Multi-method Fusion of Cross-Subject Emotion Recognition Based on High-Dimensional EEG Features

Fu Yang [1,2], Xingcong Zhao [1,2], Wenge Jiang [1,2], Pengfei Gao [1,2] and Guangyuan Liu [1,2*]

[1] College of Electronic Information and Engineering, Southwest University, Chongqing, China, [2] Chongqing Key Laboratory of Nonlinear Circuit and Intelligent Information Processing, Chongqing, China

Emotion recognition using electroencephalogram (EEG) signals has attracted significant research attention. However, it is difficult to improve the emotional recognition effect across subjects. In response to this difficulty, in this study, multiple features were extracted for the formation of high-dimensional features. Based on the high-dimensional features, an effective method for cross-subject emotion recognition was then developed, which integrated the significance test/sequential backward selection and the support vector machine (ST-SBSSVM). The effectiveness of the ST-SBSSVM was validated on a dataset for emotion analysis using physiological signals (DEAP) and the SJTU Emotion EEG Dataset (SEED). With respect to high-dimensional features, the ST-SBSSVM average improved the accuracy of cross-subject emotion recognition by 12.4% on the DEAP and 26.5% on the SEED when compared with common emotion recognition methods. The recognition accuracy obtained using ST-SBSSVM was as high as that obtained using sequential backward selection (SBS) on the DEAP dataset. However, on the SEED dataset, the recognition accuracy increased by ~6% using ST-SBSSVM from that using the SBS. Using the ST-SBSSVM, ~97% (DEAP) and 91% (SEED) of the program runtime was eliminated when compared with the SBS. Compared with recent similar works, the method developed in this study for emotion recognition across all subjects was found to be effective, and its accuracy was 72% (DEAP) and 89% (SEED).

Keywords: EEG, emotion recognition, cross-subject, multi-method fusion, high-dimensional features

1. INTRODUCTION

Emotion is essential to humans, as it contributes to the communication between people and plays a significant role in rational and intelligent behavior (Picard et al., 2001; Nie et al., 2011), which is critical to several aspects of daily life. Therefore, research on emotion recognition is necessary. It is difficult to define and classify emotion due to the complex nature and genesis of emotion (Ashforth and Humphrey, 1995; Horlings et al., 2008; Hwang et al., 2018). To classify and represent emotion, several models have been proposed. Moreover, there are two main models. The first assumes that all emotions can comprise primary emotions, similar to how all colors can comprise primary colors. Plutchik (1962) related eight basic emotions to evolutionarily valuable properties, and then reported the following primary emotions: anger, fear, sadness, disgust, surprise, curiosity, acceptance, and joy. Ekman (Power and Dalgleish, 1999; Horlings et al., 2008) reported other

emotions as a basis set and found that these primary emotions, in addition to their expressions, are universal. The Ekman list of primary emotions is as follows: anger, fear, sadness, happiness, disgust, and surprise. The second main model is composed of multiple dimensions, and each emotion is on a multi-dimensional scale. Russell (1980) divided human emotions into two dimensions: arousal and valence. Arousal represents the strength of the emotion with respect to arousal and relaxation and valence represents positive and negative levels. Among several emotional models, the Russell model (Russell, 1980) is generally adopted, in which two dimensions are represented by a vertical arousal axis and horizontal valence axis (Choi and Kim, 2018). In both dimensions of emotion, the ability to measure valence levels is essential, as the valence level is a more critical dimension for distinguishing between positive emotions (e.g., excitement, happiness, contentment, or satisfaction) and negative emotions (e.g., fear, anger, frustration, mental stress, or depression; Hwang et al., 2018). It is necessary to effectively classify and identify positive and negative emotions. For example, the accurate identification of the mental stress (a negative emotion) or emotional state of construction workers can help reduce construction hazards and improve production efficiency (Chen and Lin, 2016; Jebelli et al., 2018a). The focus of this study was on improving the classification accuracy of positive and negative emotions. In daily human life, communication and decision-making are influenced by emotional behavior. For many years, the brain-computer interface (BCI) has been a critical topic with respect to biomedical engineering research, allowing for the use of brain waves to control equipment (Nijboer et al., 2009). To achieve accurate and smooth interactions, computers and robots should be able to analyze emotions (Pessoa and Adolphs, 2010; Zheng et al., 2019). Researchers in the fields of psychology, biology, and neuroscience have directed significant attention toward emotional research. Emotional research has a preliminary development trend in the field of computer science, such as task workload assessments and operator vigilance (Shi and Lu, 2013; Zheng and Lu, 2017b). The automatic emotion recognition system simplifies the computer interface and renders it more convenient, more efficient, and more user-friendly. Human emotion recognition can be studied using questionnaires (Mucci et al., 2015; Jebelli et al., 2019), facial images, gestures, speech signals and other physiological signals (Jerritta et al., 2011). However, the questionnaire method interfered with this study. In addition, it exhibited a significant deviation and yielded inconsistent results (Jebelli et al., 2019). There was an ambiguity with respect to emotion recognition from facial images, gestures, or speech signals, as real emotions can be mimicked. To overcome the ambiguity, an electroencephalogram (EEG) could be employed for emotion recognition, as it is more accurate and more objective than emotional evaluation based on facial image and gesture-based methods (Ahern and Schwartz, 1985). Therefore, EEG has attracted significant research attention. Moreover, EEG signals can be used to effectively identify different emotions (Sammler et al., 2007; Mathersul et al., 2008; Knyazev et al., 2010; Bajaj and Pachori, 2014). For effective medical care, it is essential to consider emotional states (Doukas and Maglogiannis, 2008; Petrantonakis and Hadjileontiadis, 2011).

Due to the objectivity of physiological data and the ability to model learning principles from heterogeneous features to emotional classifiers, the use of machine learning methods for the analysis of EEG signals has attracted significant attention in the field of human emotion recognition. To improve the satisfaction and reliability of the people who interact and collaborate with machines and robots, a smart human-machine (HM) system that can accurately interpret human communication capabilities is required (Koelstra et al., 2011). Human intentions and commands mostly convey emotions in a linguistic or non-verbal manner; thus, the accurate response to human emotional behavior is critical to the realization of machine and computer adaptation (Zeng et al., 2008; Fanelli et al., 2010). At this stage, the majority of HM systems cannot accurately recognize emotional cues. Emotional classifiers were developed based on facial/sound expressions or physiological signals (Hanjalic and Xu, 2005; Kim and André, 2008). Emotion classifiers can provide temporal predictions of specific emotional states. Emotional recognition requires appropriate signal preprocessing techniques, feature extraction, and machine learning-based classifiers to carry out automatic classification. An EEG, which captures brain waves, can effectively distinguish between emotions. The EEG directly detects brain waves from the central nervous system activities (i.e., brain activities), whereas other responses (e.g., EDA, HR, and BVP) are based on peripheral nervous system activities (Zhai et al., 2005; Chanel et al., 2011; Hwang et al., 2018). In particular, central nervous system activities are related to several aspects of emotions (e.g., from displeasure to pleasure, and from relaxation to excitement); however, the peripheral nervous system activities are only associated with arousal and relaxation (Zhai et al., 2005; Chanel et al., 2011). Therefore, the EEG can provide more detailed information on emotional states than other methods (Takahashi et al., 2004; Lee and Hsieh, 2014; Liu and Sourina, 2014; Hou et al., 2015). Moreover, EEG-based emotion recognition has a greater potential with respect to research than facial and speech-based methods, given that internal nerve fluctuations cannot be deliberately masked or controlled. However, the improvement of the performance of cross-subject emotional recognition has been the focus of several studies, including this study. In previous studies, cross-subject emotion recognition was difficult to achieve when compared with intra-subject emotion recognition. In Kim (2007), the method of bimodal data fusion was investigated, and a linear discriminant analysis (LDA) was conducted to classify emotions. The best recognition accuracy across the three subjects was 55%, which was significantly lower than the 92% achieved using the intra-subject emotion recognition method (Kim, 2007). In Zhu et al. (2015), the authors employed differential entropy (DE) as features, and a linear dynamic system (LDS) was applied to carry out feature smoothing. The average cross-subject classification accuracy was 64.82%, which was significantly lower than the 90.97% of the intra-subject emotion recognition method (Zhu et al., 2015). In Zhuang et al. (2017), a method for feature extraction and emotion recognition based on empirical mode decomposition (EMD) was introduced. Using EMD, the EEG signals were automatically decomposed into intrinsic mode functions (IMFs). Based on the results, IMF1 demonstrated

