracy of 56.34%.

(3) High-order spectrum: This represents the higher order moments or cumulants of EEG signals, which has been widely utilized in phase information extraction for emotion distinguishing. Third order spectrum analysis, namely quadratic phase coupling, has promising characteristic of recognizing the nonlinear coupling between phases f_1 and f_2 .

$$\text{Bis}(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)] \quad (15)$$

where $E(\bullet)$ represents the expectation calculation. $X(f)$ is the Fourier transform of EEG series $x(n)$. * represents the complex conjugate operation.

4.1.3 Time-frequency domain feature extraction

EEG is a non-stationary signal and it is very difficult to accurately describe the changes of frequency over time using only single time domain or frequency domain information. Joint time-frequency domain analysis can dynamically reflect the changing characteristics of EEG signals over time and has been successfully used for emotion-related EEG features extraction.

(1) Short time Fourier transform (STFT): The sliding time window $\omega(n - \tau)$ with fixed length L is utilized for dynamically processing the EEG data within the time range of $[\tau - L/2, \tau + L/2]$. Fourier transform is performed on the data extracted by the sliding time window for local information analysis of non-stationary signals. Thus,

$$\text{STFT}(\tau, \omega) = \int_{-\infty}^{+\infty} \omega(n - \tau) x(n) e^{-j\omega t} dt \quad (16)$$

The length of the sliding time window in STFT typically affects the resolution in time domains and frequency domains. Information overload is

caused when the sliding window is too long, resulting in low time resolution. On the contrary, low frequency resolution occurs when the sliding time window is too short to fully maintain information. In essence, it is very important to select an appropriate window size for STFT.

(2) Wavelet transform: Once the sliding time window function is determined, its length and shape are fixed without adaptability and cannot be adjusted according to changes in the frequency. For non-stationary and slow-changing low-frequency signals, a long sliding window needs to be selected to improve the time resolution; however, for fast-changing high-frequency signals, selecting a short sliding window improves the frequency resolution.

Using the function $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$,

wavelet transform can perform time translation and scale expansion to adjust the length and shape of the sliding window according to the frequency change of EEG signals. Thus, time-frequency resolution can be improved adaptively.

Wavelet transform is an effective method that describes the underlying frequency changes over time for emotion recognition [103]. Candra et al. [104] used discrete wavelet transform to extract wavelet entropy from the DEAP database. Their research revealed that the wavelet entropy extracted from 3–12 s sliding time window yielded a recognition accuracy of 65% in the classification of valence and arousal using SVM classifier. Related research found that Daubechies fourth-order wavelet transform (db4) had the property of feature smoothing and time-frequency positioning, which was suitable for detecting changes in EEG signals [104]. Mohammadi et al. [105] applied the db4 mother wave function to discrete wavelet transform on the five pair-wise data (i.e., F3-F4, F7-F8,

FC1-FC2, FC5-FC6, and FP1-FP2) collected from DEAP database. They fed entropy and energy of Theta (4–8 Hz), Alpha (8–16 Hz), Beta (16–32 Hz), and Gamma (32–64 Hz) bands into SVM and k-nearest neighbor (KNN) models. Classification results showed that EEG signals extracted from high frequency bands, such as Beta and Gamma bands, could better classify valence and arousal.

4.1.4 Pair-wise electrodes features

Pair-wise electrode features can be used to interpret the underlying spatial distribution pattern in different emotional states [28, 82, 90]. For example, the asymmetric spatial pattern of Alpha band between left and right hemispheres of the brain is related to emotion. Negative emotions, such as fear, disgust, and sadness, produce withdrawal stimuli that activate the right prefrontal lobe and cause the decrease of Alpha band power. Positive emotions, such as happiness and excitement, produce approaching stimuli that activate the left frontal lobe and cause the decrease on the power of Alpha band [106]. Therefore, the asymmetry of power in Alpha band can be used to evaluate the emotion changes [12, 69]. Similarly, the power changes of other frequency bands also show similar asymmetric pattern under different emotional states [46, 107].

According to the spatial symmetry of electrode distribution, pair-wise electrode features can be extracted by calculating the differences and ratio features of spatial paired electrodes.

Differential asymmetry (DASM):

$$\text{DASM} = F(C_L) - F(C_R) \quad (17)$$

Differential caudality (DCAU):

$$\text{DCAU} = F(C_{\text{frontal}}) - F(C_{\text{posterior}}) \quad (18)$$

Rational asymmetry (RASM):

$$\text{RASM} = \frac{F(C_L)}{F(C_R)} \quad (19)$$

where C_L and C_R represent the symmetrically pair-wise electrodes of the left and right hemisphere, while C_{frontal} and $C_{\text{posterior}}$ represent symmetrically pair-wise electrodes of the frontal and posterior hemisphere. $F(\bullet)$ is the specified feature of the selected electrode for further calculation.

4.1.5 Connectivity features

As a complex and dynamic cognition process, emotion has been further investigated with globally coordinated information transmitting as well as functional connections and interactions within specific brain regions [108, 109]. Connectivity features from EEG signals offer a deeper insight into emotion-related decoding of neural activities, which demonstrated the outstanding performance in emotion recognition.

Typically, connectivity features are calculated using the multi-electrode EEG signals from time and frequency domains, namely Pearson correlation connectivity (PCC), mutual information (MI), and phase locking value (PLV).

(1) Pearson correlation connectivity: This measures the linear correlation between two EEG signals from different electrodes, ranging from -1 to 1 . PCC shows negative and positive connectivity relationship, where a PCC value of 0 means no linear correlation between two separated EEG time series S_i and S_j from different electrodes or brain regions. The PCC between EEG signals S_i and S_j is calculated as

$$\text{PCC}(S_i, S_j) = \frac{\sum_{k=1}^N (s_i^k - \bar{S}_i)(s_j^k - \bar{S}_j)}{\sqrt{\sum_{k=1}^N (s_i^k - \bar{S}_i)^2} \sqrt{\sum_{k=1}^N (s_j^k - \bar{S}_j)^2}} \quad (20)$$

(2) Mutual information: This quantifies the information interaction between two different electrodes EEG signals in term of entropy, defined as

$$I(S_i, S_j) = \sum p_{ab}^{S_i S_j} \log \frac{p_{ab}^{S_i S_j}}{p_a^{S_i} p_b^{S_j}} \quad (21)$$

$p_{ab}^{S_i S_j}$ is the joint probability of $p(S_i = S_i^a, S_j = S_j^b)$. When there is no correlation between EEG time series S_i and S_j , mutual information value is zero corresponding to $p_{ab}^{S_i S_j} = p_a^{S_i} p_b^{S_j}$.

(3) Phase locking value (PLV): This describes the phase synchronization of different frequency bands. PLV is a nonlinear measure of phase correlation used to characterize specific brain rhythms and couplings. PLV value is between 0 and 1 , where 0 represents no phase coupling between two time series within 0 and π , while 1 represents identical phase synchronization. Thus,

$$\text{PLV}(S_x, S_y) = \left| \frac{1}{N} \sum_{t=1}^N e^{i[\phi_x(t) - \phi_y(t)]} \right| \quad (22)$$

Here, $\phi_x(t)$ is the Hilbert phase of S_x and calculated as the ratio of Hilbert transform of the signal S_x to itself: $\phi_x(t) = \arctan \frac{\widetilde{S_x(t)}}{S_x(t)}$.

