



OPEN ACCESS

EDITED BY
Wenming Zheng,
Southeast University,
China

REVIEWED BY
Guijun Chen,
Taiyuan University of Technology,
China
Shou-Yu Chen,
Tongji University,
China

*CORRESPONDENCE
Yinzhen Huang
✉ hyz@hnit.edu.cn

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

SPECIALTY SECTION
This article was submitted to
Perception Science,
a section of the journal
Frontiers in Psychology

RECEIVED 19 December 2022
ACCEPTED 11 January 2023
PUBLISHED 27 February 2023

CITATION
Wang X, Ren Y, Luo Z, He W, Hong J and
Huang Y (2023) Deep learning-based EEG
emotion recognition: Current trends and future
perspectives.
Front. Psychol. 14:1126994.
doi: 10.3389/fpsyg.2023.1126994

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

Deep learning-based EEG emotion recognition: Current trends and future perspectives

Xiaohu Wang^{1†}, Yongmei Ren^{2†}, Ze Luo¹, Wei He², Jun Hong¹ and Yinzhen Huang^{3*}

¹School of Intelligent Manufacturing and Mechanical Engineering, Hunan Institute of Technology, Hengyang, China, ²School of Electrical and Information Engineering, Hunan Institute of Technology, Hengyang, China, ³School of Computer and Information Engineering, Hunan Institute of Technology, Hengyang, China

Automatic electroencephalogram (EEG) emotion recognition is a challenging component of human–computer interaction (HCI). Inspired by the powerful feature learning ability of recently-emerged deep learning techniques, various advanced deep learning models have been employed increasingly to learn high-level feature representations for EEG emotion recognition. This paper aims to provide an up-to-date and comprehensive survey of EEG emotion recognition, especially for various deep learning techniques in this area. We provide the preliminaries and basic knowledge in the literature. We review EEG emotion recognition benchmark data sets briefly. We review deep learning techniques in details, including deep belief networks, convolutional neural networks, and recurrent neural networks. We describe the state-of-the-art applications of deep learning techniques for EEG emotion recognition in detail. We analyze the challenges and opportunities in this field and point out its future directions.

KEYWORDS

human–computer interaction, electroencephalogram, emotion recognition, deep learning, survey

1. Introduction

Emotion recognition (or detection) is a major scientific problem in affective computing, which mainly solves the problem of computer systems accurately processing, recognizing, and understanding the emotional information expressed by human beings. Affective computing requires interdisciplinary knowledge, including psychology, biology, and computer science. As emotion plays a key role in the field of human–computer interaction (HCI) and artificial intelligence, it has recently received extensive attention in the field of engineering research. Research of emotion recognition technology can further promote the development of various disciplines, including computer science, psychology, neuroscience, human factors engineering, medicine, and criminal investigation.

As a complex psychological state, emotion is related to physical behavior and physiological activities (Cannon, 1927). Researchers have conducted numerous studies to enable computers to correctly distinguish and understand human emotions. These studies aim to enable computers to generate various emotional features similar to human beings, so as to achieve the purpose of natural, sincere, and vivid interaction with human beings. Some of these methods mainly use non physiological signals, such as speech (Zhang et al., 2017; Khalil et al., 2019; Zhang S. et al., 2021), facial expression (Alreshidi and Ullah, 2020), and body posture (Piana et al., 2016). However, their accuracy depends on people's age and cultural

characteristics, which are subjective, so accurately judging the true feelings of others is difficult. Other methods use physiological activities (or physiological clues), such as heart rate (Quintana et al., 2012), skin impedance (Miranda et al., 2018), respiration (Valderas et al., 2019) or brain signals, functional magnetic resonance imaging (Chen et al., 2019a), magnetoencephalography (Kajal et al., 2020), and electroencephalography, to identify emotional states. Some studies have shown that physiological activities and emotional expression are correlated, although the sequence of the two processes is still debated (Cannon, 1927). Therefore, the method based on calculating physiological signals is considered an effective supplement to the recognition method based on nonphysiological signals. The subject cannot control the automatically generated electroencephalogram (EEG) signal. For those who cannot speak clearly and express their feelings through natural speech or have physical disabilities and cannot express their feelings through facial expressions or body postures, emotion recognition of voice, expression, and posture becomes impossible. Therefore, EEG is an appropriate means to extract human emotions, and studying emotional cognitive mechanisms and recognizing emotional states by directly using brain activity information, such as EEG, are particularly important.

From the perspective of application prospects, EEG-based emotion recognition technology has penetrated into various fields, including medical, education, entertainment, shopping, military, social, and safe driving (Suhaimi et al., 2020). In the medical field, timely acquisition of patients' EEG signals and rapid analysis of their emotional state can help doctors and nurses to accurately understand the patients' psychological state and then make reasonable medical decisions, which has an important effect on the rehabilitation of some people with mental disorders, such as autism (Mehdizadehfar et al., 2020; Mayor-Torres et al., 2021; Ji et al., 2022), depression (Cai et al., 2020; Chen X. et al., 2021), Alzheimer's disease (Güntekin et al., 2019; Seo et al., 2020), and physical disabilities (Chakladar and Chakraborty, 2018). In terms of education, the emotion recognition technology based on EEG signals can enable teaching staff to adjust teaching methods and teaching attitudes in a timely manner in accordance with the emotional performance of different trainees in class, such as increasing or reducing the workload (Menezes et al., 2017). In terms of entertainment, such as computer games, researchers try to detect the emotional state of players to adapt to the difficulty, punishment, and encouragement of the game (Stavroulia et al., 2019). In the military aspect, the emotional status of noncommissioned officers and soldiers can be captured timely and accurately through EEG signals, so that the strategic layout can be adjusted in time to improve the winning rate of war (Guo et al., 2018). In terms of social networks, we can enhance barrier-free communication in the HCI system, increase the mutual understanding and interaction in the human-machine-human interaction channel, and avoid some unnecessary misunderstandings and frictions through the acquisition of emotional information (Wu et al., 2017). In terms of safe driving, timely detection of EEG emotional conditions can enable a vehicle to perform intelligent locking during startup to block driving or actively open the automatic driving mode to intervene in the vehicle's motion trajectory until parking at a safe position, thereby greatly reducing the occurrence of accidents (Fan et al., 2017).

Recently, automatic recognition of emotional information from EEG has become a challenging problem, and has attracted extensive attention in the fields of artificial intelligence and computer vision. The flow of emotion recognition research is shown in Figure 1. Essentially, human emotion recognition using EEG signals belongs to one type of pattern recognition research.

In the early EEG-based automated emotion recognition literature, a variety of machine learning-based studies, such as support vector machine (SVM; Lin et al., 2009; Nie et al., 2011; Jie et al., 2014; Candra et al., 2015), k-nearest neighbor (KNN; Murugappan et al., 2010; Murugappan, 2011; Kaundanya et al., 2015), linear regression (Bos, 2006; Liu et al., 2011), support vector regression (Chang et al., 2010; Soleymani et al., 2014), random forest (Lehmann et al., 2007; Donos et al., 2015; Lee et al., 2015), and decision tree (Kuncheva et al., 2011; Chen et al., 2015), have been developed.

Although the abovementioned hand-crafted EEG signal features associated with machine learning approaches can produce good domain-invariant features for EEG emotion recognition, they are still low-level and not highly discriminative. Thus, obtaining high-level domain-invariant feature representations for EEG emotion recognition is desirable.

The recently-emerged deep learning methods may present a possible solution to achieve high-level domain-invariant feature representations and high-precision classification results of EEG emotion recognition. The representative deep learning techniques contain recurrent neural networks (RNNs; Elman, 1990), long short-term memory (LSTM; Hochreiter and Schmidhuber, 1997; Zhang et al., 2019), deep belief networks (DBNs; Hinton et al., 2006), and convolutional neural networks (CNNs; Krizhevsky et al., 2012). To date, deep learning techniques have shown outstanding performance on object detection and classification (Wu et al., 2020), natural language processing (Otter et al., 2020), speech signal processing (Purwins et al., 2019), and multimodal emotion recognition (Zhou et al., 2021) due to its strong feature learning ability. Figure 2 shows the evolution of EEG emotion recognition with deep learning algorithms, emotion categories and databases.

Inspired by the lack of summarizing the recent advances in various deep learning techniques for EEG-based emotion recognition, this paper aims to present an up-to-date and comprehensive survey of EEG emotion recognition, especially for various deep learning techniques in this area. This paper highlights the different challenges and opportunities on EEG emotion recognition tasks and points out its future trends. In this survey, we have searched the published literature between January 2012, and December 2022 through Scholar, google, ScienceDirect, IEEEExplore, ACM, Springer, PubMed, and Web of Science, on the basis of the following keywords: "EEG emotion recognition," "emotion computing," "deep learning," "RNNs," "LSTM," "DBNs," and "CNNs." There is no any language restriction for the searching process.

In this work, our contributions can be summarized as follows:

1. We provide an up-to-date literature survey on EEG emotion recognition from a perspective of deep learning. To the best of our knowledge, this is the first attempt to present a comprehensive review covering EEG emotion recognition and deep learning-based feature extraction algorithms in this field.

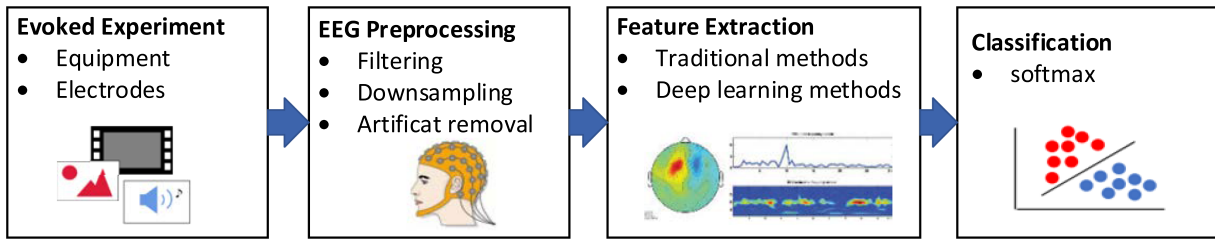


FIGURE 1
Flowchart of emotion recognition using EEG signals.