FIGURE 1 | Analysis process.
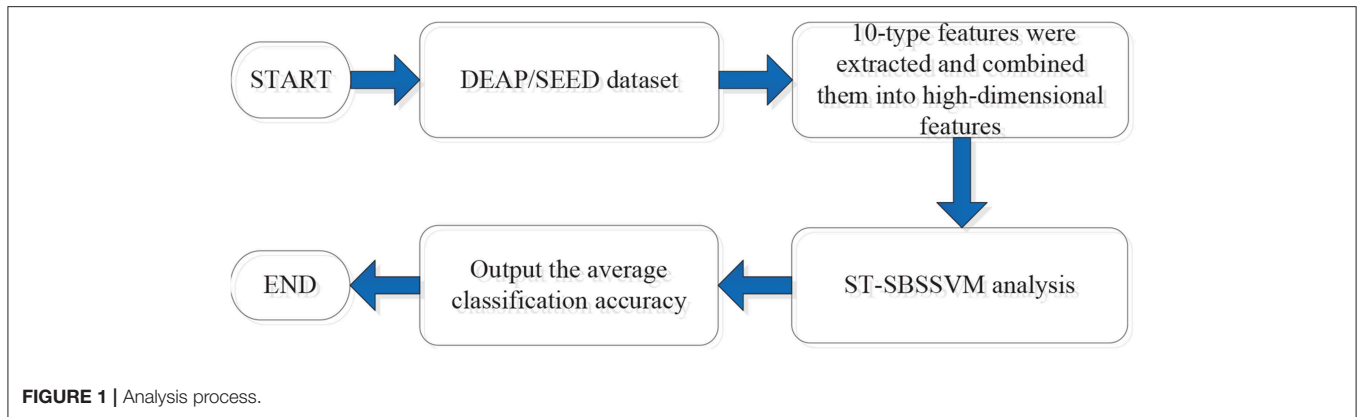
the best performance, which was 70.41% for valence (Zhuang et al., 2017). In Candra et al. (2015), an accuracy of 65% was achieved for valence and arousal using the wavelet entropy of signal segments with periods of 3–12 s. In Mert and Akan (2018), the advanced properties of EMD and its multivariate extension (MEMD) for emotion recognition were investigated. The multichannel IMFs extracted by MEMD were analyzed using various time- and frequency-domain parameters such as the power ratio, power spectral density, and entropy. Moreover, Hjorth parameters and correlation were employed as features of the valence and arousal scales of the participants. The proposed method yielded an accuracy of 72.87% for high/low valences (Mert and Akan, 2018). In Zheng and Lu (2017a), deep belief networks (DBNs) were trained using differential entropy features extracted from multichannel EEG data, and the average accuracy was 86.08%. In Yin et al. (2017), cross-subject EEG feature selection for emotion recognition was carried out using transfer recursive feature elimination. The classification accuracy was 78.75% in the valence dimension, which was higher than those reported in several studies that used the same database. However, from the calculation times of all the classifiers, it was found that the accuracy of the *t*-test/recursive feature extraction (T-RFE) increased at the expense of the training time. In Li et al. (2018), 18 linear and non-linear EEG features were extracted. In addition, the support vector machine (SVM) method and the leave-one-subject-out verification strategy were used to evaluate the recognition performance. With the automatic feature selection method, the recognition accuracy rate using the dataset for emotion analysis using physiological signals (DEAP) was a maximum of 59.06%, and the recognition accuracy using the SEED dataset was a maximum of 83.33% (Li et al., 2018). In Gupta et al. (2018), the aim of the study was to comprehensively investigate the channel specificity of EEG signals and provide an effective emotion recognition method based on the flexible analytic wavelet transform (FAWT). The average classification accuracy obtained using this method was 90.48% for positive/neutral/negative (SEED) emotion classification, and 79.99% for high valence (HV)/low valence (LV) emotion classification using EEG signals (Gupta et al., 2018). In Li et al. (2019), the accuracy of multisource supervised STM (MS-S-STM) for emotion

recognition accuracy was 88.92%, and the multisource semi-supervised selective transfer machine (STM) (MS-semi-STM) experimental data was used in a transmissive manner, with a maximum accuracy of 91.31%. The methods of emotion recognition across subjects, as employed in the previous studies, require improvements. A method for improving the accuracy of emotion classification is therefore necessary, which requires only a small computational load when applied to the analysis of high-dimensional features. In this study, multiple types of features were extracted. In addition, a two-category emotion recognition method across subjects is proposed. In particular, 10 types of linear and non-linear EEG features were first extracted, and then combined into high-dimensional features. With respect to high-dimensional features, a method for improving the emotion recognition performance across subjects based on high-dimensional features was proposed. Moreover, the significant test/sequential backward selection/support vector machine (ST-SBSSVM) fusion method was proposed and then used to identify and classify the high-dimensional EEG features of the cross-subject emotions.

2. MATERIALS AND METHODS

Figure 1 presents the analysis process in this study. First, 10 types of high-dimensional features were extracted from both the DEAP and SEED. The features were then combined into high-dimensional features, as follows:

$$DEAP, 1280(trials, rows) \times 320(features, cols) \qquad (1)$$
$$SEED, 450(trials, rows) \times 620(features, cols) \qquad (2)$$

For further details, refer to section 2.2. Furthermore, the proposed method (ST-SBSSVM) was used to analyze the high-dimensional features and output the classification and recognition accuracy of positive and negative emotions.

2.1. DEAP Dataset and SEED Dataset
Two publicly accessible datasets were employed for the analysis, namely, the DEAP and SEED. The DEAP dataset (Koelstra et al., 2011) consisted of 32 subjects. Each subject was exposed to 40 1-min long music videos as emotional stimuli

TABLE 1 | Structure of the DEAP dataset.

Array name	Array shape	Array contents
Data	$40 \times 40 \times 8,064$	(Videos/trials) $\times$ channels $\times$ data (128 Hz $\times$ 63 s)
Labels	40×4	(videos/trials) $\times$ labels (Valence, arousal, dominance, liking)

TABLE 2 | Structure of the SEED dataset.