Recently, increasing number of researches have proven the effectiveness of EEG connectivity features for emotion recognition. Chen et al. [110] extracted PCC, MI, and phase coherence connectivity for emotion recognition using the DEAP database. A binary SVM classification was performed and the results showed that the best accuracies of 76.2% for valence and 73.6% for arousal were obtained using mutual information with all frequency bands. Moon et al. [111] fed brain connectivity features into convolutional neural networks for emotion recognition. Based on PCC, PLV, and transfer entropy extracted from DEAP database, an outstanding recognition accuracy of 80.7% for valence was obtained when convolution kernel sizes was 5 and connectivity matrix of PLV was used as the input data. In order to

explore the dynamic emotion-related neural mechanism, Liu et al. [112] proposed a dynamic functional connectivity method by separately sorting the static networks constructed from phase lag index and then feeding them into temporal brain network. They validated the effectiveness of this dynamic functional connectivity analysis on SEED database and achieved recognition accuracy of 87.0% in Beta band.

4.2 EEG feature selection

High dimension EEG features may contain large numbers of unnecessary features. In order to avoid the curse of dimensionality and improve classification performance, it is necessary to select out emotion-related EEG features prior to emotion classifier modeling.

According to whether label information is used, feature selection can be simply classified into supervised and unsupervised based feature selection.

4.2.1 Supervised feature selection

(1) Linear discriminant analysis (LDA) is a supervised linear dimensionality reduction and feature selection method. The main idea of LDA is to find a suitable projection direction based on the class discriminatory information. Projection direction is determined when minimal intra-class variance and maximal inter-class variance are simultaneously achieved [20]. Thus, projection points from the same class should be as close as possible and the distance of points from different classes should be as far as possible.

The ratio of the intra-class variance to the inter-class variance is recorded as a Fisher score. The higher the Fisher score, the higher the discrimination between two groups of features. LDA selects out the feature subset with the largest Fisher score to achieve the purpose of feature selecting.

(2) Maximum relevance minimum redundancy (mRMR) adopts MI to select out emotion-related EEG features that satisfy both maximum correlation and minimum redundancy [91].

Selecting the feature subset in high relevance with labels can reduce information redundancy. First, the correlation between EEG features F and emotional label c is calculated with the help of MI.

$$I(F, c) = \iint p(F, c) \log \frac{p(F, c)}{p(F)p(c)} dFdc \quad (23)$$

The maximum correlation is obtained first by calculating the average mutual information between features F_i and class c in each feature subset S [82]. The feature subset S has high correlation with class c . It is calculated as:

$$\max[D(S, c)], \text{ where } D = \frac{1}{|S|} \sum_{F_i \in S} I(F_i, c) \quad (24)$$

The features selected by the maximum correlation may have redundant information. Therefore, the dependence between features is large. When the correlation and dependence of feature F_i and F_j is large, removing one of them will not affect the recognition performance [113].

Therefore, the features with low dependency can be further filtered using the minimum redundancy calculation, which can effectively reduce the redundancy and dimensionality of features. Thus,

$$\min[R(S)], \text{ where } R = \frac{1}{|S|^2} \sum_{F_i, F_j \in S} I(F_i, F_j) \quad (25)$$

The purpose of mRMR feature selection is to find the feature subset that simultaneously satisfies maximum relevance and minimum redundancy. Hence,

$$\max[\Phi(D, R)], \text{ where } \Phi = D - R \quad (26)$$

or

$$\max[\Phi(D, R)], \text{ where } \Phi = \frac{D}{R} \quad (27)$$

Here, feature subset S_{k-1} that satisfy the objective function mentioned above is filtered first. Then, forward search method [113] is used for next feature searching from the remaining feature set $\{F - S_{k-1}\}$. Feature subset S_k is finally selected as follows:

$$\max_{F_j \in \{F - S_{k-1}\}} \left[I(F_i, c) - \frac{1}{k-1} \sum_{F_i \in S_{k-1}} I(F_j, F_i) \right] \quad (28)$$

4.2.2 Unsupervised feature selection

(1) Principal components analysis (PCA) is one of the commonly used unsupervised feature selection methods. Orthogonal transformation is used for linear transformation. By projecting samples into a low-dimension space, a series of linearly independent principal components are obtained. PCA preserves as much data information as possible by minimizing the reconstruction error during feature selection.

(2) Kernel principal component analysis (KPCA) is a nonlinear extension of PCA. Nonlinear mapping function Γ is used to map the data to a higher dimensional Hilbert space to make them linearly separable. Then, PCA is applied for further dimension reduction. The process of KPCA calculation is as follows:

a) Feature dataset \mathbf{X} is mapped to the high-dimensional space $\Gamma(\mathbf{X}) = [\Gamma(\mathbf{X}_1), \dots, \Gamma(\mathbf{X}_n)]$ using the nonlinear mapping function Γ .

b) The kernel matrix \mathbf{K} is calculated as follows: $\mathbf{K} = \Gamma(\mathbf{X})^T \Gamma(\mathbf{X})$.

c) The transformation matrix \mathbf{A} that satisfies $\max_{A^T A = \mathbf{E}} \text{tr}(\mathbf{A}^T \mathbf{K} \mathbf{H} \mathbf{K}^T \mathbf{A})$ is determined, where \mathbf{E} is the identity matrix; the centering matrix is calculated as $\mathbf{H} = \mathbf{E} - \frac{1}{n} \mathbf{I}$, where \mathbf{I} is a $n \times n$ matrix of ones.

d) Finally, feature subset is selected

as $\mathbf{Z} = \mathbf{A}^T \mathbf{K}$.

5 EEG-based emotion classifiers

Modeling is one of the significant processes in affective computing, where machine learning algorithms can effectively learn the underlying relationship between different emotional states and EEG features. The commonly used supervised learning, unsupervised learning, and semi-supervised learning classifiers on video-triggered EEG-based databases for emotion recognition are reported in Table 4.

5.1 Supervised learning classifiers

Supervised learning classifiers establish the feature-label mapping relationship and modify the parameters of classification model under the guidance of label information. Generally, higher recognition accuracy can be obtained by using supervised learning-based classification models. Emotion classification with supervised learning algorithms such as naive Bayes, Probabilistic neural networks (PNNs), SVM, and KNN have been widely verified on the publicly available video-triggered EEG-emotion databases.

5.1.1 Naive Bayes

Naive Bayes is a probabilistic classification model for supervised learning. This algorithm learns the joint probability distribution between features of the training set and labels by assuming that the feature data are independent from each other. Test data are fed into the probability distribution model and then predicted labels are obtained according to posterior probability. The advantages of naive Bayes include its simple algorithm, high computational efficiency, high accuracy, and less sensitivity to missing data [120]. Considering the outstanding performance of naive Bayes on

Table 4 Emotion classifiers and recognition accuracies in databases for video-triggered EEG-based emotion recognition.