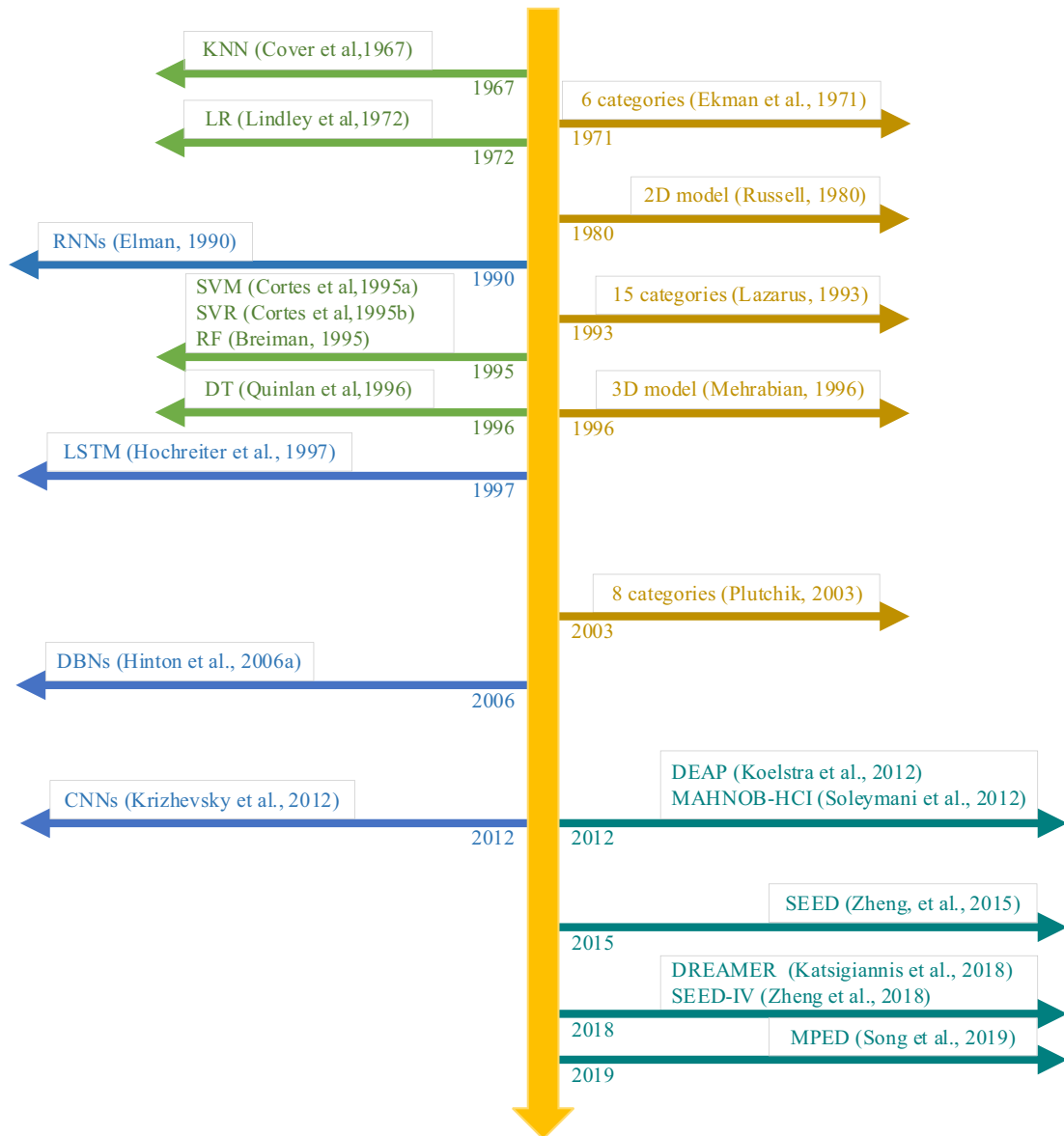


FIGURE 2
The evolution of EEG emotion recognition with deep learning algorithms, emotion categories and databases.

2. We analyze and discuss the challenges and opportunities faced to EEG emotion recognition and point out future directions in this field.

The organization of this paper is as follows. We first present the preliminaries and basic knowledge of EEG emotion recognition. We review benchmark datasets and deep learning techniques in

detail. We show the recent advances of the applications of deep learning techniques for EEG emotion recognition. We give a summary of open challenge and future directions. We provide the concluding remarks.

2. Preliminaries and basic knowledge

2.1. Definition of affective computing

Professor Picard (1997) of the MIT and his team clearly defined affective computing, that is, the calculation of factors triggered by emotion, related to emotion, or able to affect and determine emotional change. In accordance with the research results in the field of emotion, emotion is a mechanism gradually formed in the process of human adaptation to social environment. When different individuals face the same environmental stimulus, they may have the same or similar emotional changes or they may have different emotional changes due to the difference in individual living environment. This psychological mechanism can play a role in seeking advantages and avoiding disadvantages. Although computers have strong logic computing ability, human beings cannot communicate more deeply when interacting with computers due to the lack of psychological mechanisms similar to human beings. Emotion theory is an effective means to solve this problem. Therefore, an effective method to realize computer intelligence is to combine logical computing with emotional computing, which is a research topic that many researchers focus on at present (Zhang S. et al., 2018).

2.2. Classification of emotional models

Many researchers cannot reach a unified emotional classification standard when conducting emotional computing research due to the high complexity and abstractness of emotion. At present, researchers usually divide emotion models into discrete model and dimensional space model.

In the discrete model, each emotion is distributed discretely, and these discrete emotions combine to form the human emotional world. In the discrete model, designers have different definitions of emotions, and they are divided into different emotional categories. American psychologist Ekman and Friesen (1971) divided human emotions into six basic emotions, namely, anger, disgust, fear, happiness, sadness, and surprise, by analyzing human facial expressions. Lazarus (Lazarus, 1993), one of the modern representatives of American stress theory, divided emotions into 15 categories, such as anger, anxiety and happiness, and each emotional state has a corresponding core related theme. Psychologist Plutchik (2003) divided emotions into eight basic categories: anger, fear, sadness, disgust, expectation, surprise, approval, and happiness. These discrete emotion classification methods are relatively simple and easy to understand, and have been widely used in many emotion recognition studies.

The dimensional space model of emotion can be divided into 2D and 3D. The 2D expression model of emotion was first proposed by psychologist Russell (1980). It uses 2D coordinate axis to describe the polarity and intensity of emotion. Polar axis is used to describe the positive and negative types of emotion, and intensity coordinate axis refers to the intensity of emotion. The 2D emotion model is consistent

with people's cognition of emotion. Currently, the VA model that divides human emotions into two dimensions is widely used, which are the valency dimension and arousal dimension, as shown in Figure 3.

Considering that the 2D space representation of emotions cannot effectively distinguish some basic emotions, such as fear and anger, Mehrabian (1996) proposed a 3D space representation of emotions, and its three dimensions are pleasure, activation, and dominance, as shown in Figure 4. Centered on the origin, pleasure (P) represents the difference between positive and negative emotions; arousal (A) indicates the activation degree of human emotions; dominance (D) indicates the degree of human control over current things. At the same time, the coordinate values of the three dimensions can be used to describe specific human emotions.

2.3. Deep learning techniques

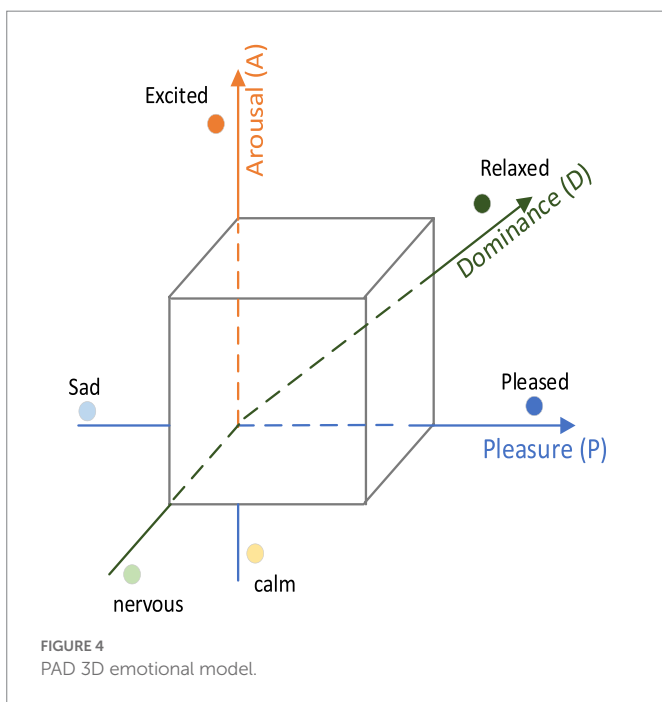
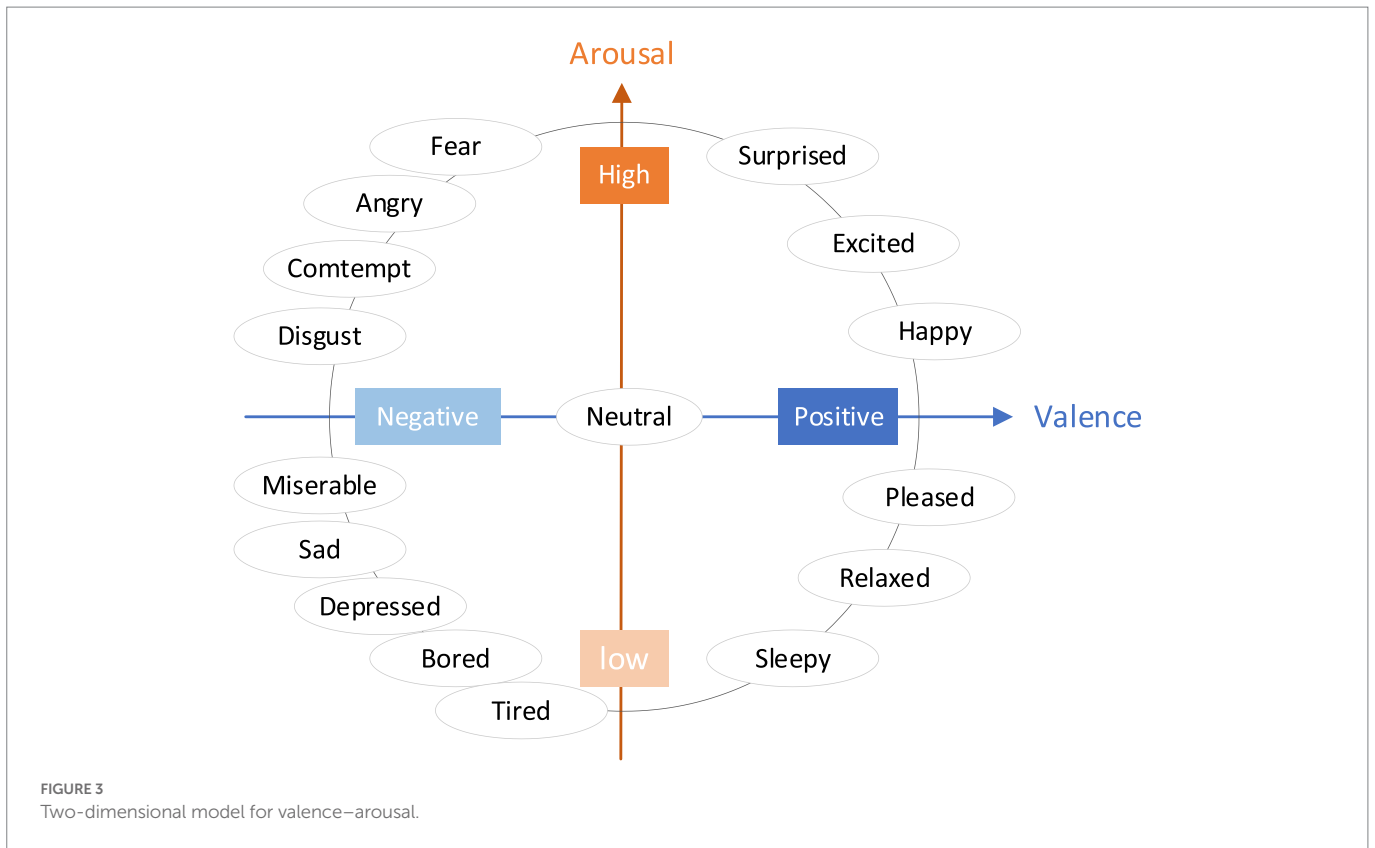
2.3.1. DBNs

DBNs proposed by Hinton et al. (2006) are a generative model aim to train the weights among its neurons and make the entire neural network generate training data in accordance with the maximum probability.

At present, DBN has been applied to many areas of life, such as voice, graphics, and other visual data classification tasks, and achieved good recognition results. Tong et al. used a DBN model to classify hyperspectral remote sensing images, improved the training process of DBN, and used the hyperspectral remote sensing image dataset Salinas to verify the proposed method (Tong et al., 2017). Compared with traditional model classification methods, the classification accuracy of DBN model can reach more than 90%. In the text classification event, Payton L et al. proposed an ME learning algorithm for DBN (Payton et al., 2016). This algorithm is specially designed to deal with limited training data. Compared with the maximum likelihood learning method, the method of maximizing the entropy of parameters in DBN has more effective generalization ability, less data distribution deviation, and robustness to over fitting. It achieves good classification effect on Newsgroup, WebKB, and other datasets. The DBN model also achieves good classification results in speech classification events. Wen et al. tried to recognize human emotions from speech signals (Wen et al., 2017) using a random DBN (RDBN) integrating method. The experimental results on the benchmark speech emotion database show that the accuracy of RDBN is higher than that of KNN and other speech emotion recognition methods. Kamada S and Ichimura T extended the learning algorithm of adaptive RBM and DBN to time series analysis by using the idea of short-term memory and used the adaptive structure learning method to search the optimal network structure of DBN in the training process. This method was applied to MovingMNIST, a benchmark dataset for video recognition, and its prediction accuracy exceeded 90% (Kamada and Ichimura, 2019).