Array name	Array shape	Array contents
Data	$3 \times 15 \times 62 \times 48,000$	(Experiments) $\times$ (Videos/trials) $\times$ channels $\times$ data (200 Hz $\times$ 240 s)
Labels	$3 \times 15 \times 3$	(Experiments) $\times$ (videos/trials) $\times$ labels (positive, neutral, negative)

TABLE 3 | Ten-type EEG features.

Feature type	Extracted features		
The linear features	1. Hjorth activity	2. Hjorth mobility	3. Hjorth complexity
	4. The standard deviation	5. PSD-Alpha	6. PSD-Beta
	7. PSD-Gamma	8. PSD-Theta	
The non-linear features	9. Sample entropy	10. Wavelet entropy	

while their physiological signals were recorded. The resulting dataset includes 32 channels of EEG signals, four-channel electrooculography (EOG), four-channel electromyography (EMG), respiration, plethysmography, galvanic skin response and body temperature. Each subject underwent 40 EEG trials, each of which corresponded to an emotion triggered by a music video. After watching each video, the participants were asked to score their real emotions on a five-level scale: (1) valence (related to the level of pleasure), (2) arousal (related to the level of excitement), (3) dominance (related to control), (4) like (related to preference), and (5) familiarity (related to the awareness of stimuli). The score ranged from 1 (weakest) to 9 (strongest), with the exception of familiarity, which ranged from 1 to 5. The EEG signal was recorded using Biosemi ActiveTwo devices at a sampling frequency of 512 Hz and down-sampling frequency of 128 Hz. The data structure of DEAP is shown in **Table 1**. The SEED (Zheng and Lu, 2017a) consisted of 15 subjects. Movie clips were selected to induce (1) positive emotions, (2) neutral emotions, and (3) negative emotions; each of which were distributed over five segments of each movie. All subjects underwent three EEG recordings, with two consecutive recordings conducted at a two-week interval. At each stage, each subject was exposed to 15 movie clips, each of which was ∼4 min long, to induce specific emotions. The same 15-segment movie clip was used in all three recording sessions. The resulting data contained 15 EEG trials. Each subject underwent 15 trials with 5 trials per emotion. The EEG signals were recorded using a 62-channel NeuroScan electric source imaging (ESI) device at a sampling rate of 1,000 Hz and down-sampling rate of 200 Hz. The data structure of SEED is shown in **Table 2** (the duration of the SEED videos varied: each video was about 4 min = 240 s; thus, the data were about $200 Hz \times 240 s = 48,000$). In this study, only experiments with positive emotions and negative emotions were carried out to evaluate the ability of the proposed method to distinguish between these two emotions. For consistency with the DEAP, 1 min of data extracted at the middle of each trial was employed using the SEED.

2.2. Data Processing
2.2.1. Data Preprocessing
The EEG signal considered in this study was a neurophysiological signal with a high dimensionality, redundancy, and noise. After the EEG data were collected, the original data were pre-processed, i.e., the removal of EOG, EMG artifacts, and down-sampling; to reduce the computational overhead of feature extraction. For the DEAP, the default pre-processing technique was as follows: (1) the data was down-sampled to 128 Hz; (2) the EOG artifacts were removed, as achieved in Koelstra et al. (2011); (3) a bandpass filter with a throughput frequency range of 4.0–45.0 Hz was applied; (4) the data were averaged to the common reference; and (5) the data were segmented into 60-s trials and a 3-second pre-trial baseline. For the SEED, the default preprocessing technique was applied as follows: (1) the data was down-sampled to 200 Hz; (2) a bandpass filter with a throughput frequency range of 0–75 Hz was applied; and (3) the EEG segments corresponding to the duration of each movie were extracted. Prior to the extraction of the power spectral density (PSD) features, four rhythms were extracted, namely, theta (3–7 Hz), alpha (8–13 Hz), beta (14–29 Hz), and gamma (30–47 Hz) (Koelstra et al., 2011). Other features were extracted on the data preprocessed by the dataset.

2.2.2. Label Processing
For label processing using the DEAP, the subjects were divided into two categories according to the corresponding scores of the subjects with respect to valence. A score higher than 5 was set as 1, which represented positive emotions; and a score below 5 was set as 1, which represented negative emotions. In the SEED, the trials were divided into positive emotions, neutral emotions, and negative emotions. However, for consistency with the DEAP, only positive and negative emotion samples were investigated using the SEED. Moreover, binary classification tasks were employed to carry out emotional recognition across the subjects.

2.3. Feature Extraction
Ten types of linear and non-linear features were extracted, as shown in **Table 3**. Several features [Hjorth activity, Hjorth mobility, Hjorth complexity, standard deviation, sample entropy (SampEn), and wavelet entropy (WE)] were directly extracted from the dataset pre-processed EEG signals. The extraction processes of the remaining features (the four PSD frequency domain features) were divided into two steps. First, four types
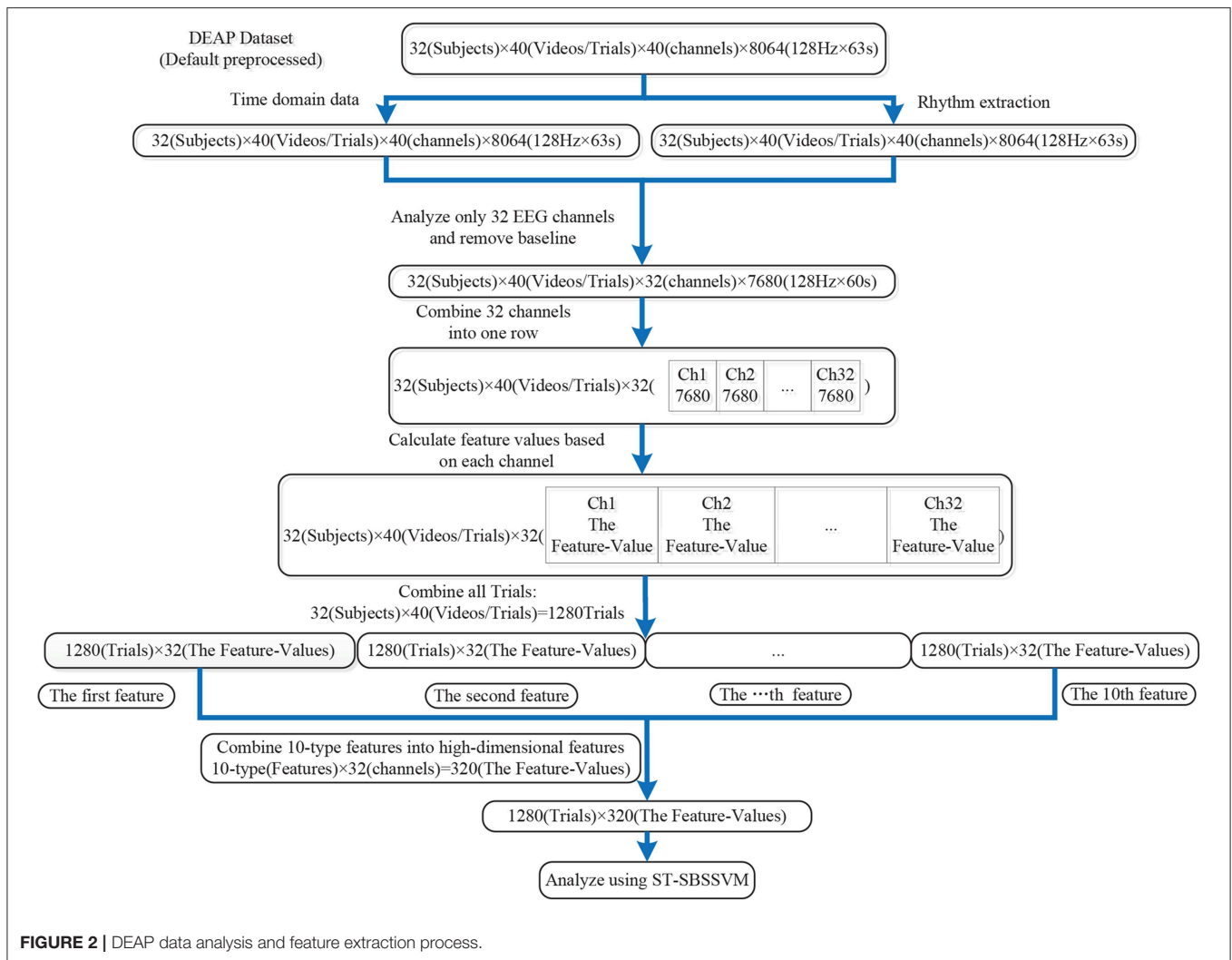
FIGURE 2 | DEAP data analysis and feature extraction process.