Ref.	Database (stimuli)	EEG Features	Classifier (Supervised/Unsupervised)	Emotion Types	Accuracy
Müller et al. [71]	DEAP (music videos)	PSD and DASM	Naive Bayes (supervised)	LALV, HALV, LAHV, and HAHV	Valence: 57.6% Arousal: 62.0% Liking: 55.4%
Nakisa et al. [114]	DEAP (music videos) MAHNOB-HCI (movie clips) Nakisa et al. database (movie clips)	Time domain: power, mean, Hjorth features, FD, etc. Frequency domain: PSD, mean Time-frequency domain: power and root mean square	PNN (supervised)	HA-P, LA-P, HA-N, and LA-N	DEAP: 67.5% MAHNOB-HCI: 97.0% Nakisa et al. database: 65%
Soleymani et al. [115]	— (20 movie clips)	PSD and DASM	SVM with RBF kernel (supervised)	Positive, neutral, and negative	Valence: 50.5% Arousal: 62.1%
Katsigiannis et al. [22]	DREAMER (movie clips)	PSD	SVM with RBF kernel (supervised)	Valence, arousal, and dominance	Valence: 62.5% Arousal: 62.2% Dominance: 61.8%
Zheng et al. [21]	SEED-IV (movie clips)	PSD and DE	SVM with linear kernel (supervised)	Happy, sad, fear, and neutral	Happy: 80.0% Sad: 63.0% Fear: 65.0% Neutral: 78.0%
Mohammadi et al. [105]	DEAP (music videos)	Energy and entropy of extracted frequency bands	KNN ($k = 3$) (supervised)	LALV, HALV, LAHV, and HAHV	Valence: 86.8% Arousal: 84.1%
Hwang et al. [116]	SEED (Chinese movie clips)	DE	CNN (unsupervised)	Positive, neutral, and negative	Positive: 96.0% Neutral: 92.0% Negative: 83.0%
Jang et al. [117]	DEAP (music video)	Energy and entropy of extracted frequency bands	GCNN (unsupervised)	—	65.27%
Song et al. [118]	SEED (Chinese movie clips) DREAMER (movie clip)	DE, PSD, DASM, and RASM	DGCNN (unsupervised)	SEED positive, neutral, and negative DREAMER valence, arousal, and dominance	SEED subject-dependent: 90.4% subject-independent: 80.0% DREAMER Valence: 86.2% Arousal: 84.5% Dominance: 85.0%
Liang et al. [86]	DEAP (music video)	Time domain: power, mean, standard deviation, Hjorth features, FD, etc. Frequency domain: PSD Time-Frequency domain: energy and Shannon entropy extracted from the detail coefficients of level 4 to 7	Hypergraph Theory (unsupervised)	Valence, arousal, dominance, and liking	Valence: 56.3% Arousal: 62.3% Dominance: 64.2% Liking: 66.1%

(Continued)

Ref.	Database (stimuli)	EEG Features	Classifier (Supervised/Unsupervised)	Emotion Types	Accuracy
Zheng et al. [23]	SEED (Chinese movie clips)	DE	DBN (semi-supervised)	Positive, neutral, and negative	86.08%
Xu et al. [119]	DEAP (music video)	PSD	Stacked denoising auto-encoder (semi-supervised) DBN (semi-supervised)	LALV, HALV, LAHV, and HAHV	Stacked denoising auto-encoder Valence: 82.4% Arousal: 81.5% Liking: 81.1% DBN Valence: 87.1% Arousal: 86.7% Liking: 86.7%

PSD, power spectral density; DASM, differential asymmetry; FD, fractal dimension; PNN, probabilistic neural networks; HA-P, high arousal-positive emotions; LA-P, low arousal-positive emotions; HA-N, high arousal-negative emotions; LA-N, low arousal-negative emotions; SVM, support vector machines; RBF, radial basis function; DE, differential entropy; KNN, k-nearest neighbor; CNN, convolutional neural networks; GCNN, graph convolutional neural networks; DGCNN, dynamical graph convolutional neural networks; DBN, deep belief networks.

imbalanced data size within categories in small datasets, Koelstra et al. [100] used naive Bayes classifier to train emotion recognition model using DEAP database they established. The recognition accuracies of valence, arousal, and liking were respectively 57.6%, 62.0%, and 55.4%.

5.1.2 Probabilistic neural network

Based on the Bayesian decision-making rules of Bayesian network, Specht proposed a supervised learning algorithm known as the PNN [121]. PNN is a feedforward network consisting of input layer, pattern layer, summation layer, and output layer with great characteristics of better generalization and fault tolerance for outliers [114]. Using radial basis function kernel as the activation function of pattern layer, PNN is closely equivalent to Bayesian optimal classification. Thus, PNN are often applied to emotion decoding on EEG signals from different modalities [114]. In addition, PNN, simple in structure and fast in training, is very suitable for real-time emotion recognition [122]. Based on EEG signals from MAHNOB-HCI, DEAP, and

their own emotion database, Nakisa et al. [114] used PNN to verify the effectiveness of evolutionary computation algorithm for EEG feature selection. Zhang et al. [122] compared the performance of PNN and SVM in emotion recognition on the DEAP dataset. Their research results showed that PNN reached a comparable classification accuracy of SVM, but has much simpler network structure and faster training process.

5.1.3 Support vector machine

SVM classifier is less sensitive to outliers and can also achieve considerable performance on small-sample training sets. With better robustness and generalization, SVM is one of the most efficient and alternative classifiers for emotion recognition [41, 123]. Katsigiannis and Ramzan [22] adopted SVM with radial basis function (RBF) kernel for emotion recognition on data from the DREAMER database, which were collected by wireless and portable EEG recording devices. The classification results of DEAP (with accuracy of 57.6% for valence and

accuracy of 62.0% for arousal), MAHNOB-HCI (with accuracy of 57.0% for valence and accuracy of 52.4% for arousal), and DECAF (a MEG-based multimodal dataset for decoding user physiological responses to affective multimedia content, with accuracy of 59.0% for valence and accuracy of 62.0% for arousal) databases collected by non-portable devices were applied for comparison. In conclusion, their results revealed that there was no significant classification difference of EEG signals collected by wireless and portable devices in emotion recognition (with accuracy of 62.5% for valence and accuracy of 62.2% for arousal). To verify the effectiveness of portable EEG recording devices in emotion recognition experiment, this research explored the application prospects of wireless portable devices for emotion recognition and provided supportive guidance for the construction of portable emotion-based human-computer interface systems.

5.1.4 *K*-nearest neighbor

K-nearest neighbor is a non-parametric supervised learning classifier, and often used as a baseline method for evaluating the performance of other classifiers. Mohammadi et al. [105] utilized SVM and KNN to build the emotion recognition model using the EEG signals from DEAP database. When $k = 3$, KNN classifier reached the highest recognition accuracy of 86.75% for valence, and 84.05% for arousal. Similarly, Li et al. [124] applied KNN ($k = 3$) for valence and arousal classification on DEAP database. Their research result showed that promising recognition result was achieved when using EEG features extracted from Gamma band (with recognition accuracy of 95.70% for valence and 95.69% for arousal). Thus, they inferred that Gamma band may contain more emotion-related information comparing to other frequency

bands of EEG signals.

