IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 July 2023, accepted 21 August 2023, date of publication 25 August 2023, date of current version 31 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3308811

RESEARCH ARTICLE

EEG-Based Emotion Recognition Using Spatial-Temporal Connectivity

WENHAO CHU[1,2], BAOLE FU[1,2], YUXIAO XIA[1,2], AND YINHUA LIU[1,2,3]

[1]School of Automation, Qingdao University, Qingdao 266071, China
[2]Institute for Future, Qingdao University, Qingdao 266071, China
[3]Shandong Key Laboratory of Industrial Control Technology, Qingdao 266071, China

Corresponding author: Yinhua Liu (liuyinhua@qdu.edu.cn)

This work was supported by the National Key Research and Development Program of China under Grant 2020YFB1313604.

ABSTRACT The connectivity properties between EEG signaling channels were found to play an important role in emotion recognition. During the generation and change of emotions, the connectivity between EEG signal channels is not only manifested in the spatial domain at a specific time, but also interconnected between different time intervals, which is often overlooked. A novel approach is proposed to exploit the shapes of spatial-temporal connectivity for EEG-based emotion recognition. By quantifying the connectivity strength of EEG signal channels within and across time intervals, spatial-temporal connectivity features are extracted, and these features can be represented by shapes in the three-dimensional space of EEG signals. Through these shapes, a mapping of different emotional states and brain activity is constructed. Subsequently, a parallel multi-scale convolutional neural network is employed to extract discriminative features from these connectivity shapes, facilitating the classification and identification of emotional states. Experimental results on the DEAP dataset show that our method achieves excellent performance, with an average classification accuracy of 93.25 % and 93.16 % for the two emotion dimensions of valence and arousal, respectively.

INDEX TERMS Emotion recognition, electroencephalogram, spatial-temporal connectivity.

I. INTRODUCTION

Emotion is a psychological phenomenon that revolves around subjective needs and exerts an influence on diverse cognitive and behavioral processes. The comprehension and precise recognition of emotions hold substantial significance in fields such as psychology, human-computer interaction, and affective computing [1]. The production of electroencephalogram (EEG) signals stems from the spontaneous activity of the brain's neural and physiological regulatory systems. Due to its inherent objectivity, EEG signals are difficult to artificially control or manipulate, making them a valuable avenue for studying emotions [2]. Furthermore, EEG signals offer a non-invasive means of directly measuring brain activity, facilitating a more convenient and accurate assessment of emotional states [3]. Consequently, the exploration of EEG-based automatic emotion recognition has emerged as a prominent subject within brain-computer interface research [4]. Nonetheless, the complexity and dynamic nature of emotional states present notable challenges for EEG-based emotion recognition.

In the realm of EEG-based emotion recognition, the process of feature extraction assumes a crucial role in capturing relevant information from EEG signals for subsequent classification [5]. Numerous types of features have been investigated to characterize the emotional states encoded within EEG data, with these features broadly categorized as temporal, spatial, and connectivity features.

In recent decades, extensive research has been conducted on the identification of key frequency bands and channels for EEG-based emotion recognition. Zheng et al. [6] demonstrated that the most optimal performance was achieved by combining all frequency bands. Another study in [7] demonstrated the representation of EEG signals by combining features of different frequency bands. This approach has proven to be very effective at accurately identifying emotional states in humans. To further leverage the spatial information embedded in EEG signals, Zhong et al. [8] introduced

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Huang.

a regularized graph neural network (RGNN) to capture relationships between different EEG channels, enhancing the capabilities of emotion recognition. Furthermore, the temporal dynamics of EEG signals were recognized as important for accurate emotion recognition. Tao et al. [9] introduced an attention-based convolutional recurrent neural network (ACRNN). This approach fuses the self-attention mechanism with a recurrent neural network (RNN) to effectively capture temporal features. Connectivity features focus on evaluating interactions among distinct brain regions, capturing functional connectivity based on EEG signals. Neuroscience research has confirmed that human emotions are correlated with functional connectivity between regions of the brain [10], [11]. Song et al. [12] employed graph convolutional neural networks(GCNN) to dynamically learn the internal relationships between multiple EEG signal channels. Li et al. [13] utilized multi-scale residual networks to characterize connectivity features and explore their relationship with emotional states. Additionally, several studies have combined multiple feature extraction methods to achieve enhanced results. For instance, Xiao et al. [14] proposed a 4D-based neural network that fused information from spatial, spectral, and temporal domains, capturing discriminative patterns within EEG signals in a four-dimensional space. Li et al. [15] extracted temporal, spatial, and connectivity features from EEG signals for emotion recognition.