2.3.2. CNNs

The concept of neural networks originated from the neural mathematical model first proposed in 1943 (McCulloch and Pitts, 1943). However, the artificial neural networks confined to the shallow network architecture fell into a low tide in the late 1960s due to the constraints of early computing power, data and other practical conditions. The real rise of neural network method began with AlexNet (Krizhevsky et al., 2012) proposed by Hinton et al. (2012). This CNN model won the Image Net Large Scale Visual Recognition Challenge (Deng et al., 2009) with a huge



advantage of 10.9%. Therefore, deep neural networks (DNNs) have gradually attracted extensive attention from the industry and academia. In a broad sense, the DNNs can be divided into feedforward neural networks (Rumelhart et al., 1986; Lecun and Bottou, 1998) and RNNs (Williams and Zipser, 1989; Yi, 2010; Yi, 2013) in accordance with the difference in connection modes between neurons. In accordance with the differences in use scenarios and HCI methods, the DNNs can

be divided into single input networks and multi-input networks, which are continuously extended to a variety of HCI scenarios and have achieved breakthrough results (Wang et al., 2018), allowing AI products to be actually implemented in practical applications.

2.3.3. RNNs

Compared with CNNs, RNNs are better in processing data with sequence characteristics and can obtain time-related information in data (Lecun et al., 2015). Researchers have widely used RNNs in natural language processing, including machine translation, text classification, and named entity recognition (Schuster and Paliwal, 1997). RNNs have achieved outstanding performance in audio-related fields and made great breakthroughs. They have been widely used in speech recognition, speech synthesis, and other fields. Considering that RNN only considers the preorder information and ignores the postorder information, a bidirectional RNN (BRNN) was proposed (Schuster and Paliwal, 1997). To solve the problems of gradient disappearance and gradient explosion in the training process of the long sequence of RNN, researchers improved the structure of RNN and built a LSTM (Hochreiter and Schmidhuber, 1997). LSTM has modified the internal structure of RNN in the current time step, making the hidden layer architecture more complex, which can have a better effect in longer sequences and is a more widely used RNN in the general sense.

The bidirectional LSTM (BILSTM) can be obtained by combining LSTM and BRNN (Graves and Schmidhuber, 2005). It replaces the original RNN neuron structure in BRNN with the neuron structure of LSTM and combines the forward LSTM and backward LSTM to form a network. BILSTM retains the advantages of BRNN and LSTM at the same time. It can retain the context information of the current time node and record the relationship between the front and back features. Therefore, BILSTM improves the generalization ability of the network

model and the ability to handle long sequences, and avoids the problem of gradient explosion and gradient disappearance according to the difference of connection modes between neurons.

In recent years, gated recurrent unit (GRU) networks (Cho et al., 2014) were proposed. They discard the three LSTM gated (force gate, input gate, and output gate) networks, selects reset gate and update gate, and combines the current state of neurons and the hidden layer state, which are uniformly expressed as h_t . Compared with LSTM, the GRU model is simpler, with fewer parameters and easier convergence.

3. Benchmark datasets

The proposed EEG emotion recognition algorithms should be verified on EEG data with emotion ratings or labels. However, some researchers are limited by conditions and cannot build a special experimental environment. Most researchers are interested in verifying their algorithms and comparing with relevant studies on the recognized benchmark datasets. Hence, a variety of open-source EEG emotion databases have been developed for EEG emotion recognition. Table 1 presents a brief summary of existing speech emotion databases. In this section, we describe briefly these existing EEG emotion databases as follows.

3.1. Database for emotion analysis using physical signals

DEAP is a large multimodal physiological and emotional database jointly collected and processed by Koelstra and other research institutions of four famous universities (Queen Mary University in London, Twente University in the Netherlands, Geneva University in Switzerland, and Swiss Federal Institute of Technology; Koelstra et al., 2012). The collection scene for the DEAP database is shown in Figure 5, and it is an open-source data set for analyzing human emotional states. The DEAP database collected 32 participants for the experiment, where 16 of them were male and 16 were female. In the experiment, the EEG and peripheral physiological signals of the participants were collected, and the frontal facial expression videos of the first 22 participants were recorded. The participants read the instructions of the experiment process and wore the detection equipment before starting the data acquisition experiment. Each participant watched 40 music video clips with a duration of 1 min in the experiment. The subjective emotional experience in induction experiments was self-evaluated and rated on assessment scales that

cover four emotional dimensions, namely, arousal, valence, dominance, and like. During self-assessment, the participants saw the content displayed on the screen and clicked to select the option that matched their situation at that time. The EEG and peripheral physiological signals were recorded by using a Biosemi ActiveTwo system. The EEG information was collected by using electrode caps with 32 AgCl electrodes. The EEG sampling rate was 512 Hz. The data set recorded 40 channels in total, the first 32 channels were EEG signal channels, and the last 8 channels were peripheral signal channels.

3.2. Multimodal database for affect recognition and implicit tagging

MAHNOB-HCI is a multimodal physiological emotion database collected by Soleymani et al. (2014) through a reasonable and normal experimental paradigm. The MAHNOB-HCI dataset collected EEG signals and peripheral physiological signals from 30 volunteers with different cultural and educational backgrounds using emotional stimulation videos. Among the 30 young healthy adult participants, 17 were women and 13 were men, and the age ranged from 19 to 40. Thirty participants watched 20 different emotional video clips selected from movies and video websites. These video clips can stimulate the subjects to have five emotions: disgust, amusement, fear, sadness, and joy. The duration of watching videos was 35 to 117 s. The participants evaluated the arousal and potency dimensions rated on assessment scales after watching each video clip. In the data collection experiment, six cameras were used to record the facial expressions of the subjects at a frame rate of 60 frames per second. The collection scene for the MAHNOB-HCI database is shown in Figure 6.

3.3. SJTU emotion EEG dataset

The SEED is an EEG emotion dataset released by the BCMI Research Center in Shanghai Jiaotong University (SJTU; Zheng and Lu, 2015), and the protocol used in the emotion experiment is shown in Figure 7. The SEED dataset selected 15 people (7 men and 8 women) as the subjects of the experiment and collected data of 62 EEG electrode channels from the participants. In the experiment, 15 clips of Chinese movies were selected for the subjects to watch. These videos contained three types of emotions: positive, neutral, and negative. Each genre had five clips, and each clip was about 4 min. Clips containing different emotions appeared alternately. In the experiment, the subjects

TABLE 1 Description of public datasets.

| Name | Participants | Documented Signals | Stimulus | Task models/ Emotions |
|------------|--------------|--|-----------------------------|---|
| DEAP | 32 | EEG, EMG, EOG, GSR, Temperature, and Face Video | 40 Video clips | VAD model |
| MAHNOB-HCI | 27 | EEG, ECG, GSR, ERG, Respiration Amplitude, Skin Temperature, Face Video, Audio Signals, and Eye Gaze | 20 Video clips and Pictures | VAD model |
| SEED | 15 | EEG, Face Video, and Eye tracking | 15 Video clips | Positive, Neutral, and Negative |
| DREAMER | 23 | EEG, ECG | 18 Video clips | VAD model |
| SEED-IV | 15 | EEG, and EM | 168 Video clips | Happiness, Sadness, Fear and Neutrality |
| MPED | 23 | EEG, ECG, RSP, and GSR | 28 Video clips | Joy, Funny, Anger, Fear, Disgust, Sadness, and Neutrality |

had a 5-s prompt before watching each video. The subjects conducted a 45 s self-assessment, followed by a 15 s rest. During the experiment, the subjects were asked to complete three experiments repeatedly in a week

or even longer. Each subject watched the same 15 clips of video and recorded their self-evaluation emotions.



FIGURE 5
Collection scene for the Deap database (Koelstra et al., 2012).
Reproduced with permission from IEEE. Licence ID: 1319273-1.



FIGURE 6
Collection scene for the MAHNOB-HCI database (Soleymani et al., 2014).
Reproduced with permission from IEEE. Licence Number: 5493400200365.

3.4. Dreamer

The Dreamer database (Katsigiannis and Ramzan, 2018) is a multimodal physiological emotion database released by the University of Western Scotland in 2018. It contains 18 audio-visual clips during the emotion induction experiments and collects the EEG and electrocardiogram (ECG) signals simultaneously. The video duration is between 65 and 393 s, with an average duration of 199 s. Twenty-three subjects with an average age of 26.6 years were invited to participate in the experiment. The subjects were asked to conduct a self-assessment between 1 and 5 points in the emotional dimensions of valence, arousal, and dominance after each emotional induction experiment.

3.5. Seed-iv

SEED-IV is another version of the SEED dataset released by SJTU (Zheng et al., 2018), which has been widely used in recent related work. The protocol of SEED-IV for four emotions is shown in Figure 8. Forty-four participants (22 women, all college students) were recruited to self-evaluate their emotions during the induction experiment, and 168 film clips were selected as the material library of four emotions (happiness, sadness, fear, and neutrality). It follows the experimental paradigm adopted in SEED, 62-channel EEG of 15 selected subjects were recorded in the three tests. They chose 72 film clips with four different emotional labels (neutral, sad, fear, and happy). Each subject watched six film clips in each session, resulting in 24 trials in total.

3.6. Multi-modal physiological emotion database

The MPED contains four physiological signals of 23 subjects (10 men and 13 women) and records seven types of discrete emotion (joy, funny, anger, fear, disgust, sadness, and neutrality) when they watch 28

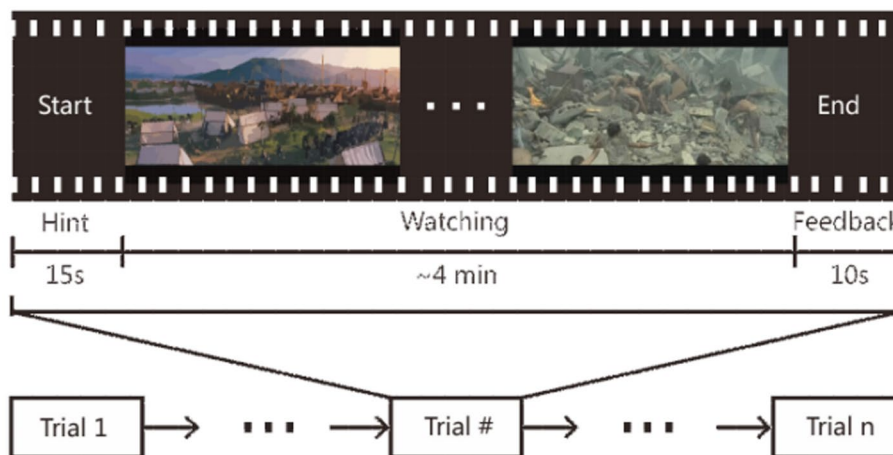
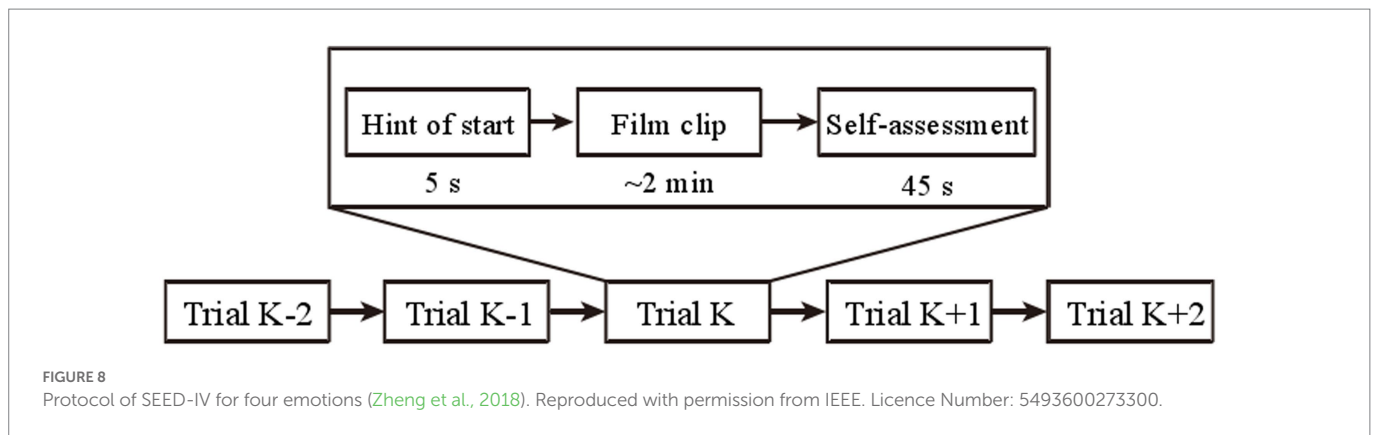


FIGURE 7
Protocol used in the emotion experiment (Zheng and Lu, 2015).
Reproduced with permission from IEEE. Licence Number: 5493480479146.



video clips (Song et al., 2019). The experiments are divided into two sessions with an interval of at least 24 h. Twenty-one EEG data are used as training data, and the remaining 7 EEG data are used as test data.