5.2 Unsupervised learning classifiers

Contrary to supervised learning, no label information is available in unsupervised learning. By adopting clustering or feature analysis methods for unlabeled data, unsupervised learning classifiers detect and learn the underlying connections and regulation among EEG features, and then infer possible category information from these unlabeled EEG features. Thus, it can better adapt to individual differences in cross-subjects classifier building [125]. Unsupervised learning model is specifically designed for unlabeled data classification, which is of great significance in solving practical problems, such as classification of EEG signals with missing labels, difficulties in obtaining label set, and time-consuming manual labeling.

5.2.1 *Convolutional neural networks*

Convolutional neural networks (CNN) is a feedforward network with convolutional layers and deep structure. Excellent in fault tolerance, self-adaptability, and generalization, CNN can detect the underlying mapping relation between category and raw data and extract class-related deep features from large numbers of unlabeled raw data. It has been widely used in image recognition and achieved remarkable results. Researchers seldom directly feed multi-dimensional EEG data into CNN model for emotion classification due to the limitation that EEG signals are spatially discrete and temporally non-stationary. Hwang et al. [112] innovatively proposed a topology-preserving DE feature-based CNN classifier for emotion recognition. The new proposed method was evaluated on SEED database. Firstly, topology-reserving DE images were generated while keeping the spatial information among

electrodes. Specifically, Hwang et al. first calculated DE features of each electrode in the Delta, Theta, Alpha, Beta, and Gamma bands from SEED database. Then, Azimuthal Equidistant Projection (AEP) [126] was performed to generate an EEG topology map that represented DE features. Finally, the generated topology map was fed into CNN for emotion recognition to learn the spatial information within multiple electrodes. The EEG topology map generated by AEP can well preserve the spatial information of EEG signals, thus, solving the problem of low spatial resolution and achieving a promising classification on the SEED database. The accuracies for positive, neutral, and negative emotions were 96%, 92%, and 83% respectively.

5.2.2 Graph convolutional neural networks

Graph convolutional neural networks (GCNN) is a deep learning network that combines CNN and spectrogram theory to find the relationship within different nodes from graph signals. GCNN is an effective method to extract features from discrete spatial signals [127, 128], which can be used to explore the spatial connection of multi-dimensional EEG signals for emotion recognition [118]. Jang et al. [117] utilized GCNN for emotion recognition based on DEAP database. First of all, intra-band graphs for each electrode were created by calculating the power and entropy of Delta (0–3 Hz), Theta (4–7 Hz), low Alpha (8–10 Hz), high Alpha (10–12 Hz), low Beta (13–16 Hz), middle Beta (17–20 Hz), and high Beta (21–29 Hz) while considering the relationship between electrodes. Then, they merged all intra-band graphs into a larger graph and fed this larger graph into the GCNN classifier for emotion recognition. Compared with traditional algorithms of KNN (with highest recognition accuracy of 48.5%) and random forest (with highest recognition

accuracy of 51.3%), GCNN obtained the higher recognition accuracy of 65.27%.

5.2.3 Dynamical graph convolutional neural networks

To further promote the development of GCNN, Song et al. proposed dynamical graph convolutional neural networks (DGCNN) for multi-dimensional EEG-based emotion recognition [118]. On the basis of GCNN, DGCNN dynamically calculate and update adjacency matrix of graph nodes according to the changes of graph signals. Song et al. [118] pointed out that dynamically learning the inner relationship of multi-electrode EEG signals would efficiently improve the emotion recognition accuracy of EEG signals. Their research results showed that DGCNN achieved better classification performance on DREAMER and SEED databases. Classification accuracies of valence, arousal, and dominance were respectively 86.23%, 84.54%, and 85.02% on DREAMER database. Based on the DE feature of SEED database, classification accuracy of 90.4% was achieved in subject-dependent classification model and the accuracy of 79.95% for subject-independent validation. Similarly, based on the EEG signals from SEED database, Wang et al. [129] compared DGCNN with SVM (86.08% \pm 8.34%), DBN (83.99% \pm 9.72%), GCNN (87.40% \pm 9.20%). Unsurprisingly, DGCNN achieved better emotion recognition performance in multi-channels EEG signals with classification accuracy of 90.40% \pm 8.49%.

5.3 Semi-supervised learning classifiers

Semi-supervised learning is a classifier that combines the benefits of both unsupervised learning and supervised learning. By simultaneously feeding in a small amount of labeled data and large numbers of unlabeled data for model training, semi-supervised learning classifier can improve the classification accuracy of unlabeled

data, and deal with the problem of time-consuming manual data labeling. At the same time, semi-supervised learning algorithm makes up for the poor generalization of supervised learning classifiers and the low classification accuracy of unsupervised learning classifiers with significant practical applications.

Deep belief networks (DBN) is a probabilistic generation model with deep structure that can be used to extract high related features from EEG, and is one of typical and state-of-the-art semi-supervised learning classifiers. DBN mainly follows unsupervised pre-training, unsupervised fine-tuning, and supervised fine-tuning [38, 130]. Firstly, a deep model with a stacking structure is generated by performing unsupervised greedy layer-wise pre-training. Secondly, in the processing of unsupervised fine-tuning, parameters of n restricted Boltzmann machines (RBM) that make up DBN are updated with backpropagation. The main purpose of unsupervised fine-tuning is to make the reconstructed visible unit as close as possible to the input visible unit by adjusting the connection weight and bias between each layer. Finally, label information is added to the highest layer and the weights will be updated through error backpropagation to realize supervised fine-tuning of parameters.

DBN can efficiently combine feature extraction with feature learning to obtain better classification performance and has been widely applied to the research of EEG-based emotion recognition [120, 131]. Based on SEED database, Zheng et al. [23] established positive, neutral, and negative emotion recognition models using DBN to explore the critical frequency bands and crucial channels for improving the performance of emotion recognition. Their research discovered that Beta and Gamma band might contain more information that better characterized electrophysiological changes of the brain

under different emotional states. In comparison with the classification results of SVM (83.99%), logistic regression (82.70%), and KNN (72.60%), DBN with accuracy of 86.08% could effectively recognize positive, neutral, and negative emotion from multi-channels EEG signals.

6 Summary and outlook

Based on the current video-triggered EEG-based emotion recognition databases, this paper mainly summarized EEG-based emotion recognition methods in recent years. According to the typical processing pipeline, different emotion evoking methods were compared and the details about publicly available databases for video-triggered EEG-based emotion recognition were introduced. Moreover, a systematic introduction of EEG features extraction, selection, and modeling methods used in the existing video-triggered EEG-emotion public databases was presented.

With the beneficial contributions of these public video-triggered EEG-emotion databases, great progresses on emotion recognition using EEG signals have been made in recent years. However, there are still several limitations about the public databases that requires further improvement.

(1) Low quantity and small size. As shown in Table 2, the existing public EEG-based emotion recognition databases merely include DEAP, MAHNOB-HCI, SEED series databases (i.e., SEED, SEED-IV, and SEED-VIG). Meanwhile, the number of participants in each database is relatively small, which could lead to both poor classification performance and weak generalization of emotion recognition modeling. The individual difference could not be properly examined. Also, the application of deep learning networks in robust emotion classification modeling will be limited.