Various machine learning methods, such as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN), have been employed by researchers for classifying extracted emotional features [6], [16]. However, these conventional machine learning techniques often fail to fully capture the intricate emotional characteristics embedded within EEG recordings. To address this limitation, deep learning methods were introduced, demonstrating the potential to enhance the accuracy of emotion recognition [17]. Zhang [18] proposed a spatial-temporal recurrent neural network (STRNN) that outperforms traditional methods in recognizing emotions. Wang et al. [19] developed a deep multi-task convolutional neural network using a self-supervised approach, leading to notable improvements in accuracy. On the other hand, Li et al. [15] employed multi-scale residual networks to effectively represent connectivity patterns present in EEG signals. These findings substantiate the beneficial impact of deep learning methods in accurately identifying emotional states.

Although significant progress has been made using existing approaches, they tend to overlook the intricate interplay of spatial and temporal connectivity among brain channels during emotional processing. It is important to recognize that emotions are not confined to specific brain regions or isolated temporal instances. They are intricate and dynamic phenomena that involve interactions among multiple brain pathways [20]. Effectively and accurately integrating the temporal and spatial representations of multi-channel EEG connectivity features becomes crucial in the context of EEG-based emotion recognition.

To address the aforementioned limitations, a novel method for EEG-based emotion recognition is proposed, which leverages spatial-temporal connectivity shapes(ST-C shapes). Three connectivity metrics, namely Pearson correlation coefficient (PCC) [21], phase lock value (PLV) [22], and phase lag index (PLI) [23], are employed to extract connectivity features from the original EEG signals and construct a spatial-temporal connectivity feature matrix. In order to investigate the relationship between connectivity features and emotional states, we employ multi-scale convolution kernels to analyze the connectivity features of EEG signals across multiple scales. By incorporating multi-scale convolutional neural networks with varying receptive fields, we facilitate a comprehensive exploration of dynamic interactions among different brain regions.

The main contributions are as follows:

(1) Distinct emotional states manifest varied EEG signal fluctuations, thereby reflecting alterations in the connectivity characteristics between different brain regions over time. A novel approach is presented herein to enhance EEG-based emotion recognition by leveraging ST-C shapes. In this methodology, the conventional two-dimensional connectivity features are extended by incorporating the time dimension. The dynamic changes in connectivity between different EEG channels over time are captured to construct ST-C shapes within a three-dimensional space. These shapes comprehensively capture the spatial-temporal traits of brain connectivity during diverse emotional states, quantifying the strength of connections between EEG channels within and across time intervals. The introduction of ST-C shapes represents an innovative advancement in EEG signal feature extraction, augmenting the performance of connectivity features in the temporal domain.

(2) By employing multi-scale convolution kernels, our approach comprehensively captures the dynamic interactions among brain regions under different emotional states, using information from various receptive fields. This enables the extraction of multi-scale spatial-temporal connectivity features, which in turn facilitate the identification of ST-C shapes associated with different emotions. Experiments were conducted using the DEAP dataset, and the results showcase the superior performance of the method in accurately recognizing emotional states.

The remainder of this paper is organized as follows: Section II outlines the methodology for constructing the ST-C shapes model. In Section III, experimental results obtained by applying the proposed method are presented and discussed. Finally, the paper concludes by summarizing the key findings and addressing potential avenues for future research in the field of job optimization.

II. METHODOLOGY

A. FRAMEWORK

Currently, a considerable amount of research has been dedicated to emotion recognition based on electroencephalogram