4. Review of EEG emotion recognition techniques

4.1. Shallow machine learning methods for EEG emotion recognition

The emotion recognition method of EEG signals based on machine learning is usually divided into two steps: manual feature extraction and classifier selection (Li et al., 2019). The feature extraction methods mainly include time domain analysis, frequency domain analysis, time frequency domain analysis, multivariate statistical analysis, and nonlinear dynamic analysis (Donos et al., 2018; Rasheed et al., 2020). Principal component analysis, linear discriminant analysis (LDA; Alotaiby et al., 2017), and independent component analysis (Ur et al., 2013; Acharya et al., 2018a; Maimaiti et al., 2021) are widely used unsupervised time-domain methods to summarize EEG data. Frequency domain features include spectral center, coefficient of variation, power spectral density, signal energy, spectral moment, and spectral skewness, which can provide key information about data changes (Yuan et al., 2018; Acharya et al., 2018a). The abovementioned time-domain or frequency-domain methods have limitations and cannot provide accurate frequency or time information at a specific time point. Wavelet transformation (WT) is usually used to decompose EEG signals into their frequency components to express the relationship between signal information and time. Time frequency signal processing algorithms, such as discrete wavelet transform analysis and continuous wavelet transform, are a necessary means to solve different EEG behavior, which can be described in the time and frequency domains (Martis et al., 2012; UR et al., 2013). Statistical parameters, such as mean, variance, skewness, and kurtosis, have been widely used to extract feature information from EEG signals. Variance represents the distribution of data, skewness represents the symmetry information of data, and kurtosis provides the peak information in data (Acharya et al., 2018a).

In classifier selection, previous work mainly used shallow machine learning methods, such as LDA, SVM, and KNN, to train emotion recognition models based on manual features. Although the method of “manual features+shallow classifier” has made some progress in previous emotion recognition systems, the design of manual features requires

considerable professional knowledge, and the extraction of some features (such as linear features) is time consuming.

4.2. Deep learning for EEG emotion recognition

Traditional machine learning techniques extract EEG features manually, which not only have high redundancy in the extracted features, but also have poor universality. Therefore, manual feature extraction techniques can not achieve the ideal results in EEG emotion recognition. Obviously, with the increasing progress of deep learning technology (Chen B. et al., 2021), EEG emotion recognition research ground on various neural networks has gradually become a research hotspot. Different from shallow classifier, deep learning has the advantages of strong learning ability and good portability, which can automatically learn good feature representations instead of manually design. Recently, various deep learning models, such as DNN, CNN, LSTM, and RNN models, were tested on public datasets. Compared with CNN, RNN is more suitable for processing sequence-related tasks. LSTM has been proven to be capable of capturing time information in the field of emotion recognition (Bashivan et al., 2015; Ma et al., 2019). As a type of sequence data, most studies on EEG are based on RNN and LSTM models. Li et al. (2017) designed a hybrid deep learning model by combining CNN and RNN to mine inter-channel correlation. The results demonstrated the effectiveness of the proposed methods, with respect to the emotional dimensions of Valence and Arousal. Zhang T. et al. (2018) proposed a spatial-temporal recurrent neural network (STRNN) for emotion recognition, which integrate the feature learning from both spatial and temporal information of signal sources into a unified spatial-temporal dependency model, as shown in Figure 9. Experimental results on the benchmark emotion datasets of EEG and facial expression show that the proposed method is significantly better than those state-of-the-art methods. Nath et al. (2020) compared the emotion recognition effects of LSTM with KNN, SVM, DT, and RF. Among them, LSTM has the best robustness and accuracy.

EEG signals are essentially multichannel time series signals. Thus, a more effective method for emotional recognition of EEG signals is to obtain the long-term dependence of EEG signals based on RNN. Li et al. (2020) proposed a BiLSTM network framework based on multimodal attention, which is used to learn the best time characteristics, and inputted the learned depth characteristics into the DNN to predict the emotional output probability of each channel. A

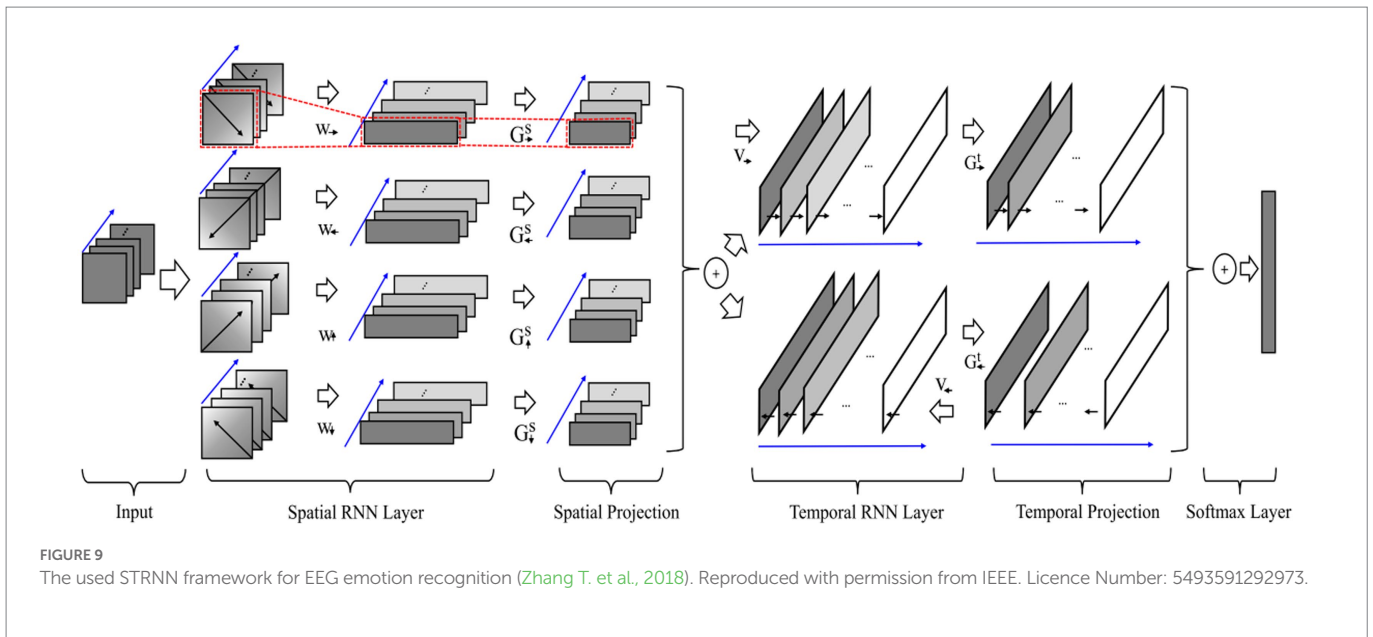


FIGURE 9 The used STRNN framework for EEG emotion recognition (Zhang T. et al., 2018). Reproduced with permission from IEEE. Licence Number: 5493591292973.

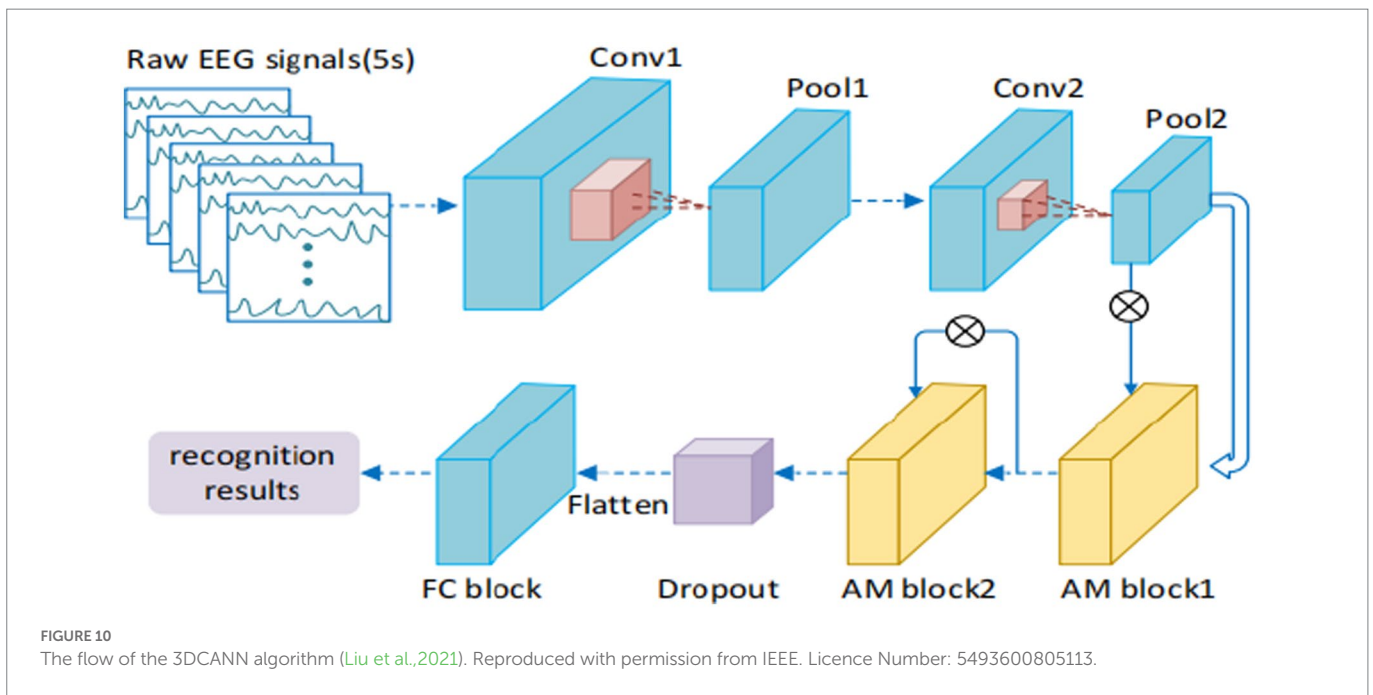


FIGURE 10 The flow of the 3DCANN algorithm (Liu et al., 2021). Reproduced with permission from IEEE. Licence Number: 5493600805113.

decision-level fusion strategy is used to predict the final emotion. The experimental results on AMIGOS dataset show that this method is superior to other advanced methods. Liu et al. (2017) proposed an emotion recognition algorithm model ground on multi-layer long short-term memory recurrent neural network (LSTM-RNN), which combines temporal attention (TA) and band attention (BA). Experiments on Mahnob-HCI database demonstrate the proposed method achieves higher accuracy and boosts the computational efficiency. Liu et al. (2021) studied an original algorithm named three-dimension convolution attention neural network (3DCANN) for EEG emotion recognition, which is composed of spatio-temporal feature extraction module and EEG channel attention weight learning module. Figure 10 presents the details of the used 3DCANN scheme. Alhagry et al. (2017) proposed an end-to-end deep learning neural network to identify emotions from original EEG signals. It uses LSTM-RNN to

learn features from EEG signals and uses full connection layer for classification. Li et al. (2022) proposed a C-RNN model using CNN and RNN, and used multichannel EEG signals to identify emotions. Although the method based on RNN has great advantages in processing time series data and has made great achievements, it still has shortcomings in the face of multichannel EEG data. GRU or LSTM can connect the relationship between different channels through multichannel fusion, but this processing ignores the spatial distribution of EEG channels and cannot reflect the dynamics of the relationship between different channels.