of rhythms were extracted from the EEG signals pre-processed using the dataset, and the PSD features were then extracted from the four rhythms. The detailed analysis of the data and feature extraction is shown in **Figures 2**, **3**.

2.3.1. The Linear Feature

Hjorth parameters were indicators of statistical properties used in signal processing in the time domain, as introduced by Hjorth (1970). The parameters are as follows: activity, mobility, and complexity. They were commonly used in the analysis of electroencephalography signals for feature extraction. The parameters are normalized slope descriptors (NSDs) used in EEGs. The standard deviation feature was the standard value of the EEG time-series signal. The four PSD Features were extracted as follows: PSD-alpha was extracted from the alpha rhythm, PSD-beta was extracted from the beta rhythm, PSD-gamma was extracted from the gamma rhythm, and PSD-theta was extracted from the theta rhythm. The power spectrum $S_{xx}(f)$ of a time series $x(t)$ describes the power distribution with respect to the frequency components that compose that signal (Fanelli et al.,

2010). According to Fourier analysis, any physical signal can be decomposed into several discrete frequencies, or a spectrum of frequencies over a continuous range. The statistical average of a certain signal or signal type (including noise), as analyzed with respect to its frequency content, is referred to as its spectrum. When the energy of the signal is concentrated around a finite time interval, especially if its total energy is finite, the energy spectral density can be computed. Moreover, the PSD (power spectrum) is more commonly used, which applies to signals existing over a sufficiently large time period (especially in relation to the duration of a measurement) that can be considered as an infinite time interval. The PSD refers to the spectral energy distribution per unit time, given that the total energy of such a signal over an infinite time interval would generally be infinite. The summation or integration of the spectral components yield the total power (for a physical process) or variance (in a statistical process), which correspond to the values that are obtained by integrating $x^2(t)$ over the time domain, as dictated by Parseval's theorem (Snowball, 2005). For continuous signals over a quasi-infinite time interval, such as stationary processes, the PSD)
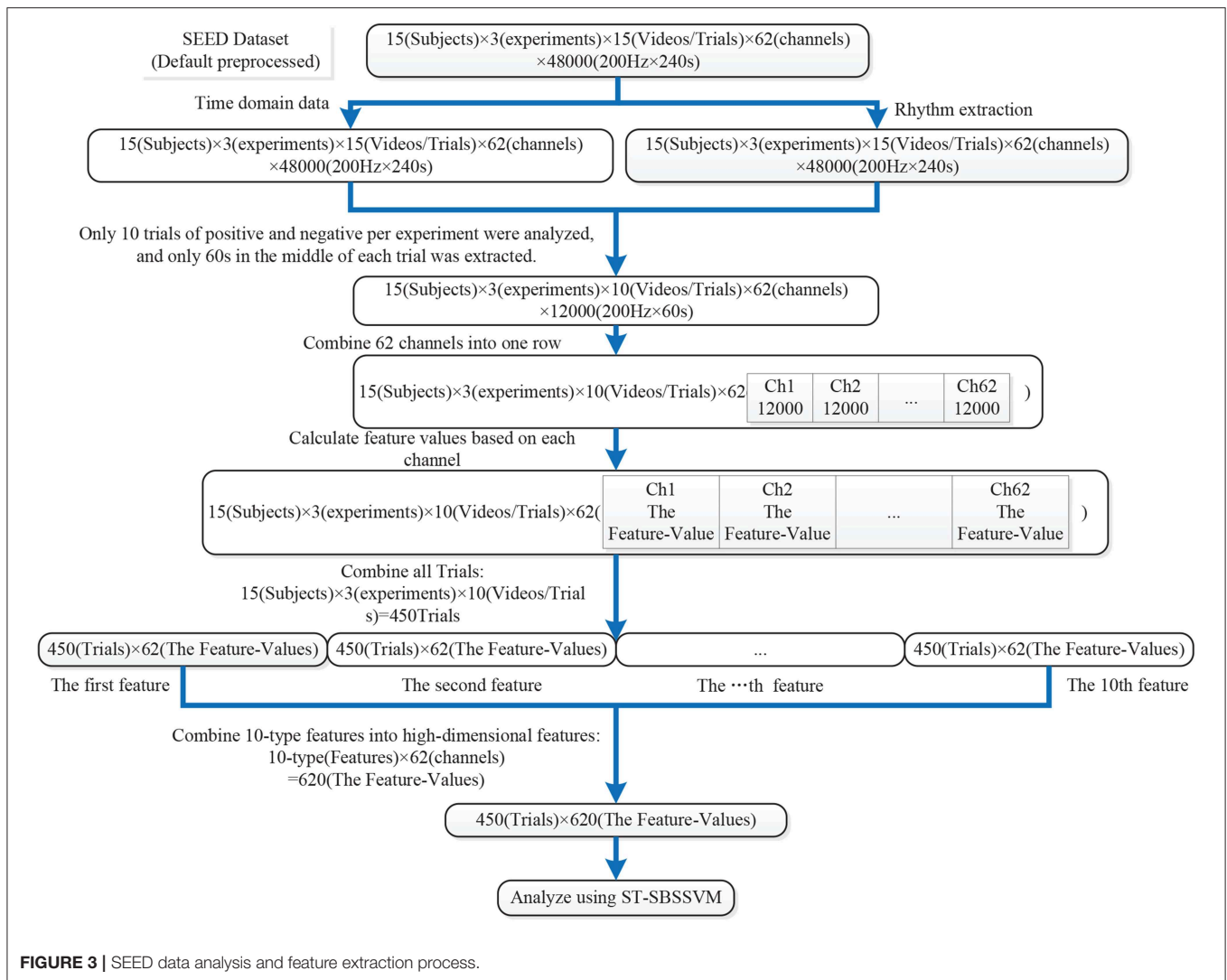
FIGURE 3 | SEED data analysis and feature extraction process.