(2) Unbalanced classes distribution. In DEAP database, Koelstra et al. [20] found that, without consideration of unbalance problem in class distribution, the classification accuracy would be much higher (For arousal, accuracy: 62.0%, F1: 58.3%. For valence, accuracy: 57.6%, F1: 56.3%. For liking, accuracy: 55.4%, F1: 50.2%). The sample imbalance in the class distribution will directly affect the credibility of the classification results, and the robustness and generalization ability of the trained classification model will be decreased.

(3) No standard self-assessment methods. The collected self-assessment results in the existing databases are not consistent. Currently, two self-assessment labeling methods are mainly adopted in the existing public EEG-emotion databases. One method is based on quantitative indicators, such as valence, arousal, and dominance for labeling (e.g., DEAP and MAHNOB-HCI database); another self-assessment method is based on discrete emotion labels, such as happiness, sadness, fear, disgust, and anger (e.g., SEED and SEED-IV database). Moreover, most of the existing databases (such as DEAP and MAHNOB-HCI) used the SAM scale; the SEED database used the Philippot questionnaire as self-assessment scale while the SEED-IV database used the PANAS scale. Due to these differences, it is a great challenge to propose a cross-database classification model, and it is hard to conduct transfer learning by transferring the data collected under different labeling conditions of self-assessment.

(4) Lack of long-term emotion triggering. Existing emotion databases mostly used short-term videos (i.e., video clips with duration shorter than 6 min) as the triggering materials. There is a lack of emotion study triggered by long-term videos (more than 10 min). However, to successfully evoke negative emotions, such as sadness, it requires a longer continuous

triggering. Moreover, conducting a long-term emotion induction experiment helps to better understand the dynamic process of emotion responses and the complex regulation mechanism in brain; it further explores the neuro-physiological mechanisms related to emotions.

In addition, emotion classification from multi-modal physiological signals is of guiding significance for proper decoding of physiological responses to emotions. Most of the existing databases used the individual self-assessment as ground truth. However, ratings from self-assessment have obvious individual differences. Thus, the reliability cannot be guaranteed. Multi-modal physiological signals (such as ECG, and eye gaze data) can be used as an index to provide more evidences about the evoked emotions. The development of multi-modality physiological signals-based emotion recognition model can improve the classification performance with promising results.

Promoting the development of emotion recognition has great application significance. EEG-based emotion recognition with videos has attracted wide attentions in many fields including psychology, cognitive science, medicine, and information technology. However, the current video-triggered EEG-based emotion recognition only tackles the recognition problem with specified emotion states. However, in real life, human emotion is complex, diverse, and dynamic. To realize the real-time emotion recognition with an online emotion monitor system, the dynamic changes of emotions from the perspective of neurophysiology should be further studied, and the underlying neuro-physiological mechanism of emotion regulation should be further explored.

Also, the establishment of robust emotion recognition models is indispensable in affective computing. However, the existing video-

triggered EEG-based emotion recognition researches are mainly based on supervised learning classifiers. The lack of labeled data and low reliability of self-assessed labeling seriously affect the classification performance. It is worth noting that unsupervised learning and semi-supervised learning have better self-adaptation and self-learning capabilities in dealing with such problems as individual differences, inaccurate label information, and lack of labels.

To sum up, improving and unifying the construction of video-triggered EEG-based emotion databases could further promote the development of emotion recognition, which benefits the development of emotion classification with a more robust and generalized EEG model. Further, the quality of life and mental health could be greatly improved with the application of real-time emotion detection and regulation.

Conflict of interests

All contributing authors have no conflict of interests.

Financial supports

This work is funded by the National Natural Science Foundation of China (Grant No. 61906122).

References

- [1] Dolan RJ. Emotion, cognition, and behavior. *Science*. 2002, **298**(5596): 1191–1194.
- [2] Bucks RS, Radford SA. Emotion processing in Alzheimer's disease. *Aging Ment Health*. 2004, **8**(3): 222–232.
- [3] Sirois BC, Burg MM. Negative emotion and coronary heart disease. *Behav Modif*. 2003, **27**(1): 83–102.
- [4] Joormann J, Gotlib IH. Emotion regulation in depression: relation to cognitive inhibition. *Cogn Emot*. 2010, **24**(2): 281–298.
- [5] Plutchik R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am Sci*. 2001, **89**(4): 344–350.
- [6] Suttles J, Ide N. Distant supervision for emotion classification with discrete binary values. In *Computational Linguistics and Intelligent Text Processing*. Gelbukh A, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.
- [7] Shu L, Xie JY, Yang MY, et al. A review of emotion recognition using physiological signals. *Sensors*. 2018, **18**(7): E2074.
- [8] Ekman P, Friesen WV, O'Sullivan M, et al. Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol*. 1987, **53**(4): 712–717.
- [9] Verma GK, Tiwary US. Affect representation and recognition in 3D continuous valence–arousal–dominance space. *Multimed Tools Appl*. 2017, **76**(2): 2159–2183.
- [10] Russell JA. A circumplex model of affect. *J Pers Soc Psychol*. 1980, **39**(6): 1161–1178.
- [11] Liu Y, Sourina O. EEG-based dominance level recognition for emotion-enabled interaction. In *2012 IEEE International Conference on Multimedia and Expo*, Melbourne, VIC, Australia, 2012, pp 1039–1044.
- [12] Alarcão SM, Fonseca MJ. Emotions recognition using EEG signals: a survey. *IEEE Trans Affect Comput*. 2019, **10**(3): 374–393.
- [13] Etkin A, Büchel C, Gross JJ. The neural bases of emotion regulation. *Nat Rev Neurosci*. 2015, **16**(11): 693–700.
- [14] Chanel G, Kierkels JJM, Soleymani M, et al. Short-term emotion assessment in a recall paradigm. *Int J Hum-Comput Stud*. 2009, **67**(8): 607–627.
- [15] Zhuang N, Zeng Y, Yang K, et al. Investigating patterns for self-induced emotion recognition from EEG signals. *Sensors*. 2018, **18**(3): 841.
- [16] Mehmood RM, Lee HJ. A novel feature extraction method based on late positive potential for emotion recognition in human brain signal patterns. *Comput Electr Eng*. 2016, **53**: 444–457.
- [17] Sohaib AT, Qureshi S, Hagelbäck J, et al. Evaluating