Among various network algorithms, CNNs have a good ability to extract features of convolution kernels. They can extract information features by transferring each part of the image with multiple kernels. They have been widely used in image processing tasks. For EEG signal, they can process raw EEG data well and can be used for spectrum

diagram. Considering that the use of CNN to train EEG data can reduce the effect of noise, most studies use CNN for the emotional recognition of EEG signals to reduce the complexity of training. Thodoroff et al. (2016) combined CNN and RNN to train robust features for automatic detection of seizures. Shamwell et al. (2016) explored a new CNN architecture with 4 convolutional layers and 3 fully connected layers to classify EEG signals. To reduce the over fitting of the model, Manor and Geva (2015) proposed a CNN model based on spatiotemporal regularization, which is used to classify single track EEG in RSVP (fast serial visual rendering). Sakhavi et al. (2015) proposed a parallel convolutional linear network, which is an architecture that can represent EEG data as dynamic energy input, and used CNN for image classification. Ren and Wu (2014) applied convolutional DBN to classify EEG signals. Hajinorozi et al. (2017) used covariance learning to train EEG data for driver fatigue prediction. Jiao et al. (2018) proposed an improved CNN method for mental workload classification tasks. Gao et al. (2020) proposed a gradient particle swarm optimization (GPSO) model to achieve the automatic optimization of the CNN model. The experimental results show that the proposed method based on the GPSO-optimized CNN model achieve a prominent classification accuracy. Figure 11 presents the details of the used GPSO scheme.

CNN can use EEG to identify many human diseases. Antoniadis et al. (2016) used deep learning to automatically generate features of EEG data in time domain to diagnose epilepsy. Page et al. (2016) conducted end-to-end learning through the maximum pool convolution neural network (MPCNN) and proved that transfer learning can be used to teach the generalized characteristics of MPCNN raw EEG data. Acharya et al. (2018b) proposed a five-layer deep CNN for detecting normal, pre seizure, and seizure categories.

The summary of recent state-of-the-art methods related to EEG-based emotion recognition system using machine learning and deep learning approaches is given in Table 2.

5. Open challenges

To date, although a number of literature related to EEG emotion recognition using deep learning technology is reported, showing its certain advance, a few challenges still exist in this area. In the following, we discuss these challenges and opportunities, and point out potential research directions in the future.

5.1. Research on the basic theory of affective computing

At present, the theoretical basis of emotion recognition mainly includes discrete model and continuous model, as shown in Figure 3. Although they are related to each other, they have not formed a unified theoretical framework. The relationship between explicit information (such as happy, sad, and other emotional categories) and implicit information (such as the signal characteristics of different frequency bands of EEG signals corresponding to happy, sad, and other emotional categories) in emotional computing is worthy of further study. Digging out the relationship between them is extremely important for understanding the different emotional states represented by EEG signals.

5.2. EEG emotion recognition data sets

For EEG-based emotion recognition, most publicly available datasets for affective computing use images, videos, audio, and other external methods to induce emotional changes. These emotional changes are passive, which are different from the emotional changes that individuals actively produce in real scenes and may lead to differences

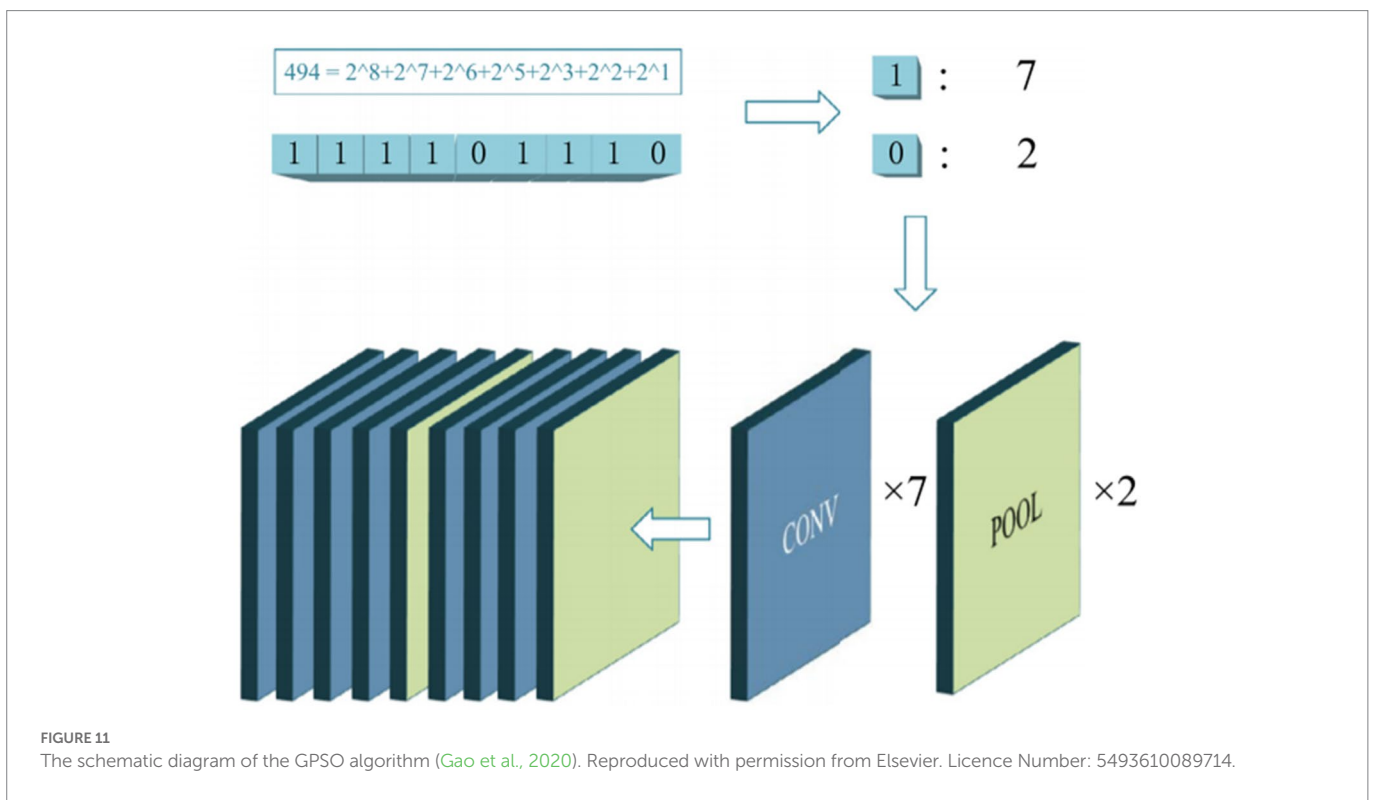


FIGURE 11 The schematic diagram of the GPSO algorithm (Gao et al., 2020). Reproduced with permission from Elsevier. Licence Number: 5493610089714.

TABLE 2 Summary of EEG emotion recognition papers using Deep learning methods from 2017 to 2022.

| Year | References | Stimulus | Classification methods | Emotion | Acc.(%) |
|------|--------------------------------|---------------------------|--|---|--|
| 2017 | Alhagry et al. (2017) | DEAP | LSTM-RNN | Valence, Arousal, and Liking | Arousal: 85.65, Valence: 85.45, Liking: 87.99 |
| 2017 | Li et al. (2017) | DEAP | CNN + RNN | Arousal and Valence | Arousal: 72.06, Valence: 74.12 |
| 2017 | Liu et al. (2017) | Mahnob-HCI | LSTM-RNN | Valence, Arousal, and F1-score | Arousal: 73.1, Valence: 74.5; F1-score: Arousal: 72.3, Valence: 73.0 |
| 2018 | Zhang T. et al. (2018) | SEED and CK+ | STRNN | Positive, Negative, and Neutral | SEED: Overall Accuracy: 89.5 CK+: Overall Accuracy: 95.4 |
| 2018 | Jiao et al. (2018) | The Sternberg memory task | Deep CNN | Types of mental load | Fused CNNs 1: 91.32 Fused CNNs 2: 92.37 |
| 2018 | Song et al. (2018) | SEED and DREAMER | DGCNN | Positive, neutral and negative; Arousal, Valence, and Dominance | SEED: 90.40, DREAMER: Arousal: 84.54 Valence: 86.23 Dominance: 85.02 |
| 2018 | Salama et al. (2018) | DEAP | 3D-CNN | Arousal and Valence | Arousal: 88.49 Valence: 87.44 |
| 2019 | Chao et al. (2019) | DEAP | CapsNet | Valence, Arousal and Dominance | Valence: 66.73 Arousal: 68.28 Dominance: 67.25 |
| 2019 | Gonzalez et al. (2019) | DEAP, IAPS and DREAMER | CNN | Valence and Arousal | Single subject: Valence: 70.26 Arousal: 72.42 |
| 2019 | Garg et al. (2019) | DEAP | Merged LSTM | Arousal, Valence, Liking and Dominance | Arousal: 83.85, Valence: 84.89 Liking: 80.72 Dominance: 84.37 |
| 2019 | Wang et al. (2019) | SEED | DNNs | Positive, Negative, and Neutral | Overall Accuracy: 93.28 |
| 2019 | Chen et al. (2019b) | DEAP | Bagging Tree (BT), SVM, LDA, Bayesian LDA, Deep CNN | Valence and Arousal | Valence: 99.97 (using CVCNN), Arousal: 99.58 (using GSLTCNN) |
| 2019 | Ma et al. (2019) | DEAP | multimodal residual LSTM | Arousal and Valence | Valence: 92.30 Arousal: 92.87 |
| 2019 | Pandey and Seeja (2019) | DEAP | DNN | Arousal and Valence | Arousal: 61.25 Valence: 62.50 |
| 2020 | Nakisa et al. (2020) | Audio-video Clips | ConvNet long short-term memory (LSTM; early and late fusion) | Low Arousal Positive, High Arousal Positive, Low Arousal Negative, High Arousal Negative | Overall accuracy: Early fusion:71.61 Late fusion: 70.17 |
| 2020 | Nath et al. (2020) | DEAP | LSTM | Arousal and Valence | Valence: 94.69 Arousal: 93.13 |
| 2020 | Gao et al. (2020) | Film clips | GPSO-optimized CNN | Fear, happiness, and sadness | Overall accuracy: 92.44 ± 3.60 |
| 2020 | Joshi and Ghongade (2020) | Own dataset | BiLSTM | Positive, neutral and negative | Overall Accuracy:72.83 |
| 2020 | Wei et al. (2020) | SEED | SRU | Positive, neutral and negative | Overall Accuracy:80.02 |
| 2020 | Sharma et al. (2020) | DEAP and SEED | LSTM | Arousal and Valence Positive, neutral and negative | DEAP: 4 classes: 82.01 Arousal: 85.21 Valence: 84.16 SEED: 90.81 |
| 2020 | Alhalaseh and Alasasfeh (2020) | DEAP | CNN, k-NN, NB, DT | Valence and Arousal | Overall accuracy: 95.20 (using CNN) |
| 2020 | Cui et al. (2020) | DEAP and DREAMR | Regional- Asymmetric Convolutional Neural Network (RACNN) | Valence and Arousal | Overall accuracy: 96.65 (Valence), 97.11 (Arousal) |

(Continued)