should be defined, which describes the power distribution of a signal or time-series with respect to frequency.
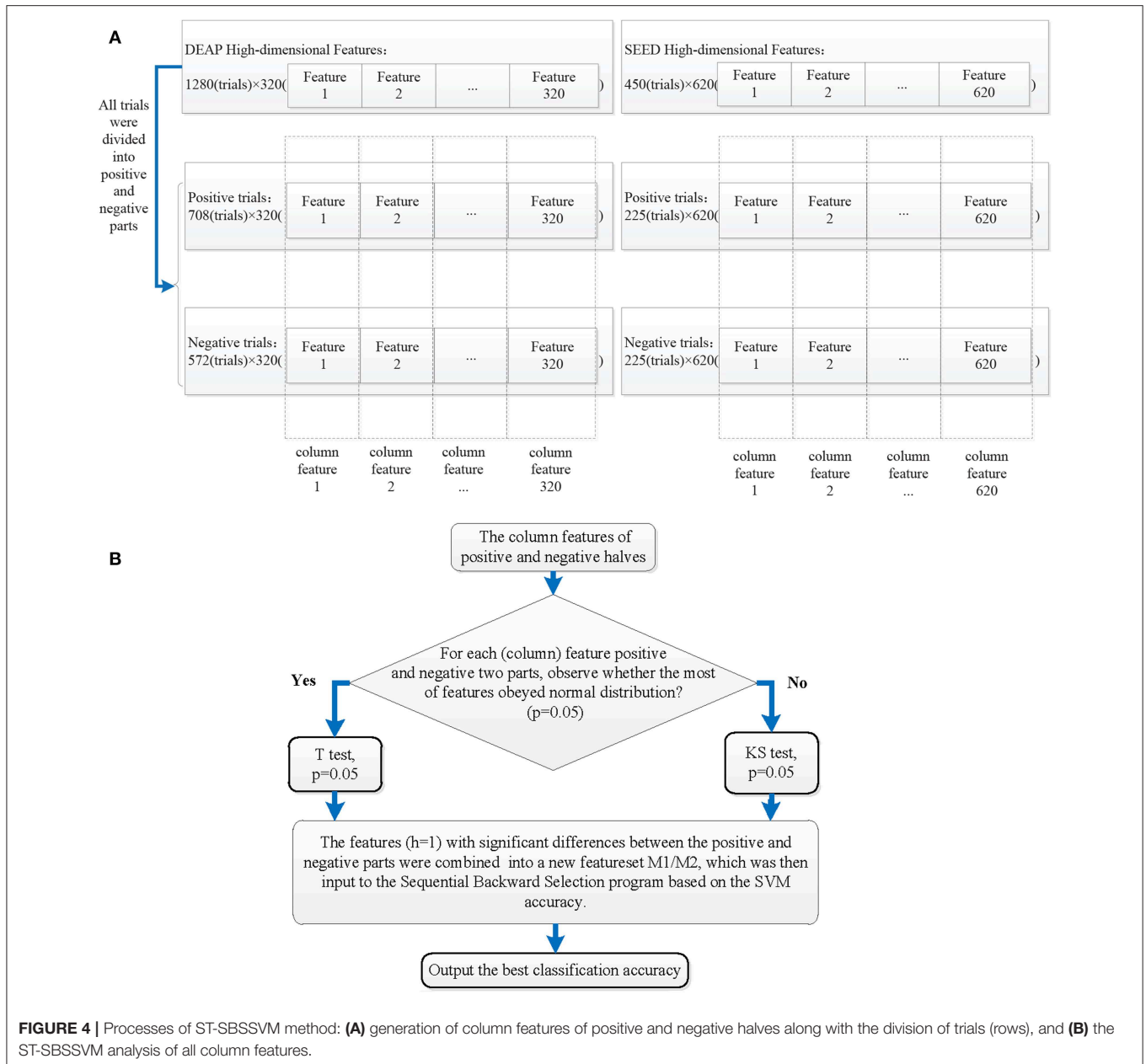
2.3.2. The Non-linear Feature

The SampEn is a modification of the approximate entropy (ApEn), and it is used for assessing the complexity of physiological time-series signals in addition to the diagnosis of diseased states (Richman and Moorman, 2000). Moreover, SampEn has two advantages over ApEn, namely, data length independence and a relatively simple implementation. Similar to ApEn, SampEn is a measure of complexity (Richman and Moorman, 2000). The Shannon entropy provides a useful criterion for the analysis and comparison of probability distributions, which can act as a measure of the information of any distribution; namely, the wavelet entropy (WE) (Blanco et al., 1998). In this study, the total WE was defined as follows:

$$S_{WT} \equiv S_{WT}(p) = \sum_{j<0} p_j \cdot In[p_j] \qquad (3)$$

The WE can be used as a measure of the degree of order/disorder of the signal; thus, it can provide useful information on the underlying dynamical process associated with the signal.

2.4. ST-SBSSVM

The ST-SBSSVM method is a combination of the significance test, sequential backward selection, and support vector machine. In this study, the SVM based on the radial basis function (RBF) kernel was employed. The detailed fusion process is shown in **Figure 4**. Ten types of features from both public datasets were extracted, and high-dimensional features [$DEAP$, $1280(trials, rows) \times 320(features, cols)$; and $SEED$, $450(trials, rows) \times 620(features, cols)$] were formed. If the sequential backward selection (SBS) method was directly employed for the analysis of the high-dimensional EEG features and SVM was used to determine the accuracy of the emotion classification of each feature combination, the computational overhead would be significantly large. Therefore, a method was developed to achieve a higher emotional

FIGURE 4 | Processes of ST-SBSSVM method: **(A)** generation of column features of positive and negative halves along with the division of trials (rows), and **(B)** the ST-SBSSVM analysis of all column features.

recognition accuracy across the subjects than the SBS, namely the ST-SBSSVM. Moreover, the proposed method requires a significantly lower computational overhead for the analysis of high-dimensional EEG features. As shown in step 1 of **Figure 4**, each trial (row) of $1280(trials, rows) \times 320(features, cols)$ and $450(trials, rows) \times 620(features, cols)$ was in one-to-one correspondence with the positive and negative emotion labels. In step 2, according to the labels, all the trials (rows) were divided into two parts. The objective was to simultaneously divide each column feature into two parts. In step 3, the significance test was carried out from the first column feature to the final column feature for the column features that were divided into two positive and negative parts (the last column of the DEAP feature was the 320th column, and the last column of the SEED

feature was the 620th column). It was then determined whether the majority of EEG column features, which were divided, were in accordance with the normal distribution. If the majority of EEGs were subject to the normal distribution, the student's t-test (T-test) was used for the divided column features; otherwise, the Kolmogorov–Smirnov (KS) test was used. The corresponding column features of the positive and negative significant difference (h = 1) were then selected. In step 4, after the significance test, the high-dimensional feature set was simplified, and the following was obtained:

$$M_1 = 1280(trials, rows) \times 68(features, cols) \qquad (4)$$

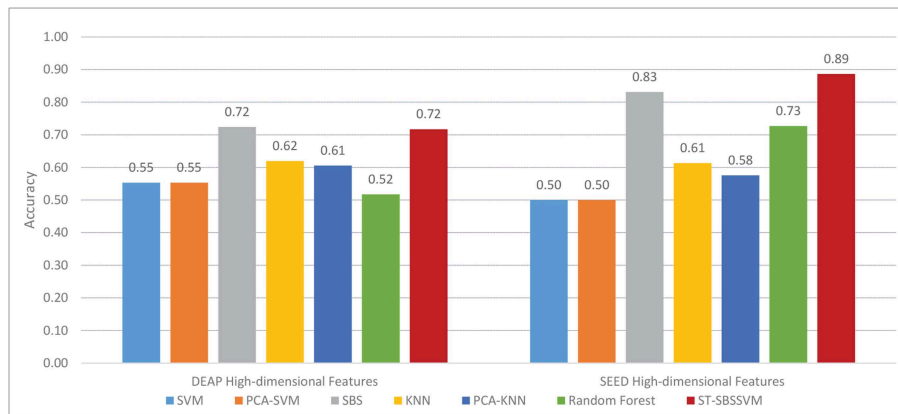$$M_2 = 450(trials, rows) \times 227(features, cols) \qquad (5)$$

FIGURE 5 | Accuracy results of valence classification using DEAP and SEED.