- classifiers for emotion recognition using EEG. In *Foundations of Augmented Cognition*. Schmorrow DD, Fidopiastis CM, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp 492–501.
- [18] Lin YP, Wang CH, Jung TP, et al. EEG-based emotion recognition in music listening. *IEEE Trans Biomed Eng.* 2010, **57**(7): 1798–1806.
- [19] Daly I, Williams D, Hallowell J, et al. Music-induced emotions can be predicted from a combination of brain activity and acoustic features. *Brain Cogn.* 2015, **101**: 1–11.
- [20] Koelstra S, Muhl C, Soleymani M, et al. DEAP: a database for emotion Analysis: Using physiological signals. *IEEE Trans Affective Comput.* 2012, **3**(1): 18–31.
- [21] Zheng WL, Liu W, Lu YF, et al. EmotionMeter: a multimodal framework for recognizing human emotions. *IEEE Trans Cybern.* 2019, **49**(3): 1110–1122.
- [22] Katsigiannis S, Ramzan N. DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J Biomed Health Inform.* 2018, **22**(1): 98–107.
- [23] Zheng WL, Lu BL. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans Auton Mental Dev.* 2015, **7**(3): 162–175.
- [24] Kothe CA, Makeig S, Onton JA. Emotion recognition from EEG during self-paced emotional imagery. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland, 2013, pp 855–858.
- [25] Salas CE, Radovic D, Turnbull OH. Inside-out: Comparing internally generated and externally generated basic emotions. *Emotion.* 2012, **12**(3): 568–578.
- [26] Poria S, Cambria E, Bajpai R, et al. A review of affective computing: From unimodal analysis to multimodal fusion. *Inf Fusion.* 2017, **37**: 98–125.
- [27] Ellard KK, Farchione TJ, Barlow DH. Relative effectiveness of emotion induction procedures and the role of personal relevance in a clinical sample: a comparison of film, images, and music. *J Psychopathol Behav Assess.* 2012, **34**(2): 232–243.
- [28] Soleymani M, Asghari-Esfeden S, Fu Y, et al. Analysis of EEG signals and facial expressions for continuous emotion detection. In *IEEE Trans Affective Comput.* 2015, **7**(1): 17–28.
- [29] Castellano G, Kessous L, Caridakis G. Emotion recognition through multiple modalities: face, body gesture, speech. In *Affect and Emotion in Human-Computer Interaction*. Peter C, Beale R, Eds. Berlin, Heidelberg: Springer, 2008.
- [30] Kessous L, Castellano G, Caridakis G. Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis. *J Multimodal User Interfaces.* 2010, **3**(1/2): 33–48.
- [31] Gunes H, Schuller B, Pantic M, et al. Emotion representation, analysis and synthesis in continuous space: a survey. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, Santa Barbara, CA, USA, 2011, pp 827–834.
- [32] Bethel CL, Salomon K, Murphy RR, et al. Survey of Psychophysiology Measurements Applied to Human-Robot Interaction. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, Jeju, South Korea, 2007, pp 732–737.
- [33] Abadi MK, Subramanian R, Kia SM, et al. DECAF: MEG-based multimodal database for decoding affective physiological responses. *IEEE Trans Affective Comput.* 2015, **6**(3): 209–222.
- [34] Khosrowabadi R, Quek HC, Wahab A, et al. EEG-based Emotion Recognition Using Self-Organizing Map for Boundary Detection. In *2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey, 2010, pp 4242–4245.
- [35] Lisetti CL, Nasoz F. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP J Adv Signal Process.* 2004, **2004**(11): 929414.
- [36] Healey JA, Picard RW. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans Intell Transport Syst.* 2005, **6**(2): 156–166.
- [37] Kreibig SD. Autonomic nervous system activity in emotion: a review. *Biol Psychol.* 2010, **84**(3): 394–421.

- [38] Agrafioti F, Hatzinakos D, Anderson AK. ECG pattern analysis for emotion detection. *IEEE Trans Affective Comput.* 2012, **3**(1): 102–115.
- [39] Das P, Khasnobish A, Tibarewala DN. Emotion recognition employing ECG and GSR signals as markers of ANS. In *2016 Conference on Advances in Signal Processing (CASP)*, Pune, India, 2016, pp 37–42.
- [40] Zhang Q, Chen XX, Zhan QY, et al. Respiration-based emotion recognition with deep learning. *Comput Ind.* 2017, **92**: 84–90.
- [41] Nie D, Wang XW, Duan RN, et al. A survey on EEG based emotion recognition. *Chinese Journal of Biomedical Engineering.* 2012, **31**(4): 595–606.
- [42] Fairclough SH. Fundamentals of physiological computing. *Interact Comput.* 2009, **21**(1/2): 133–145.
- [43] Waugh CE, Shing EZ, Avery BM. Temporal dynamics of emotional processing in the brain. *Emot Rev.* 2015, **7**(4): 323–329.
- [44] Thiruchselvam R, Blechert J, Sheppes G, et al. The temporal dynamics of emotion regulation: an EEG study of distraction and reappraisal. *Biol Psychol.* 2011, **87**(1): 84–92.
- [45] Mauss IB, Robinson MD. Measures of emotion: a review. *Cogn Emot.* 2009, **23**(2): 209–237.
- [46] Davidson PR, Jones RD, Peiris MT. EEG-based lapse detection with high temporal resolution. *IEEE Trans Biomed Eng.* 2007, **54**(5): 832–839.
- [47] Niemic CP. Studies of emotion: A theoretical and empirical review of psychophysiological studies of emotion. 2002, **1**(1): 15–18.
- [48] Alotaiby T, El-Samie FEA, Alshebeili SA, et al. A review of channel selection algorithms for EEG signal processing. *EURASIP J Adv Signal Process.* 2015, **2015**: 66.
- [49] Davidson RJ, Abercrombie H, Nitschke JB, et al. Regional brain function, emotion and disorders of emotion. *Curr Opin Neurobiol.* 1999, **9**(2): 228–234.
- [50] Raichle ME, MacLeod AM, Snyder AZ, et al. A default mode of brain function. *PNAS.* 2001, **98**(2): 676–682.
- [51] Güntekin B, Basar E. Emotional face expressions are differentiated with brain oscillations. *Int J Psychophysiol.* 2007, **64**(1): 91–100.
- [52] Wang XW, Nie D, Lu BL. Emotional state classification from EEG data using machine learning approach. *Neurocomputing.* 2014, **129**: 94–106.
- [53] Schmidt LA, Trainor LJ. Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cogn Emot.* 2001, **15**(4): 487–500.
- [54] Schutter DJLG, Putman P, Hermans E, et al. Parietal electroencephalogram beta asymmetry and selective attention to angry facial expressions in healthy human subjects. *Neurosci Lett.* 2001, **314**(1/2): 13–16.
- [55] Lang PJ, Bradley MM. Emotion and the motivational brain. *Biol Psychol.* 2010, **84**(3): 437–450.
- [56] Kim MK, Kim M, Oh E, et al. A review on the computational methods for emotional state estimation from the human EEG. *Comput Math Methods Med.* 2013, **2013**: 1–13.
- [57] Mühl C, Allison B, Nijholt A, et al. A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. *Brain - Comput Interfaces.* 2014, **1**(2): 66–84.
- [58] Ray W, Cole H. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science.* 1985, **228**(4700): 750–752.
- [59] Gordon R, Ciorciari J, van Laer T. Using EEG to examine the role of attention, working memory, emotion, and imagination in narrative transportation. *Eur J Marketing.* 2017, **52**(1/2): 92–117.
- [60] Pedley TA. Electroencephalography: Basic principles, clinical applications, and related fields. Niedermeyer E, da Silva FL, Eds. Baltimore: Williams & Wilkins, 2005.
- [61] Bonanni E, di Coscio E, Maestri M, et al. Differences in EEG delta frequency characteristics and patterns in slow-wave sleep between dementia patients and controls. *J Clin Neurophysiol.* 2012, **29**(1): 50–54.
- [62] Palop JJ, Mucke L. Epilepsy and cognitive impairments in Alzheimer disease. *Arch Neurol.* 2009, **66**(4): 435–440.
- [63] Sammler D, Grigutsch M, Fritz T, et al. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology.* 2007, **44**(2): 293–304.
- [64] Finelli LA, Baumann H, Borbély AA, et al. Dual electroencephalogram markers of human sleep homeostasis: correlation between theta activity in