TABLE 2 (Continued)

| Year | References | Stimulus | Classification methods | Emotion | Acc.(%) |
|------|-------------------------|--------------------------------|--|--|---|
| 2020 | Hassouneh et al. (2020) | Own dataset | LSTM | Happy, fear, anger, sad, Surprise and disgust | Overall accuracy: 7.25 |
| 2020 | Liu et al. (2020) | SEED | DECNN | Positive and negative | Overall accuracy: 97.56 |
| 2020 | Li et al. (2020) | AMIGOS | Bidirectional LSTM-RNNs | Valence and Arousal | Arousal: F1-Score: 61.3, ACC: 73.5; Valence: F1-Score: 58.3, ACC: 67.8 |
| 2021 | Topic and Russo (2021) | DEAP,DREAMER, SEED and AMIGOS. | CNN + SVM | Arousal and Valence; Positive and negative | DEAP: Arousal:77.7 andValence: 76.6 DREAMER: Arousal: 90.4 andValence: 88.2 AMIGOS: Arousal: 90.5 andValence: 78.4 SEED: 88.5 |
| 2021 | Liu and Fu (2021) | DEAP | multi-channel feature fusion | SROCC and PLCC | SROCC: 78.9, PLCC: 84.3 |
| 2021 | Sakalle et al. (2021) | Own dataset, DEAP and SEED | LSTM | Disgust, sadness, surprise, and anger Positive, negative, and neutral | DEAP: 91.38 SEED: 89.34 Own dataset: 4 class: 94.12 3 class: 92.66 |
| 2021 | Huang et al. (2021) | DEAP | BiDCNN | Arousal and Valence | Subject-dependent Arousal:94.72 Valence: 94.38 Subject-independent Arousal: 63.94 Valence: 68.14 |
| 2022 | Chowdary et al. (2022) | Own dataset | RNN, LSTM, and GRU | positive, negative, and neutral | average accuracy: RNN: 95, LSTM: 97, GRU:96 |
| 2022 | Algarni et al. (2022) | DEAP | Bi-LSTM | arousal, valence and liking | average accuracy: valence: 99.45, arousal: 96.87, liking: 99.68 |
| 2022 | Tuncer et al. (2022) | DREAMER | LEDPatNet19 | arousal, dominance, and valence | valence: 94.58, arousal: 92.86, arousal: 94.44 |
| 2022 | Li et al. (2022) | DEAP and SEED | ensemble learning | arousal and valence | DEAP average accuracy: Arousal: 65.70, valence: 64.22 SEED average accuracy: 84.44 |
| 2022 | He et al. (2022) | DREAMER and DEAP | adversarial discriminative-temporal convolutional networks (AD-TCNs) | arousal and valence | DEAP average accuracy: Arousal: 64.33, valence: 63.25 DREAMER average accuracy: Arousal: 66.56, valence: 63.69 |
| 2022 | Wang et al. (2022) | DEAP | 2D CNN | arousal and valence | Average accuracy: Arousal: 99.99, valence: 99.98 |

in their EEG signals. Therefore, how to solve the difference between the external-induced emotional change and the internal active emotional change is a subject worthy of study.

Different individuals may not induce the same emotion for the same emotion-inducing video due to the differences in the physiology and psychology between different subjects. Although the same emotion is generated, the EEG signals will have some differences due to the physiological differences between individuals. To effectively solve the problem of individual differences, we can build a personalized emotional computing model from the perspective of individuals. However, building an emotion recognition model with better generalization ability is a relatively more economical solution because

the collection and annotation of physiological signals will bring about a large cost. An effective method to improve the generalization ability of affective computing models is transfer learning (Pan et al., 2011). Therefore, how to combine the agent independent classifier model with the transfer learning technology may be a point worth being considered in the future.

The privacy protection of users' personal information is an important ethical and moral issue in the Internet era. The EEG and other physiological signals collected in emotional computing belong to users' private information, so privacy protection should be paid attention. At present, research in this area has only started (Cock et al., 2017; Agarwal et al., 2019).

5.3. EEG signal preprocessing and feature extraction

In the EEG signal acquisition experiment, many equipment are needed, and the noise acquisition should be minimized. However, EEG signal acquisition is more complex, and the acquisition results are often vulnerable to external factors. Therefore, collecting EEG signals with high efficiency and quality is an important part of affective computing. Effective preprocessing can remove the noise in the original EEG signal, improve the signal quality, and help feature extraction, which is another important part for affective computing.

The common features of EEG signal include power spectral density, differential entropy, asymmetric difference of differential entropy, asymmetric quotient of differential entropy, discrete wavelet analysis, empirical mode decomposition, empirical mode decomposition sample entropy (EMD_SampEn), and statistical features (mean, variance, etc.). How to extract appropriate features or fuse different features will have an important effect on affective computing models.

6. Conclusion

Multiple recent studies using deep learning have been conducted for EEG emotion recognition associated with promising performance due to the strong feature learning and classing ability of deep learning. This paper attempts to provide a comprehensive survey of existing EEG emotion recognition methods. The common open data sets of EEG-based affective computing are introduced. The deep learning techniques are summarized with specific focus on the common methods of emotional calculation of EEG signals, related algorithms. The challenges faced by emotional computing based on EEG signals and the problems to be solved in the future are analyzed and summarized.

References

- Acharya, U. R., Hagiwara, Y., and Adeli, H. (2018a). Automated seizure prediction. *Epilepsy Behav.* 88, 251–261. doi: 10.1016/j.yebeh.2018.09.030
- Acharya, U. R., Adeli, et al. (2018b). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput. Biol. Med.* 100, 270–278. doi: 10.1016/j.compbiomed.2017.09.017
- Alhagry, S., Aly, A., and Reda, A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. *Int. J. Adv. Comput. Sci. Appl.* 8, 355–358. doi: 10.14569/IJACSA.2017.081046
- Alhalaseh, R., and Alasasfeh, S. (2020). Machine-learning-based emotion recognition system using EEG signals. *Computers* 9:95. doi: 10.3390/computers9040095
- Algarni, M., Saeed, F., Al-Hadhrani, T., et al. (2022). Deep learning-based approach for emotion recognition using electroencephalography (EEG) signals using bi-directional long short-term memory (bi-LSTM). *Sensors* 22:2976. doi: 10.3390/s22082976
- Alreshidi, A., and Ullah, M. (2020). Facial emotion recognition using hybrid features. *Informatics* 7:6. doi: 10.3390/informatics7010006
- Alotaiby, T. N., Alshebeili, S. A., Alotaibi, F. M., and Alrshoud, S. R. (2017). Epileptic seizure prediction using CSP and LDA for scalp EEG signals. *Comput. Intell. Neurosci.* 2017, 1–11. doi: 10.1155/2017/1240323
- Antoniades, A., Spyrou, L., Took, C. C., and Sanei, Saeid (2016). “Deep learning for epileptic intracranial EEG data”. in *2016 IEEE 26th international workshop on machine learning for signal processing (MLSP)*. IEEE.
- Agarwal, A., Dowsley, R., McKinney, N. D., Wu, D., Lin, C. T., de Cock, M., et al. (2019). Protecting privacy of users in brain-computer interface applications. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 1546–1555. doi: 10.1109/TNSRE.2019.2926965
- Bashivan, P., Rish, I., Yeasin, M., and Codella, N. (2015). Learning representations from EEG with deep recurrent-convolutional neural networks. *Computer Science*. doi: 10.48550/arXiv.1511.06448
- Bos, D. O. (2006). EEG-based emotion recognition. The influence of visual and auditory stimuli. *Computer Science* 56, 1–17.
- Breiman, L. (1995). Better subset regression using the nonnegative garrote. *Technometrics* 37, 373–384. doi: 10.1080/00401706.1995.10484371
- Cannon, W. B. (1927). The James-Lange theory of emotions: a critical examination and an alternative theory. *Am. J. Psychol.* 39, 106–124. doi: 10.2307/1415404
- Candra, H., Yuwono, M., Chai, R., Handojoseno, Ardi, Elamvazuthi, Irraivan, Nguyen, Hung T., and Su, Steven (2015) “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)*. pp. 7250–7253. IEEE.
- Cai, H., Qu, Z., Li, Z., Zhang, Y., Hu, X., and Hu, B. (2020). Feature-level fusion approaches based on multimodal EEG data for depression recognition. *Inf. Fusion* 59, 127–138. doi: 10.1016/j.inffus.2020.01.008
- Chakladar, D. D., and Chakraborty, S. (2018). EEG based emotion classification using “correlation based subset selection”. *Biol. Inspired Cogn. Archit.* 24, 98–106. doi: 10.1016/j.bica.2018.04.012
- Chang, C. Y., Zheng, J. Y., and Wang, C. J. (2010). “Based on support vector regression for emotion recognition using physiological signals”. in *The 2010 international joint conference on neural networks (IJCNN)*. pp. 1–7. IEEE.
- Chen, X., Xu, L., Cao, M., Zhang, T., Shang, Z., and Zhang, L. (2021). Design and implementation of human-computer interaction systems based on transfer support vector machine and EEG signal for depression patients’ emotion recognition. *J. Med. Imaging Health Infor.* 11, 948–954. doi: 10.1166/jmhi.2021.3340
- Chen, J., Hu, B., Moore, P., Zhang, X., and Ma, X. (2015). Electroencephalogram-based emotion assessment system using ontology and data mining techniques. *Appl. Soft Comput.* 30, 663–674. doi: 10.1016/j.asoc.2015.01.007
- Chen, X., Zeng, W., Shi, Y., Deng, J., and Ma, Y. (2019a). Intrinsic prior knowledge driven CICA FMRI data analysis for emotion recognition classification. *IEEE Access* 7, 59944–59950. doi: 10.1109/ACCESS.2019.2915291
- Chen, J. X., Zhang, P. W., Mao, Z. J., Huang, Y. F., Jiang, D. M., and Zhang, Y. N. (2019b). Accurate EEG-based emotion recognition on combined features using deep convolutional

Author contributions

XW and YR contributed to the writing and drafted this article. ZL and JH contributed to the collection and analysis of existing literature. YH contributed to the conception and design of this work and revised this article. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Hunan Provincial Natural Science Foundation of China (under Grant Nos. 2022JJ50148 and 2021JJ50082), Hunan Provincial Education Department Science Research Fund of China (under Grant Nos. 22A0625 and 21B0800), and National Innovation and Entrepreneurship Training for University of PRC (under Grant Nos. 202211528036 and 202211528023).