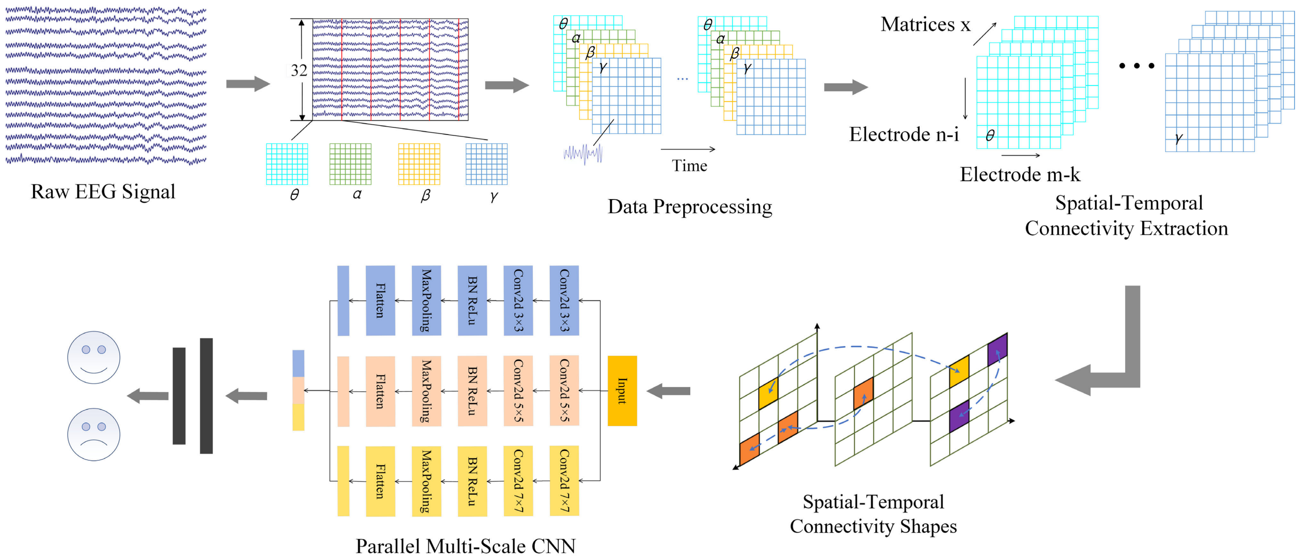

FIGURE 1. **The framework of proposed ST-C Shapes model.**

using connectivity features between brain regions to classify emotional states [11]. It is worth noting that the connectivity between EEG channels undergoes corresponding changes as emotional states fluctuate [24]. In the brain, the dynamic changes in connectivity between EEG channels over time can be represented and translated into connecting lines within the three-dimensional space of the EEG. The combination of these lines gives rise to distinctive shapes in three-dimensional space, capable of depicting EEG activity during various emotional states. Termed as ST-C shapes, these formations epitomize the intricate and evolving interconnections between channels during the course of emotion generation.

To construct these ST-C shapes, our proposed model quantifies the connectivity strengths between EEG channels within and across time intervals. By doing so, the model captures the variations in connectivity features across different emotional states. The shapes obtained accurately depict the complex dynamics of emotions, thereby enhancing the ability of classification models to discriminate between different types of emotions. The ST-C shapes serve as valuable inputs for a parallel multi-scale convolutional neural network. This network plays a crucial role in training and evaluating the model, utilizing the shapes to analyze and extract meaningful information. Finally, sentiment classification is performed using a softmax classifier.

In the proposed framework, the first step involves preprocessing the original EEG signal. The preprocessed EEG signals are divided into four distinct frequency bands within each time window. Subsequently, a cross-temporal connectivity matrix is constructed, enabling the capture of the interconnection of EEG channel connectivity properties over time. This methodology facilitates the extraction of spatial-temporal connectivity features, leading to the formation of ST-C shapes that depict connectivity patterns over time in different emotional states.

A parallel multi-scale convolutional neural network is then employed, utilizing multi-scale convolution kernels to process the obtained connectivity shapes and train the network to discern the variations in shapes associated with different emotional states. The results obtained from the four frequency bands are integrated, considering the comprehensive mapping relationship between different ST-C shapes and emotional states across the frequency bands. Finally, two Fully Connected layers are used to classify the different emotions based on the learned features. The overall framework is illustrated in FIGURE.1.

B. DATA PREPROCESSING

Data preprocessing involves several crucial steps, including removing baseline signals, segmenting time windows, and filtering signal. Typically, raw EEG signals contain baseline signals that are unrelated to the specific emotional state of interest [25]. To improve the success rate of EEG emotion recognition, it is common practice to remove the baseline signal. Baseline correction was performed by subtracting the average value of the signal during the preceding 3-second baseline period from the entire original EEG signal recording. Additionally, raw EEG data often contain artifacts such as eye movement artifacts and electrocardiogram(ECG) artifacts, which can introduce experimental errors. To mitigate these artifacts, blind source separation techniques are employed for artifact removal [26].