TABLE 4 | Comparison of Valence classification accuracy between ST-SBSSVM and common methods.

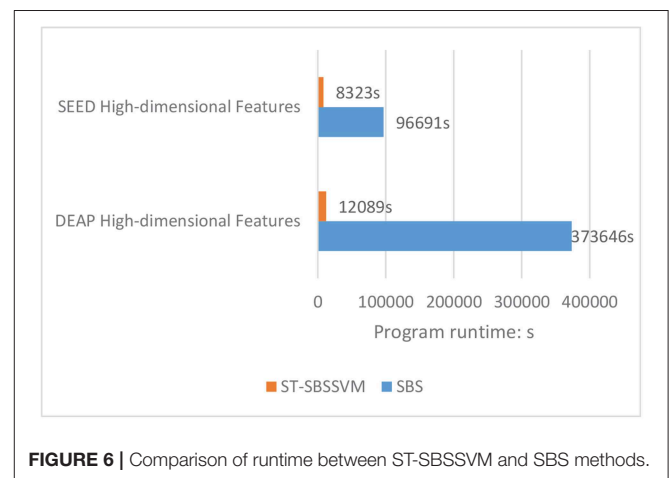|	Difference from ST-SBSSVM accuracy (DEAP)	Difference from ST-SBSSVM accuracy (SEED)
SVM	+17%	+39%
PCA-SVM	+17%	+39%
SBS	−0.42%	+6%
KNN	+10%	+28%
PCA-KNN	+11%	+31%
RF	+20%	+16%
The average difference from ST-SBSSVM accuracy	+12.4%	+26.5%

FIGURE 6 | Comparison of runtime between ST-SBSSVM and SBS methods.

In step 5, M1 and M2 were inputted into the SVM-based SBS program. Sequential backward selection is a process that decreases the number of features, in which a feature is repeatedly eliminated until a final feature is remaining. In this manner, all the feature combinations were separately classified by the SVM. The data was normalized prior to the use of SVM modeling for emotion classification recognition, which helped to improve the convergence rate and accuracy of the model. In the SVM-based SBS program, a "leave-one-subject-out" verification strategy was employed. During each process, the data of one subject was considered as the test set, and the data of the other subjects were considered as the training set. The feature selection was carried out on the training set, and the performance was then evaluated on the test set. This procedure was iterated until the data of each subject had been tested. Moreover, this strategy can eliminate the risk of "overfitting". In step 6, the average classification accuracy of the employed "leave-one-subject-out" verification strategy and SVM-based SBS program was outputted.

3. RESULTS

Figure 5 and **Table 4** present a comparison between the valence classification recognition results of the ST-SBSSVM and those of common methods using the DEAP and SEED. For the consistency of the analysis of the two datasets, only cross-subject emotional recognition was carried out for the valence classification. The ST-SBSSVM method is an improvement of the SBS method; thus, the two methods were compared. **Figure 5** and **Table 4** present a comparison of the recognition accuracies of the two emotions. Therefore, **Figure 6** presents a runtime comparison between the ST-SBSSVM and corresponding SBS program (using the corresponding method for emotion recognition on the same computer "DELL, intel(R) Core (TM) i5-4590 CPU @ 3.30 GHz, RAM-8.00 GB"). As shown in **Figure 5** and **Table 4**, with respect to high-dimensional features, the accuracy of the ST-SBSSVM was improved by 12.4% (DEAP) and 26.5% (SEED) when compared with the common emotion recognition methods. From **Table 4**, it can be seen
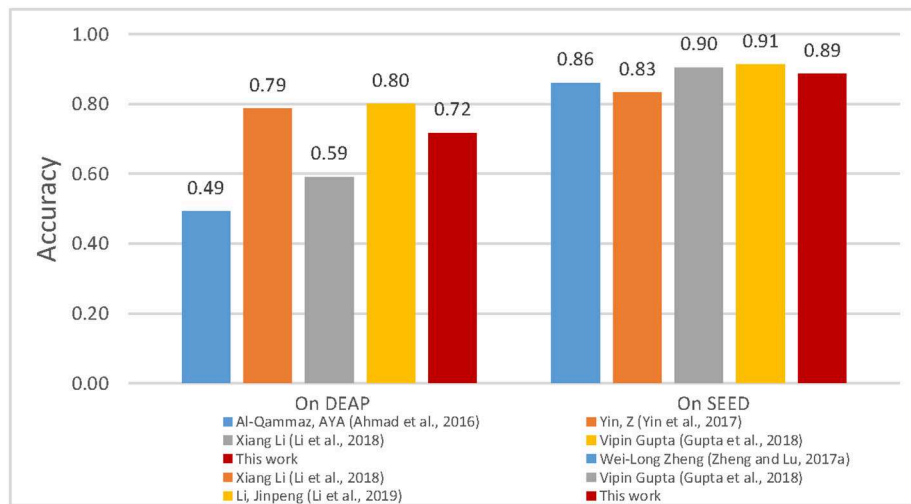
FIGURE 7 | Comparison between valence classification accuracies of similar studies.

that with respect to high-dimensional features, the cross-subject emotion recognition accuracy of the ST-SBSSVM decreased by 0.42% (almost unchanged) using the DEAP, and it improved by 6% using the SEED. **Figure 6** shows that the ST-SBSSVM decreased the program runtime by ∼97 and 91% when compared with the SBS method.