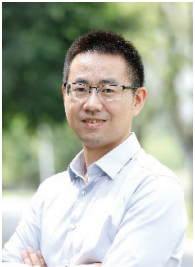
- waking and slow-wave activity in sleep. *Neuroscience*. 2000, **101**(3): 523–529.
- [65] Kawasaki M, Kitajo K, Yamaguchi Y. Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. *Eur J Neurosci*. 2010, **31**(9): 1683–1689.
- [66] Bazanova OM, Vernon D. Interpreting EEG alpha activity. *Neurosci Biobehav Rev*. 2014, **44**: 94–110.
- [67] Choi D, Sekiya T, Minote N, et al. Relative left frontal activity in reappraisal and suppression of negative emotion: Evidence from frontal alpha asymmetry (FAA). *Int J Psychophysiol*. 2016, **109**: 37–44.
- [68] Nelson BD, Kessel EM, Klein DN, et al. Depression symptom dimensions and asymmetrical frontal cortical activity while anticipating reward. *Psychophysiology*. 2018, **55**(1): DOI 10.1111/psyp.12892.
- [69] Yuvaraj R, Murugappan M, Ibrahim NM, et al. Emotion classification in Parkinson's disease by higher-order spectra and power spectrum features using EEG signals: a comparative study. *J Integr Neurosci*. 2014, **13**(1): 89–120.
- [70] Gola M, Magnuski M, Szumska I, et al. EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *Int J Psychophysiol*. 2013, **89**(3): 334–341.
- [71] Müller MM, Keil A, Gruber T, et al. Processing of affective pictures modulates right-hemispheric gamma band EEG activity. *Clin Neurophysiol*. 1999, **110**(11): 1913–1920.
- [72] Huang D, Guan CT, Ang KK, et al. Asymmetric Spatial Pattern for EEG-based emotion detection. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, Brisbane, Australia, 2012, pp 1–7.
- [73] Park K, Choi H, Lee KJ, et al. Emotion recognition based on the asymmetric left and right activation. *Int J Med Med Sci*. 2011, **3**(6): 201–209.
- [74] Lang PJ, Bradley MM, Cuthbert BN. International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL, USA, 2008.
- [75] Marchewka A, Zurawski Ł, Jednoróg K, et al. The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav Res Methods*. 2014, **46**(2): 596–610.
- [76] Bradley MM, Lang PJ. The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. Technical report B-3. University of Florida, Gainesville, FL, USA, 2008.
- [77] Soleymani M, Lichtenauer J, Pun T, et al. A multimodal database for affect recognition and implicit tagging. *IEEE Trans Affective Comput*. 2012, **3**(1): 42–55.
- [78] Becker H, Fleureau J, Guillotel P, et al. Emotion recognition based on high-resolution EEG recordings and reconstructed brain sources. *IEEE Trans Affective Comput*. 2020, **11**(2): 244–257.
- [79] Song TF, Zheng WM, Lu C, et al. MPED: a multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access*. 2019, **7**: 12177–12191.
- [80] Li Y, Zheng WM, Cui Z, et al. EEG emotion recognition based on graph regularized sparse linear regression. *Neural Process Lett*. 2019, **49**(2): 555–571.
- [81] Liu YJ, Yu MJ, Zhao GZ, et al. Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans Affective Comput*. 2018, **9**(4): 550–562.
- [82] Zheng WL, Zhu JY, Lu BL. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Trans Affective Comput*. 2019, **10**(3): 417–429.
- [83] Zhu JY, Zheng WL, Lu BL. Cross-subject and cross-gender emotion classification from EEG. In *World Congress on Medical Physics and Biomedical Engineering*, Toronto, Canada, 2015, pp 1188–1191.
- [84] Zheng WL, Lu BL. A multimodal approach to estimating vigilance using EEG and forehead EOG. *J Neural Eng*. 2017, **14**(2): 026017.
- [85] Tripathi S, Acharya S, Sharma RD, et al. Using deep and convolutional neural networks for accurate emotion classification on deap dataset. In *Twenty-Ninth IAAI Conference*, San Francisco, California, USA, 2017, pp 4746–4752.
- [86] Liang Z, Oba S, Ishii S. An unsupervised EEG decoding system for human emotion recognition. *Neural Netw*. 2019, **116**: 257–268.
- [87] Chai X, Wang QS, Zhao YP, et al. Unsupervised

- domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition. *Comput Biol Med.* 2016, **79**: 205–214.
- [88] Teplan M. Fundamentals of EEG measurement. *Meas Sci Rev.* 2002, **2**(2): 1–11.
- [89] Kroupi E, Yazdani A, Ebrahimi T. EEG correlates of different emotional states elicited during watching music videos. In *International Conference on Affective Computing and Intelligent Interaction*, Berlin, Heidelberg, 2011, pp 457–466.
- [90] Liu JX, Meng HY, Li MZ, et al. Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction. *Concurrency Computat Pract Exper.* 2018, **30**(23): e4446.
- [91] Atkinson J, Campos D. Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Syst Appl.* 2016, **47**: 35–41.
- [92] Hajcak G, MacNamara A, Olvet DM. Event-related potentials, emotion, and emotion regulation: an integrative review. *Dev Neuropsychol.* 2010, **35**(2): 129–155.
- [93] Fraedrich EM, Lakatos K, Spangler G. Brain activity during emotion perception: the role of attachment representation. *Attach Hum Dev.* 2010, **12**(3): 231–248.
- [94] Martini N, Menicucci D, Sebastiani L, et al. The dynamics of EEG gamma responses to unpleasant visual stimuli: from local activity to functional connectivity. *Neuroimage.* 2012, **60**(2): 922–932.
- [95] Bozhkov L, Georgieva P, Santos I, et al. EEG-based subject independent affective computing models. *Procedia Comput Sci.* 2015, **53**: 375–382.
- [96] Hjorth B. EEG analysis based on time domain properties. *Electroencephalogr Clin Neurophysiol.* 1970, **29**(3): 306–310.
- [97] Hausdorff JM, Lertratanakul A, Cudkowicz ME, et al. Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. *J Appl Physiol.* 2000, **88**(6): 2045–2053.
- [98] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Phys D: Nonlinear Phenom.* 1988, **31**(2): 277–283.
- [99] Affinito M, Carrozzi M, Accardo A, et al. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biol Cybern.* 1997, **77**(5): 339–350.
- [100] Kemp AH, Silberstein RB, Armstrong SM, et al. Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *Neuroimage.* 2004, **21**(2): 632–646.
- [101] Duan RN, Zhu JY, Lu BL. Differential entropy feature for EEG-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. San Diego, CA, USA, 2013, pp 81–84.
- [102] Shi LC, Jiao YY, Lu BL. Differential entropy feature for EEG-based vigilance estimation. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013, pp 1–28.
- [103] Candra H, Yuwono M, Chai R, et al. Investigation of window size in classification of EEG-emotion signal with wavelet entropy and support vector machine. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan, Italy, 2015, pp 7250–7253.
- [104] Subasi A. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Syst Appl.* 2007, **32**(4): 1084–1093.
- [105] Mohammadi Z, Frounchi J, Amiri M. Wavelet-based emotion recognition system using EEG signal. *Neural Comput & Applic.* 2017, **28**(8): 1985–1990.
- [106] Winkler I, Jäger M, Mihajlović V, et al. Frontal EEG Asymmetry based classification of emotional valence using common spatial patterns. *World Academy of Science, Engineering and Technology.* 2010, **70**: 373–378.
- [107] Balconi M, Lucchiari C. Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis. *Int J Psychophysiol.* 2008, **67**(1): 41–46.
- [108] Bastos AM, Schoffelen JM. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front Syst Neurosci.* 2015, **9**: 175.
- [109] Wu X, Zheng WL, Lu BL. Investigating EEG-based functional connectivity patterns for multimodal emotion recognition. *arXiv preprint.* 2020: arXiv: 2004.01973.
- [110] Chen M, Han J, Guo L, et al. Identifying valence