To capture the temporal dynamics of emotional states, a sliding time window approach was adopted to segment the EEG signals into individual time intervals. This segmentation strategy enables the continuous data to be divided into non-overlapping time windows of fixed duration. Previous research has shown that using a 3-second sliding window can yield satisfactory classification accuracy [13]. Consequently, a 3-second time window was selected to divide the signal into a series of non-overlapping segments, facilitating

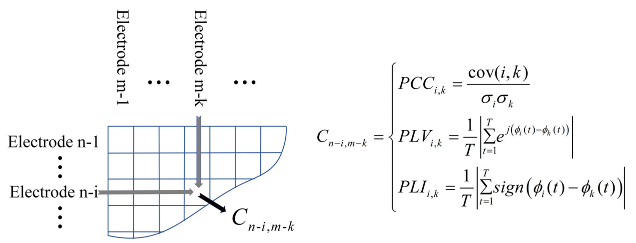

FIGURE 2. **The construction of adjacency matrix.**

the analysis of the EEG signal's temporal characteristics. Moreover, to examine the EEG activity in different frequency bands, a bandpass filter was applied to the EEG signal [27]. This filtering process allowed the isolation of four distinct frequency bands: theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), and gamma (30-50Hz). These preprocessing steps ensure the extraction of relevant temporal and spectral information for subsequent emotion recognition analysis.

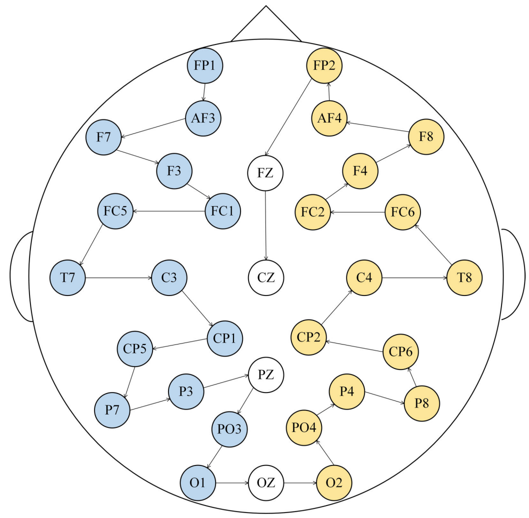

FIGURE 3. **Adjacency matrix channel selection.**

C. ST-C SHAPES

The brain's response to emotions relies on the coordination among multiple brain pathways, highlighting its complex network nature. Consequently, it is valuable to conceptualize the brain as a dynamic network of connectivity to investigate the dynamic interactions among different brain regions. In our proposed ST-C shapes model, the connectivity features is employed to process the preprocessed EEG signals obtained from the four frequency bands at different time intervals. This method quantifies the connection strength between EEG channels within and across time intervals, enabling the construction of a spatial-temporal connectivity matrix. Specifically, after preprocessing, the EEG signal can be represented as $S = S_1, S_2, \ldots, S_N \in R^{T \times F \times V}$ where T denotes the number of experimental time windows, F represents the filtered four frequency bands, and V indicates the number of EEG signal channels. Subsequently, the spatial-temporal connectivity matrix is constructed for each of the four frequency bands using the connectivity method.

As depicted in FIGURE.2, the 32-channel EEG signals within the same frequency band are paired in a one-to-one manner, subsequently leading to the construction of an adjacency matrix through the utilization of connectivity method. This process facilitates the extraction of spatial-temporal connectivity features. By pairing channels within the same time window, the connectivity characteristics between brain regions during that specific time interval are obtained. Moreover, the adjacency matrix across different time intervals enables the extraction of dynamic changes in the connectivity characteristics between channels over time.

The 32 channels of time slice n and time slice m, randomly chosen from the same frequency band, are arranged according to a predetermined structure. Through the utilization of connectivity method, the connectivity between channel i and channel k is calculated, leading to the construction of a time slice adjacency matrix for time slices n and m. This process is repeated across various time slices and EEG channels, resulting in the construction of different spatial-temporal connectivity matrices. The collection of these matrices, denoted as X, represents the dynamic connections between brain regions. The connectivity interlink different EEG channels across different time slices, creating interconnected lines in the three-dimensional space of EEG activities, thereby forming ST-C shapes that describe distinct emotional states. These shapes encapsulate the spatial-temporal connectivity signatures of EEG signals associated with specific emotional states.

To ensure the construction of a coherent and smooth connectivity matrix, it is important to appropriately arrange the EEG channels. The volume conduction effect suggests that brain signals obtained from neighboring brain regions tend to exhibit similarity [15]. Thus, when organizing EEG channels to construct a spatial-temporal connectivity matrix, it is desirable for adjacent channels in the matrix to correspond to adjacent electrodes on the scalp, promoting the construction of a smooth connectivity matrix. The process of electrode selection commences from the left frontal area and proceeds towards the nearest electrode, with a preference for electrodes within the same hemisphere. A total of 32 EEG channels are selected to construct an adjacency matrix, and the resultant arrangement is illustrated in FIGURE.3. This sorting strategy ensures that adjacent channels in the matrix align with adjacent electrodes on the scalp, facilitating the creation of a connectivity matrix characterized by smooth transitions and consistent spatial relationships.