4. DISCUSSIONS

In this paper, a method that can effectively promote emotion recognition is proposed, namely, the ST-SBSSVM method. The proposed method was used to effectively analyze the high-dimensional EEG features extracted from the DEAP and SEED. The results of this study confirmed that the ST-SBSSVM method offers two advantages. First, the ST-SBSSVM can classify and identify emotions, with an improved emotion recognition accuracy. Because ST-SBSSVM performed the Significant Test by comparing the same column feature that had significant difference between positive and negative trials, a "leave-one-subject-out" verification strategy and SVM-based SBS program were then employed to carry out feature selection for those features with significant differences, and the best emotion classification accuracy was obtained. Second, the ST-SBSSVM and SBS methods exhibited similar emotion recognition results to those of the common emotion classification methods. Moreover, when using ST-SBSSVM and SBS to analyze high-dimensional features, ST-SBSSVM decreased the program runtime by ∼90% when compared with SBS. The limitations of this study were as follows. The features extracted were relatively common, and these features were not the new features that significantly promoted emotion recognition in the most recent studies. In future work, the new features combined with ST- SBSSVM will be employed to investigate emotion recognition among subjects. In recent years, several EEG devices and data technologies were developed, such as using wearable EEG devices, for the collection of data in actual working environments (Jebelli et al., 2017b, 2018b;

Chen et al., 2018). High quality brainwaves can then be extracted from the data collected by wearable EEG devices (Jebelli et al., 2017a). A stress recognition framework was proposed, which can effectively process and analyze EEG data collected from wearable EEG devices in real work environments (Jebelli et al., 2018c). These new developments comprise the scope of future research. Similar works are as follows. In (Ahmad et al., 2016), the empirical results revealed that the proposed genetic algorithm (GA) and least squares support vector machine (LS-SVM) (GA-LSSVM) increased the classification accuracy to 49.22% for valence using the DEAP. In Zheng and Lu (2017a), DBNs were trained using differential entropy features extracted from multichannel EEG data. As shown in **Figure 7**, the proposed method demonstrated a good performance and its accuracy was similar to those of achieved in similar studies with respect to the emotion classification recognition of the cross subjects using the same datasets. In summary, when compared with the most recent studies, this method developed in this study was found to be effective for emotional recognition across subjects.

5. CONCLUSIONS

For emotion recognition, a method is proposed in this paper that can significantly enhance the two-category emotion recognition effect; with a small computational overhead when using the corresponding program to analyze high-dimensional features. In this study, 10 types of EEG features were extracted to form high-dimensional features, and the proposed ST-SBSSVM method was employed, which can rapidly analyze high-dimensional features and effectively improve the accuracy of cross-subject emotion recognition, namely the ST-SBSSVM. The results of this work revealed that ST-SBSSVM demonstrates better accuracy with respect to emotion recognition than common classification methods. Compared with the SBS method, the ST-SBSSVM exhibited a higher accuracy of emotion recognition and significantly decreased the program runtime. In comparison

to recent similar methods, the method proposed in this study is effective for emotional recognition across subjects. In summary, the proposed method can effectively promote the emotional recognition across subjects. This method can therefore contribute to the research of health therapy and intelligent human-computer interactions.

DATA AVAILABILITY

The data set generated for this study can be provided to the corresponding author upon request.

AUTHOR CONTRIBUTIONS

FY developed the ST-SBSSVM method, performed all the data analysis, and wrote the manuscript. XZ, WJ, PG, and GL advised data analysis and edited the manuscript.

FUNDING

REFERENCES

Ahern, G. L., and Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: Eeg spectral analysis. *Neuropsychologia* 23, 745–755.

Ahmad, F. K., Al-Qammaz, A. Y. A., and Yusof, Y. (2016). Optimization of least squares support vector machine technique using genetic algorithm for electroencephalogram multi-dimensional signals. *Jurnal Teknologi* 78, 5–10. doi: 10.11113/jt.v78.8842

Ashforth, B. E., and Humphrey, R. H. (1995). Emotion in the workplace: a reappraisal. *Hum. Rel.* 48, 97–125.

Bajaj, V., and Pachori, R. B. (2014). "Human emotion classification from eeg signals using multiwavelet transform," in *2014 International Conference on Medical Biometrics* (Indore: IEEE), 125–130.

Blanco, S., Figliola, A., Quiroga, R. Q., Rosso, O., and Serrano, E. (1998). Time-frequency analysis of electroencephalogram series. iii. wavelet packets and information cost function. *Phys. Rev. E* 57:932.

Candra, H., Yuwono, M., Chai, R., Handojoseno, A., Elamvazuthi, I., Nguyen, H. T., et al. (2015). Investigation of window size in classification of eeg-emotion signal with wavelet entropy and support vector machine. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2015, 7250–7253. doi: 10.1109/EMBC.2015.7320065

Chanel, G., Rebetez, C., Bétrancourt, M., and Pun, T. (2011). Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Trans. Syst. Man Cybernet. Part A. Syst. Hum.* 41, 1052–1063. doi: 10.1109/TSMCA.2011.2116000

Chen, J., and Lin, Z. (2016). "Assessing working vulnerability of construction labor through eeg signal processing," in *16th International Conference on Computing in Civil and Building Engineering* (Hong Kong), 1053–1059.

Chen, J., Lin, Z., and Guo, X. (2018). "Developing construction workers'mental vigilance indicators through wavelet packet decomposition on eeg signals," in *Construction Research Congress 2018: Safety and Disaster Management, CRC 2018* (Hong Kong: American Society of Civil Engineers), 51–61.

Choi, E. J., and Kim, D. K. (2018). Arousal and valence classification model based on long short-term memory and deap data for mental healthcare management. *Healthcare Inform. Res.* 24, 309–316. doi: 10.4258/hir.2018.24.4.309

Doukas, C., and Maglogiannis, I. (2008). Intelligent pervasive healthcare systems, advanced computational intelligence paradigms in healthcare. *Stud. Comput. Intell.* 107, 95–115. doi: 10.1007/978-3-540-77662-8_5

Fanelli, G., Gall, J., Romsdorfer, H., Weise, T., and Van Gool, L. (2010). A 3-d audio-visual corpus of affective communication. *IEEE Trans. Multim.* 12, 591–598. doi: 10.1109/TMM.2010.2052239

Gupta, V., Chopda, M. D., and Pachori, R. B. (2018). Cross-subject emotion recognition using flexible analytic wavelet transform from eeg signals. *IEEE Sensors J.* 19, 2266–2274. doi: 10.1109/JSEN.2018.2883497

Hanjalic, A., and Xu, L.-Q. (2005). Affective video content representation and modeling. *IEEE Trans. Multim.* 7, 143–154. doi: 10.1109/TMM.2004.840618

Hjorth, B. (1970). Eeg analysis based on time domain properties. *Electroencephal. Clin. Neurophysiol.* 29, 306–310.

Horlings, R., Datcu, D., and Rothkrantz, L. J. (2008). "Emotion recognition using brain activity," in *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing* (Delft: ACM), 6.

Hou, X., Liu, Y., Sourina, O., and Mueller-Wittig, W. (2015). "Cognimeter: Eeg-based emotion, mental workload and stress visual monitoring," in *2015 International Conference on Cyberworlds (CW)* (Sydney, NSW: IEEE), 153–160.

Hwang, S., Jebelli, H., Choi, B., Choi, M., and Lee, S. (2018). Measuring workers'emotional state during construction tasks using wearable eeg. *J. Constr. Eng. Manage.* 144:04018050. doi: 10.1061/(ASCE)CO.1943-7862.0001506

Jebelli, H., Hwang, S., and Lee, S. (2017a). Eeg signal-processing framework to obtain high-quality brain waves from an off-the-shelf wearable eeg device. *J. Comput. Civil Eng.* 32:04017070. doi: 10.1061/(ASCE)CP.1943-5487.0000719

Jebelli, H., Hwang, S., and Lee, S. (2017b). Feasibility of field measurement of construction workers'valence using a wearable eeg device. *Comput. Civil Eng.* 99–106. doi: 10.1061/9780784480830.013

Jebelli, H., Hwang, S., and Lee, S. (2018a). Eeg-based workers'stress recognition at construction sites. *Autom. Construct.* 93, 315–324. doi: 10.1016/j.autcon.2018.05.027

Jebelli, H., Khalili, M. M., Hwang, S., and Lee, S. (2018b). A supervised learning-based construction workers'stress recognition using a wearable electroencephalography (eeg) device. *Constr. Res. Congress* 2018, 43–53. doi: 10.1061/9780784481288.005

Jebelli, H., Khalili, M. M., and Lee, S. (2018c). A continuously updated, computationally efficient stress recognition framework using electroencephalogram (eeg) by applying online multi-task learning algorithms (omtl). *IEEE J. Biomed. Health Inform.* doi: 10.1109/JBHI.2018.2870963. [Epub ahead of print].