Conflict of interest

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

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- neural networks. *IEEE Access* 7, 44317–44328. doi: 10.1109/ACCESS.2019.2908285
- Chen, B., Liu, Y., Zhang, Z., Li, Y., Zhang, Z., Lu, G., et al. (2021). Deep active context estimation for automated COVID-19 diagnosis. *ACM Trans. Multimed. Comput. Appl.* 17, 1–22. doi: 10.1145/3457124
- Chao, H., Dong, L., Liu, Y., and Lu, B. (2019). Emotion recognition from multiband eeg signals using capsnet. *Sensors* 19:2212. doi: 10.3390/s19092212
- Cho, K., Merriënboer, B. V., Gulcehre, C., Bahdanau, D., Schwenk, F., Holger, et al. (2014). “Learning phrase representations using RNN encoder-decoder for statistical machine translation”. in *Conference on empirical methods in natural language processing (EMNLP 2014)*.
- Chowdhury, M. K., Anitha, J., and Hemanth, D. J. (2022). Emotion recognition from EEG signals using recurrent neural networks. *Electronics* 11:2387. doi: 10.3390/electronics11152387
- Cui, H., Liu, A., Zhang, X., Chen, X., Wang, K., and Chen, X. (2020). EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network. *Knowl.-Based Syst.* 205:106243. doi: 10.1016/j.knsys.2020.106243
- Cock, M., Dowsley, R., McKinney, N., Nascimento, A. C. A., and Wu, D. D. (2017). “Privacy preserving machine learning with EEG data”. In: *Proceedings of the 34th international conference on machine learning*. Sydney, Australia.
- Cortes, C., and Vapnik, V. (1995). Support vector machine. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018
- Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). “Imagenet: a large-scale hierarchical image database”. in *2009 IEEE conference on computer vision and pattern recognition (pp. 248–255)*. IEEE.
- Donos, C., Maliia, M. D., Dümpelmann, M., and Schulze-Bonhage, A. (2018). Seizure onset predicts its type. *Epilepsia* 59, 650–660. doi: 10.1111/epi.13997
- Donos, C., Dümpelmann, M., and Schulze-Bonhage, A. (2015). Early seizure detection algorithm based on intracranial EEG and random forest classification. *Int. J. Neural Syst.* 25:1550023. doi: 10.1142/S0129065715500239
- Elman, J. L. (1990). Finding structure in time. *Cogn. Sci.* 14, 179–211. doi: 10.1016/0364-0213(90)90002-E
- Ekman, P., and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* 17, 124–129. doi: 10.1037/h0030377
- Fan, J., Wade, J. W., Key, A. P., Warren, Z. E., and Sarkar, N. (2017). EEG-based affect and workload recognition in a virtual driving environment for ASD intervention. *IEEE Trans. Biomed. Eng.* 65, 43–51. doi: 10.1109/TBME.2017.2693157
- Garg, A., Kapoor, A., Bedi, A. K., and Sunkaria, Ramesh K. (2019). International Conference on Data Science and Engineering (ICDSE) (2019). “Merged LSTM model for emotion classification using EEG signals”. in *2019 international conference on data science and engineering (ICDSE)*. pp. 139–143. IEEE.
- Gao, Z., Li, Y., Yang, Y., Wang, X., Dong, N., and Chiang, H. D. (2020). A GPSS-optimized convolutional neural networks for EEG-based emotion recognition. *Neurocomputing* 380, 225–235. doi: 10.1016/j.neucom.2019.10.096
- Guo, J., Fang, F., Wang, W., and Ren, F. (2018). “EEG emotion recognition based on granger causality and capsnet neural network”. in *2018 5th IEEE international conference on cloud computing and intelligence systems (CCIS)*. pp. 47–52. IEEE.
- Güntekin, B., Hanoğlu, L., Aktürk, T., Fide, E., Emek-Savaş, D. D., Ruşen, E., et al. (2019). Impairment in recognition of emotional facial expressions in Alzheimer’s disease is represented by EEG theta and alpha responses. *Psychophysiology* 56:e13434. doi: 10.1111/psyp.13434
- Gonzalez, H. A., Yoo, J., and Elfadel, I. M. (2019). “EEG-based emotion detection using unsupervised transfer learning”. in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. pp. 694–697. IEEE.
- Graves, A., and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* 18, 602–610. doi: 10.1016/j.neunet.2005.06.042
- Hajinoroozi, M., Zhang, J. M., and Huang, Y. (2017). “Driver’s fatigue prediction by deep covariance learning from EEG”. in *2017 IEEE international conference on systems, man, and cybernetics (SMC)*. pp. 240–245. IEEE.
- Hassouneh, A., Mutawa, A. M., and Murugappan, M. (2020). Development of a real-time emotion recognition system using facial expressions and EEG based on machine learning and deep neural network methods. *Inform. Med. Unlocked* 20:100372. doi: 10.1016/j.imu.2020.100372
- He, Z., Zhong, Y., and Pan, J. (2022). An adversarial discriminative temporal convolutional network for EEG-based cross-domain emotion recognition. *Comput. Biol. Med.* 141:105048. doi: 10.1016/j.compbiomed.2021.105048
- Hinton, G. E., Osindero, S., and Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527–1554. doi: 10.1162/NECO.2006.18.7.1527
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Huang, D., Chen, S., Liu, C., Zheng, L., Tian, Z., and Jiang, D. (2021). Differences first in asymmetric brain: a bi-hemisphere discrepancy convolutional neural network for EEG emotion recognition. *Neurocomputing* 448, 140–151. doi: 10.1016/j.neucom.2021.03.105
- Ji, S., Niu, X., Sun, M., Shen, T., Xie, S., Zhang, H., and Liu, H. (2022). “Emotion recognition of autistic children based on EEG signals”. in *International conference on computer engineering and networks*. pp. 698–706. Springer, Singapore.
- Jiao, Z., Gao, X., Wang, Y., Li, J., and Xu, H. (2018). Deep convolutional neural networks for mental load classification based on EEG data. *Pattern Recogn.* 76, 582–595. doi: 10.1016/j.patcog.2017.12.002
- Jie, X., Cao, R., and Li, L. (2014). Emotion recognition based on the sample entropy of EEG. *Biomed. Mater. Eng.* 24, 1185–1192. doi: 10.3233/BME-130919
- Joshi, V. M., and Ghongade, R. B. (2020). IDEA: intellect database for emotion analysis using EEG signal. *J. King Saud Univ. - Comput. Inf. Sci.* 34, 4433–4447. doi: 10.1016/j.jksuci.2020.10.007
- Katsigiannis, S., and Ramzan, N. (2018). DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J. Biomed. Health Inform.* 22, 98–107. doi: 10.1109/JBHI.2017.2688239
- Kaundanya, V. L., Patil, A., and Panat, A. (2015). Performance of k-NN classifier for emotion detection using EEG signals. in *2015 international conference on communications and signal processing (ICCSP)*. pp. 1160–1164. IEEE.
- Kamada, S., Ichimura, T. (2019). A Video Recognition Method by using Adaptive Structural Learning of Long Short Term Memory based Deep Belief Network. In *2019 IEEE 11th International Workshop on Computational Intelligence and Applications (IWCIA)*. IEEE, 2019, 21–26. doi:10.1109/IWCIA47330.2019.8955036.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Proc. Adv. Neural Inf. Process. Syst.* 60, 84–90. doi: 10.1145/3065386
- Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., and Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: a review. *IEEE Access* 7, 117327–117345. doi: 10.1109/ACCESS.2019.2936124
- Koelstra, S., Muhl, C., Soleymani, M., Jong-Seok Lee, , Yazdani, A., Ebrahimi, T., et al. (2012). DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Kuncheva, L. I., Christy, T., Pierce, I., and Mansoor, S.P. (2011) Multi-modal biometric emotion recognition using classifier ensembles. in *International conference on industrial, engineering and other applications of applied intelligent systems*. pp. 317–326. Springer, Berlin, Heidelberg.
- Lazarus, R. S. (1993). From psychological stress to the emotions: a history of changing outlooks. *Annu. Rev. Psychol.* 44, 1–22. doi: 10.1146/annurev.ps.44.020193.000245
- Lee, J. K., Kang, H. W., and Kang, H. B. (2015). Smartphone addiction detection based emotion detection result using random Forest. *J. IKEEE* 19, 237–243. doi: 10.7471/ikee.2015.19.2.237
- Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R. E., Wahlund, L. O., et al. (2007). Application and comparison of classification algorithms for recognition of Alzheimer’s disease in electrical brain activity (EEG). *J. Neurosci. Methods* 161, 342–350. doi: 10.1016/j.jneumeth.2006.10.023
- Lecun, Y., and Bottou, L. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791
- Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Li, C., Bao, Z., Li, L., and Zhao, Z. (2020). Exploring temporal representations by leveraging attention-based bidirectional LSTM-RNNs for multi-modal emotion recognition. *Inf. Process. Manag.* 57:102185. doi: 10.1016/j.ipm.2019.102185
- Li, X., Song, D., Zhang, P., Yu, Guangliang, Hou, Yuxian, and Hu, Bin (2017). Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. in *2016 IEEE international conference on bioinformatics and biomedicine (BIBM)*. pp. 352–359. IEEE.
- Li, Y., Zheng, W., Wang, L., Zong, Y., and Cui, Z. (2019). From regional to global brain: a novel hierarchical spatial-temporal neural network model for EEG emotion recognition. *IEEE Trans. Affect. Comput.* 13, 568–578. doi: 10.1109/TAFFC.2019.2922912
- Li, R., Ren, C., Zhang, X., and Hu, B. (2022). A novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition. *Comput. Biol. Med.* 140:105080. doi: 10.1016/j.compbiomed.2021.105080
- Liu, S., Wang, X., Zhao, L., Zhao, J., Xin, Q., and Wang, S. H. (2020). Subject-independent emotion recognition of EEG signals based on dynamic empirical convolutional neural network. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 18, 1710–1721. doi: 10.1109/TNSRE.2020.3019063
- Liu, Y., and Fu, G. (2021). Emotion recognition by deeply learned multi-channel textual and EEG features. *Futur. Gener. Comput. Syst.* 119, 1–6. doi: 10.1016/j.future.2021.01.010
- Liu, S., Wang, X., Zhao, Z., Li, B., Hu, W., Yu, J., et al. (2021). 3DCANN: A Spatio-Temporal Convolution Attention Neural Network for EEG Emotion Recognition. *IEEE Journal of Biomedical and Health Informatics*. doi: 10.1109/JBHI.2021.3083525
- Liu, Y., Sourina, O., and Nguyen, M. K. (2011). *Real-time EEG-based emotion recognition and its applications*. Springer Berlin Heidelberg.
- Liu, J., Su, Y., and Liu, Y. (2017). *Multi-modal emotion recognition with temporal-band attention based on LSTM-RNN*. Springer, Cham.
- Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, Shyh-Kang, and Chen, Jyh-Horng (2009). “EEG-based emotion recognition in music listening: a comparison of schemes for multiclass support