- and arousal levels via connectivity between EEG channels. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. Xi'an, China, 2015, pp 63–69.
- [111] Moon SE, Chen CJ, Hsieh CJ, et al. Emotional EEG classification using connectivity features and convolutional neural networks. *Neural Networks*. 2020, **132**: 96–107.
- [112] Liu XC, Li T, Tang C, et al. Emotion recognition and dynamic functional connectivity analysis based on EEG. *IEEE Access*. 2019, **7**: 143293–143302.
- [113] Jenke R, Peer A, Buss M. Feature extraction and selection for emotion recognition from EEG. *IEEE Trans Affective Comput*. 2014, **5**(3): 327–339.
- [114] Nakisa B, Rastgoo MN, Tjondronegoro D, et al. Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. *Expert Syst Appl*. 2018, **93**: 143–155.
- [115] Soleymani M, Pantic M, Pun T. Multimodal emotion recognition in response to videos. *IEEE Trans Affective Comput*. 2012, **3**(2): 211–223.
- [116] Hwang S, Hong K, Son G, et al. Learning CNN features from DE features for EEG-based emotion recognition. *Pattern Anal Applic*. 2020, **23**(3): 1323–1335.
- [117] Jang S, Moon SE, Lee JS. EEG-based video identification using graph signal modeling and graph convolutional neural network. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, Canada, 2018, pp 3066–3070.
- [118] Song TF, Zheng WM, Song P, et al. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans Affective Comput*. 2020, **11**(3): 532–541. <https://doi.org/10.1109/taffc.2018.2817622>
- [119] Xu HY, Platanotis KN. Affective states classification using EEG and semi-supervised deep learning approaches. In *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, Montreal, QC, Canada, 2016, pp 1–6.
- [120] Li XW, Zhao QL, Liu L, et al. Improve affective learning with EEG approach. *Comput Inform*. 2010, **29**: 557–570.
- [121] Specht DF. Probabilistic neural networks. *Neural Networks*. 1990, **3**(1): 109–118.
- [122] Zhang J, Chen M, Hu S, et al. PNN for EEG-based emotion recognition. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, 2016, pp 2–30.
- [123] Subasi A, Ismail Gursoy M. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst Appl*. 2010, **37**(12): 8659–8666.
- [124] Li M, Xu HP, Liu XW, et al. Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. *Technol Heal Care*. 2018, **26**: 509–519.
- [125] Jiang JF, Zeng Y, Lin ZM, et al. Review on EEG-based emotion assessment (in Chinese). *J Inform Eng Univ*, 2016, **17**: 686–693.
- [126] Bashivan P, Rish I, Yeasin M, et al. Learning representations from EEG with deep recurrent-convolutional neural networks. *arXiv preprint*. 2015: arXiv:1511.06448.
- [127] Such FP, Sah S, Dominguez MA, et al. Robust spatial filtering with graph convolutional neural networks. *IEEE J Sel Top Signal Process*. 2017, **11**(6): 884–896.
- [128] Wang ZM, Tong Y, Heng X. Phase-locking value based graph convolutional neural networks for emotion recognition. *IEEE Access*. 2019, **7**: 93711–93722.
- [129] Wang XH, Zhang T, Xu XM, et al. EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Madrid, Spain, 2018, pp 1240–1244.
- [130] Wulsin DF, Gupta JR, Mani R, et al. Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement. *J Neural Eng*. 2011, **8**(3): 036015.
- [131] Jia XW, Li K, Li XY, et al. A novel semi-supervised deep learning framework for affective state recognition on EEG signals. In *2014 IEEE International Conference on Bioinformatics and Bioengineering*, Boca Raton, FL, USA, 2014, pp 30–37.



Wanrou Hu is currently a graduate student pursuing her Master's degree in the School of Biomedical Engineering, Health Science Center, Shenzhen University, China. She received her B.Eng. degree from Guangzhou Medical University, Guangzhou, China, in 2019. Her research interest focuses on affective computing. E-mail: huwanrou2019@email.szu.edu.cn



Gan Huang received his Ph.D. degree from Shanghai Jiao Tong University, China (2013) and then received his postdoctoral training in Université Catholique de Louvain in Belgium. He is now an assistant professor in School of Biomedical Engineering, Shenzhen University, China. He has published several high-quality papers on journals including *Neuroimage*, *Neural Network*, and *Neurocomputing*. His current research interests focus on neural modulation, brain-computer interface and neuroprosthesis. E-mail: huanggan1982@gmail.com



Linglin Li received her Ph. D. degree from the University of Chinese Academy of Sciences, China (2017) and then received her postdoctoral training in University of Shenzhen, China. She is now an associate researcher in School of Biomedical Engineering, Health Science Center, University of Shenzhen, China. Her current research interests focus on the development of EEG neurofeedback technique and related neuroimaging mechanisms. E-mail: lilinling@szu.edu.cn



Li Zhang received his B.Sc. degree from the Macau University of Science and Technology in 2010, M. Sc. and Ph.D. degrees from the University of Hong Kong in 2011 and 2016, respectively. He is currently a research associate professor in School of Biomedical Engineering, Shenzhen University. His researches mainly focus on biomedical signal processing, machine learning, bioinformatics, neuroinformatics and numerical optimization. E-mail: lzhang@szu.edu.cn



Zhiguo Zhang received his B.Eng. degree from Tianjin University in 2000, his M.Phil. degree from the University of Science and Technology of China in 2003, and his Ph.D. degree from the University of Hong Kong in 2008. He is now a professor with School of Biomedical Engineering, Health Science Center, Shenzhen University, China. His research interests include biomedical signal processing, neural engineering, and brain-inspired computation. E-mail: zgzhang@szu.edu.cn



Zhen Liang received her B.Eng. degree from Yanshan University, China, in 2008 and Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2013. From 2012 to 2017, she was an algorithm development scientist in NeuroSky, Inc., Hong Kong. From 2018 to 2019, she was a specially-appointed assistant professor of Graduate School of Informatics, Kyoto University, Japan. She is currently an assistant professor in the School of Biomedical Engineering, Health Science Center, Shenzhen University, China. Her current research interests include brain encoding and decoding system, affective computing, visual attention, neural engineering. E-mail: janezliang@szu.edu.cn