Jebelli, H., Khalili, M. M., and Lee, S. (2019). "Mobile EEG-based workers' stress recognition by applying deep neural network," in *Advances in Informatics and Computing in Civil and Construction Engineering*, eds I. Mutis and T. Hartmann (Cham: Springer), 173–180.

Jerritta, S., Murugappan, M., Nagarajan, R., and Wan, K. (2011). "Physiological signals based human emotion recognition: a review," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications* (Arau: IEEE), 410–415.

Kim, J. (2007). "Bimodal emotion recognition using speech and physiological changes," in *Robust Speech Recognition and Understanding* (Augsburg: IntechOpen).

Kim, J., and André, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern Analy. Mach. Intell.* 30, 2067–2083. doi: 10.1109/TPAMI.2008.26

Knyazev, G. G., Slobodskoj-Plusnin, J. Y., and Bocharov, A. V. (2010). Gender differences in implicit and explicit processing of emotional facial expressions as revealed by event-related theta synchronization. *Emotion* 10:678. doi: 10.1037/a0019175

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2011). Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15

Lee, Y.-Y., and Hsieh, S. (2014). Classifying different emotional states by means of eeg-based functional connectivity patterns. *PLoS ONE* 9:e95415. doi: 10.1371/journal.pone.0095415

Li, J., Qiu, S., Shen, Y.-Y., Liu, C.-L., and He, H. (2019). Multisource transfer learning for cross-subject eeg emotion recognition. *IEEE Transact. Cyber.* 1–13. doi: 10.1109/TCYB.2019.2904052

Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y., and Hu, B. (2018). Exploring eeg features in cross-subject emotion recognition. *Front. Neurosci.* 12:162. doi: 10.3389/fnins.2018.00162

Liu, Y., and Sourina, O. (2014). "Real-time subject-dependent eeg-based emotion recognition algorithm," in *Transactions on Computational Science XXIII*, eds M. L. Gavrilova, C. J. Kenneth Tan, X. Mao, and L. Hong (Singapore: Springer), 199–223.

Mathersul, D., Williams, L. M., Hopkinson, P. J., and Kemp, A. H. (2008). Investigating models of affect: Relationships among eeg alpha asymmetry, depression, and anxiety. *Emotion* 8:560. doi: 10.1037/a0012811

Mert, A., and Akan, A. (2018). Emotion recognition from eeg signals by using multivariate empirical mode decomposition. *Pattern Anal. Appl.* 21, 81–89. doi: 10.1007/s10044-016-0567-6

Mucci, N., Giorgi, G., Cupelli, V., Gioffrè, P. A., Rosati, M. V., Tomei, F., et al. (2015). Work-related stress assessment in a population of italian workers. the stress questionnaire. *Sci. Total Environ.* 502, 673–679. doi: 10.1016/j.scitotenv.2014.09.069

Nie, D., Wang, X.-W., Shi, L.-C., and Lu, B.-L. (2011). "Eeg-based emotion recognition during watching movies," in *2011 5th International IEEE/EMBS Conference on Neural Engineering* (Shanghai: IEEE), 667–670.

Nijboer, F., Morin, F. O., Carmien, S. P., Koene, R. A., Leon, E., and Hoffmann, U. (2009). "Affective brain-computer interfaces: psychophysiological markers of emotion in healthy persons and in persons with amyotrophic lateral sclerosis," in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (Donostia-San Sebastian: IEEE), 1–11.

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a'low road'to'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* 11:773. doi: 10.1038/nrn2920

Petrantonakis, P. C., and Hadjileontiadis, L. J. (2011). A novel emotion elicitation index using frontal brain asymmetry for enhanced eeg-based emotion recognition. *IEEE Trans. Inform. Techn. Biomed.* 15, 737–746. doi: 10.1109/TITB.2011.2157933

Picard, R. W., Vyzas, E., and Healey, J. (2001). Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 1175–1191. doi: 10.1109/34.954607

Plutchik, R. (1962). The emotions: facts, theories and a new model. *Am. J. Psychol.* 77:518.

Power, M. J., and Dalgleish, T. (1999). *Handbook of Cognition and Emotion.* Singapore: Wiley.

Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circul. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039

Russell, J. A. (1980). A circumplex model of affect. *J. Personal. Soc. Psychol.* 39:1161.

Sammler, D., Grigutsch, M., Fritz, T., and Koelsch, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 44, 293–304. doi: 10.1111/j.1469-8986.2007.00497.x

Shi, L.-C., and Lu, B.-L. (2013). Eeg-based vigilance estimation using extreme learning machines. *Neurocomputing* 102, 135–143. doi: 10.1016/j.neucom.2012.02.041

Snowball, P. S. (2005). Spectral analysis of signals. *Leber Magen Darm* 13, 57–63.

Takahashi, K. et al. (2004). "Remarks on emotion recognition from bio-potential signals," in *2nd International Conference on Autonomous Robots and Agents* (Kyoto), 1148–1153.

Yin, Z., Wang, Y., Liu, L., Zhang, W., and Zhang, J. (2017). Cross-subject eeg feature selection for emotion recognition using transfer recursive feature elimination. *Front. Neurorobot.* 11:19. doi: 10.3389/fnbot.2017.00019

Zeng, Z., Pantic, M., Roisman, G. I., and Huang, T. S. (2008). A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 39–58. doi: 10.1109/TPAMI.2008.52

Zhai, J., Barreto, A. B., Chin, C., and Li, C. (2005). "Realization of stress detection using psychophysiological signals for improvement of human-computer interactions," in *Proceedings IEEE SoutheastCon, 2005* (Miami, FL: IEEE), 415–420.

Zheng, W.-L., and Lu, B.-L. (2017b). A multimodal approach to estimating vigilance using eeg and forehead eog. *J. Neural Eng.* 14:026017. doi: 10.1088/1741-2552/aa5a98

Zheng, W. L., Liu, W., Lu, Y., Lu, B. L., and Cichocki, A. (2019). Emotionmeter: a multimodal framework for recognizing human emotions. *IEEE Trans. Cybernet.* 49, 1110–1122. doi: 10.1109/TCYB.2018.2797176

Zheng, W. L., and Lu, B. L. (2017a). Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Trans. Autonom. Ment. Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497

Zhu, J.-Y., Zheng, W.-L., and Lu, B.-L. (2015). "Cross-subject and cross-gender emotion classification from eeg," in *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada* (Shanghai: Springer), 1188–1191.

Zhuang, N., Zeng, Y., Tong, L., Zhang, C., Zhang, H., and Yan, B. (2017). Emotion recognition from eeg signals using multidimensional information in emd domain. *Biomed Res. Int.* 2017:8317357. doi: 10.1155/2017/8317357