- vector machine." in *2009 IEEE international conference on acoustics, speech and signal processing*. pp. 489–492. IEEE.
- Lindley, D. V., and Smith, A. F. M. (1972). Bayes estimates for the linear model. *J. R. Stat. Soc. Ser. B Methodol.* 34, 1–18. doi: 10.1111/j.2517-6161.1972.tb00885.x
- Mayor-Torres, J. M., Medina - De Villiers, S., Clarkson, T., Lerner, M. D., and Riccardi, G. (2021). Evaluation of interpretability for deep learning algorithms in EEG emotion recognition: a case study in autism. *arXiv* 2111.13208.
- Maimaiti, B., Meng, H., Lv, Y., Qiu, J., Zhu, Z., Xie, Y., et al. (2021). An overview of EEG-based machine learning methods in seizure prediction and opportunities for neurologists in this field. *Neuroscience* 481, 197–218. doi: 10.1016/j.neuroscience.2021.11.017
- Martis, R. J., Acharya, U. R., Tan, J. H., Petznick, A., Yanti, R., Chua, C. K., et al. (2012). Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals. *Int. J. Neural Syst.* 22:1250027. doi: 10.1142/S012906571250027X
- Ma, J., Tang, H., Zheng, W. L., and Lu, Bao-Liang (2019). "Emotion recognition using multimodal residual LSTM network" in *Proceedings of the 27th ACM international conference on multimedia*. pp. 176–183.
- Manor, R., and Geva, A. B. (2015). Convolutional neural network for multi-category rapid serial visual presentation BCI. *Front. Comput. Neurosci.* 9:146. doi: 10.3389/fncom.2015.00146
- Mcculloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5, 115–133. doi: 10.2307/2268029
- Mehdizadehfard, V., Ghassemi, F., Fallah, A., and Pouretmad, H. (2020). EEG study of facial emotion recognition in the fathers of autistic children. *Biomed. Signal Process. Control* 56:101721. doi: 10.1016/j.bspc.2019.101721
- Menezes, M. L. R., Samara, A., Galway, L., Sant'Anna, A., Verikas, A., Alonso-Fernandez, F., et al. (2017). Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset. *Pers. Ubiquit. Comput.* 21, 1003–1013. doi: 10.1007/s00779-017-1072-7
- Mehrabian, A. (1996). Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Curr. Psychol.* 14, 261–292. doi: 10.1007/BF02686918
- Miranda, J. A., Canabal, M. F., Portela García, M., and Lopez-Ongil, Celia (2018). "Embedded emotion recognition: autonomous multimodal affective internet of things". in *Proceedings of the cyber-physical systems workshop*. Vol. 2208, pp. 22–29.
- Murugappan, M. (2011). "Human emotion classification using wavelet transform and KNN". in *2011 international conference on pattern analysis and intelligence robotics*. Vol. 1, pp. 148–153. IEEE.
- Murugappan, M., Ramachandran, N., and Sazali, Y. (2010). Classification of human emotion from EEG using discrete wavelet transform. *J. Biomed. Sci. Eng.* 03, 390–396. doi: 10.4236/jbise.2010.34054
- Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F., and Chandran, V. (2020). Automatic emotion recognition using temporal multimodal deep learning. *IEEE Access* 8, 225463–225461, 225474. doi: 10.1109/ACCESS.2020.3027026
- Nath, D., Singh, M., Sethia, D., Kalra, Diksha, and Indu, S. (2020). "A comparative study of subject-dependent and subject-independent strategies for EEG-based emotion recognition using LSTM network" in *Proceedings of the 2020 the 4th international conference on computer and data analysis*. pp. 142–147.
- Nie, D., Wang, X. W., Shi, L. C., and Lu, Bao-Liang (2011). "EEG-based emotion recognition during watching movies". in *2011 5th international IEEE/EMBS conference on neural engineering*. pp. 667–670. IEEE.
- Otter, D. W., Medina, J. R., and Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 604–624. doi: 10.1109/TNNLS.2020.2979670
- Payton, L., Szu, Wei F., Syu, Siang W. (2016). Maximum entropy learning with deep belief networks. *Entropy*. 18:251–256. doi: 10.3390/e18070251.
- Page, A., Shea, C., and Mohsenin, T. (2016). Wearable seizure detection using convolutional neural networks with transfer learning. *ISCAS*. IEEE.
- Pandey, P., and Seeja, K. R. (2019). Subject independent emotion recognition from EEG using VMD and deep learning. *J. King Saud Univ. - Comput. Inf. Sci.* 34, 1730–1738. doi: 10.1016/j.jksuci.2019.11.003
- Pan, S. J., Tsang, I. W., Kwok, J. T., and Yang, Q. (2011). Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* 22, 199–210. doi: 10.1109/TNN.2010.2091281
- Piana, S., Staglianò, A., Odone, F., and Camurri, A. (2016). Adaptive body gesture representation for automatic emotion recognition. *ACM Trans. Interact. Intell. Syst.* 6, 1–31. doi: 10.1145/2818740
- Plutchik, R. (2003). *Emotions and life: Perspectives from psychology, biology, and evolution*. Washington, DC: American Psychological Association.
- Purwins, H., Li, B., Virtanen, T., Schluter, J., Chang, S. Y., and Sainath, T. (2019). Deep learning for audio signal processing. *IEEE J. Sel. Top. Signal Process.* 13, 206–219. doi: 10.1109/JSTSP.2019.2908700
- Picard, R. W., and Picard, R. (1997). *Affective computing* (Vol. 252). MIT press Cambridge. "EEG-detected olfactory imagery to reveal covert consciousness in minimally conscious state". *Brain Inj.* 29, 1729–1735.
- Quinlan, J. R. (1996). Learning decision tree classifiers. *ACM Comput. Surv.* 28, 71–72. doi: 10.1145/234313.234346
- Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., and Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: direct evidence for a relationship between the autonomic nervous system and social cognition. *Int. J. Psychophysiol.* 86, 168–172. doi: 10.1016/j.ijpsycho.2012.08.012
- Rasheed, K., Qayyum, A., Qadir, J., Sivathamboo, S., Kwan, P., Kuhlmann, L., et al. (2020). Machine learning for predicting epileptic seizures using EEG signals: a review. *IEEE Rev. Biomed. Eng.* 14, 139–155. doi: 10.1109/RBME.2020.3008792
- Ren, Y., and Wu, Y. (2014). Convolutional deep belief networks for feature extraction of EEG signal. *International joint conference on neural networks*. pp. 2850–2853. IEEE.
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39:1161.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by backpropagating errors. *Nature* 323, 533–536. doi: 10.1038/323533a0
- Sakhavi, S., Guan, C., and Yan, S. (2015). Parallel convolutional-linear neural network for motor imagery classification. in *2015 23rd European signal processing conference (EUSIPCO)*. pp. 2736–2740. IEEE.
- Salama, E. S., A.el-Khoribi, R., E.Shoman, M., and A.Wahby, M. (2018). EEG-based emotion recognition using 3D convolutional neural networks. *Int. J. Adv. Comput. Sci. Appl.* 9, 329–337. doi: 10.14569/IJACSA.2018.090843
- Sakalle, A., Tomar, P., Bhardwaj, H., Acharya, D., and Bhardwaj, A. (2021). A LSTM based deep learning network for recognizing emotions using wireless brainwave driven system. *Expert Syst. Appl.* 173:114516. doi: 10.1016/j.eswa.2020.114516
- Schuster, M., and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* 45, 2673–2681. doi: 10.1109/78.650093
- Seo, J., Laine, T. H., Oh, G., and Sohn, K. A. (2020). EEG-based emotion classification for Alzheimer's disease patients using conventional machine learning and recurrent neural network models. *Sensors* 20:7212. doi: 10.3390/s20247212
- Shamwell, J., Lee, H., Kwon, H., Marathe, A., Lawhern, V., and Nothwang, W. (2016). "Single-trial EEG RSVP classification using convolutional neural networks" in *Micro-and nanotechnology sensors, systems, and applications VIII*, vol. 9836 eds. T. George, A. K. Dutta, and M. S. Islam (Proc. of SPIE), 373–382.
- Sharma, R., Pachori, R. B., and Sircar, P. (2020). Automated emotion recognition based on higher order statistics and deep learning algorithm. *Biomed. Signal Process. Control* 58:101867. doi: 10.1016/j.bspc.2020.101867
- Soleymani, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* 3, 42–55. doi: 10.1109/T-AFFC.2011.25
- Soleymani, M., Asghari-Esfeden, S., Pantic, M., and Fu, Yun (2014). "Continuous emotion detection using EEG signals and facial expressions". in *2014 IEEE international conference on multimedia and expo (ICME)*. pp. 1–6. IEEE.
- Song, T., Zheng, W., Song, P., and Cui, Z. (2018). EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans. Affect. Comput.* 11, 532–541. doi: 10.1109/T-AFFC.2018.2817622
- Song, T., Zheng, W., Lu, C., Zong, Y., Zhang, X., and Cui, Z. (2019). MPED: a multimodal physiological emotion database for discrete emotion recognition. *IEEE Access* 7, 12177–12191. doi: 10.1109/ACCESS.2019.2891579
- Suhaimi, N. S., Mountstephens, J., and Teo, J. (2020). EEG-based emotion recognition: a state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* 2020, 1–19. doi: 10.1155/2020/8875426
- Stavroulia, K. E., Christofi, M., Baka, E., Michael-Grigoriou, D., Magnenat-Thalmann, N., and Lanitis, A. (2019). Assessing the emotional impact of virtual reality-based teacher training. *Int. J. Inf. Learn. Technol.* 36, 192–217. doi: 10.1108/IJILT-11-2018-0127
- Thodoroff, P., Pineau, J., and Lim, A. (2016). "Learning robust features using deep learning for automatic seizure detection". in *Machine learning for healthcare conference*. pp. 178–190. PMLR.
- Tuncer, T., Dogan, S., and Subasi, A. (2022). LEDPatNet19: automated emotion recognition model based on nonlinear LED pattern feature extraction function using EEG signals. *Cogn. Neurodyn.* 16, 779–790. doi: 10.1007/s11571-021-09748-0
- Topic, A., and Russo, M. (2021). Emotion recognition based on EEG feature maps through deep learning network. *Eng. Sci. Technol. Int. J.* 24, 1442–1454. doi: 10.1016/j.jestch.2021.03.012
- Ur, A., Yanti, R., Swapna, G., Sree, V. S., Martis, R. J., and Suri, J. S. (2013). Automated diagnosis of epileptic electroencephalogram using independent component analysis and discrete wavelet transform for different electroencephalogram durations. *Proc. Inst. Mech. Eng. H J. Eng. Med.* 227, 234–244. doi: 10.1177/0954411912467883
- Valderas, M. T., Bolea, J., Laguna, P., Fellow IEEE Bailón, R., and Vallverdú, M. (2019). Mutual information between heart rate variability and respiration for emotion characterization. *Physiol. Meas.* 40:084001. doi: 10.1088/1361-6579/ab310a
- Wang, J., Ju, R., Chen, Y., Zhang, L., Hu, J., Wu, Y., et al. (2018). Automated retinopathy of prematurity screening using deep neural networks. *EBioMedicine* 35, 361–368. doi: 10.1016/j.ebiom.2018.08.033
- Wang, Y., Zhang, L., Xia, P., Wang, P., Chen, X., du, L., et al. (2022). EEG-based emotion recognition using a 2D CNN with different kernels. *Bioengineering* 9:231. doi: 10.3390/bioengineering9060231
- Wang, Y., Qiu, S., Zhao, C., Yang, W., Li, J., Ma, X., and He, H. (2019). EEG-based emotion recognition with prototype-based data representation. in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. pp. 684–689. IEEE.

- Wei, C., Chen, L., Song, Z., Lou, X. G., and Li, D. D. (2020). EEG-based emotion recognition using simple recurrent units network and ensemble learning. *Biomed. Signal Process. Control* 58:101756. doi: 10.1016/j.bspc.2019.101756
- Wen, G., Li, H., Huang, J., Li, D., and Xun, E. (2017). Random deep belief networks for recognizing emotions from speech signals. *Comput. Intell. Neurosci.* 2017, 1–9. doi: 10.1155/2017/1945630
- Wu, C., Zhang, Y., Jia, J., and Zhu, W. (2017). Mobile contextual recommender system for online social media. *IEEE Trans. Mob. Comput.* 16, 3403–3416. doi: 10.1109/TMC.2017.2694830
- Wu, X., Sahoo, D., and Hoi, S. (2020). Recent advances in deep learning for object detection. *Neurocomputing* 396, 39–64. doi: 10.1016/j.neucom.2020.01.085
- Williams, R. J., and Zipser, D. (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.* 1, 270–280. doi: 10.1162/neco.1989.1.2.270
- Yi, Z. (2013). *Convergence analysis of recurrent neural networks*, vol. 13 Kluwer Academic Publishers.
- Yi, Z. (2010). Foundations of implementing the competitive layer model by lotka–volterra recurrent neural networks. *IEEE Trans. Neural Netw.* 21, 494–507. doi: 10.1109/TNN.2009.2039758
- Yuan, S., Zhou, W., and Chen, L. (2018). Epileptic seizure prediction using diffusion distance and bayesian linear discriminate analysis on intracranial EEG. *Int. J. Neural Syst.* 28:1750043. doi: 10.1142/S0129065717500435
- Zhang, S., Zhang, S., Huang, T., and Gao, W. (2017). Speech emotion recognition using deep convolutional neural network and discriminant temporal pyramid matching. *IEEE Trans. Multimed.* 20, 1576–1590. doi: 10.1109/TMM.2017.2766843
- Zhang, S., Zhang, S., Huang, T., Gao, W., and Tian, Q. (2018). Learning affective features with a hybrid deep model for audio–visual emotion recognition. *IEEE Trans. Circuits Syst. Video Technol.* 28, 3030–3043. doi: 10.1109/TCSVT.2017.2719043
- Zhang, S., Zhao, X., and Tian, Q. (2019). Spontaneous speech emotion recognition using multiscale deep convolutional LSTM. *IEEE Trans. Affect. Comput.* 13, 680–688. doi: 10.1109/TAFFC.2019.2947464
- Zhang, S., Tao, X., Chuang, Y., and Zhao, X. (2021). Learning deep multimodal affective features for spontaneous speech emotion recognition. *Speech Comm.* 127, 73–81. doi: 10.1016/j.specom.2020.12.009
- Zhang, T., Zheng, W., Cui, Z., Zong, Y., and Li, Y. (2018). Spatial–temporal recurrent neural network for emotion recognition. *IEEE Trans. Cybern.* 49, 839–847. doi: 10.1109/TCYB.2017.2788081
- Zhang, Y., Cheng, C., and Zhang, Y. (2021). Multimodal emotion recognition using a hierarchical fusion convolutional neural network. *IEEE Access* 9, 7943–7951. doi: 10.1109/ACCESS.2021.3049516
- Zheng, W. L., Liu, W., Lu, Y., Lu, B. L., and Cichocki, A. (2018). Emotionmeter: a multimodal framework for recognizing human emotions. *IEEE Trans. Cybern.* 49, 1110–1122. doi: 10.1109/TCYB.2018.2797176
- Zheng, W. L., and Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Ment. Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497
- Zhou, H., du, J., Zhang, Y., Wang, Q., Liu, Q. F., and Lee, C. H. (2021). Information fusion in attention networks using adaptive and multi-level factorized bilinear pooling for audio-visual emotion recognition. *IEEE/ACM Trans. Audio Speech Lang. Process.* 29, 2617–2629. doi: 10.1109/TASLP.2021.3096037