In EEG-based emotion recognition tasks, various measures have been employed to capture the correlation between EEG signal pairs recorded from different channels, as highlighted in [28]. Notably, the Pearson correlation coefficient (PCC), phase-locked value (PLV), and phase-lag index (PLI) have been recognized as effective methods for quantifying the dynamic connections between EEG channels. Building upon this knowledge, our study utilized PCC, PLV, and PLI to construct spatial-temporal connectivity matrix, allowing for a comprehensive characterization of the connectivity dynamics between EEG channels.

PCC is a linear correlation coefficient. By utilizing PCC, the linear correlation between EEG signals from different channels is captured and characterized, thereby revealing complex functional connectivity patterns among brain regions involved in emotional processing.The calculation formula for PCC is as follows:

$$PCC = \frac{cov(i,k)}{\sigma_i \sigma_k} \tag{1}$$

In the formula, i and k represent the EEG signals of two different channels, $cov(i, k)$ is the covariance of i and k, and σ_i and σ_k are the standard deviations of i and k.

PLV characterizes the phase synchronization between two signals and is determined by averaging the absolute values of the phase differences. On the other hand, PLI is an alternative measure of phase synchronization, quantifying the average phase difference between two signals. The specific calculation process for PLV and PLI is as follows:

Given a time series $S_i(t)(t = 1, 2, 3, \ldots, T)$, T represents T time slices extracted from the i-th EEG channel. Now define a complex analytical signal $Z_i(t)$

$$Z_i(t) = S_i(t) + j\tilde{S}_i(t) \tag{2}$$

where $S_i(t)$is the Hilbert transform of $S_i(t)$, $j = (-1)^{1/2}$, the expression of$\tilde{S}_i(t)$ is

$$\tilde{s}_i(t) = \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{s_i(\tau)}{t-\tau} d_\tau \tag{3}$$

where PV is the Cauchy principal value. From this, the instantaneous phase $\phi_i(t)$of channel i can be obtained

$$\phi_i(t) = arctan \frac{\tilde{S}_i(t)}{S_i(t)} \tag{4}$$

From this, the phase difference $\phi_i(t) = (1, 2, 3, \ldots, 32)$ of all channels can be obtained, then the PLV and PLI of channel i and channel k can be calculated.

$$PLV_{i,k} = \frac{1}{T} \left| \sum_{t=1}^{T} e^{j(\phi_i(t)-\phi_k(t))} \right| \tag{5}$$

$$PLI_{i,k} = \frac{1}{T} \left| \sum_{t=1}^{T} sign(\phi_i(t) - \phi_k(t)) \right| \tag{6}$$

After the aforementioned steps, the spatial-temporal connectivity features between EEG channels were successfully captured. These features manifest as lines connecting different channels within the three-dimensional space of the EEG signal, emphasizing the interconnection between channels at various time points. As these lines converge, they give rise to shapes that depict the flow of connectivity features across time and space, resulting in ST-C shapes. The EEG activity of the brain, and the resulting configuration of spatial-temporal connectivity, exhibit variations across different emotional states. These shapes encapsulate the intricate interconnection of brain activity across diverse channels and time windows, establishing a mapping relationship between these shapes and emotional states. Each ST-C shape represents a distinctive arrangement of connectivity patterns, elucidating complex relationships and interactions between brain regions involved in emotional processing. By considering these shapes, a comprehensive understanding of the dynamics of relevant features can be attained, thus facilitating a more nuanced and comprehensive analysis of emotional states. Therefore, ST-C shapes can serve as informative representations that reflect the underlying emotional states.

TABLE 1. Main parameters and their values or types.

Parameter of the Proposed Model	Value/Type
Batch size	128
Learning rate	0.0001
Dropout	0.6
Number of epoch	100
Pooling layer	Max pooling
Activation function	ReLU
Optimizer	Adam
Loss Function	Cross Entropy

D. PARALLEL MULTI-SCALE CNN

Emotional states are known for their intricate and dynamic nature. To fully leverage the spatial and temporal connectivity features inherent in EEG signals, the utilization of a parallel multi-scale convolutional neural network is proposed. CNN have emerged as powerful deep learning architectures widely employed in EEG emotion recognition and various other domains [14]. In contrast to conventional CNN, the parallel multi-scale CNN model excels at extracting multi-scale data from diverse receptive fields and resolutions. This capability enhances its capacity to detect shape-related features and intricate edges, consequently elevating the accuracy and effectiveness of emotion recognition [15]. To provide further insights into the parallel multi-scale CNN model, Table 1 presents the key parameters and pertinent details such as their values or types.

The parallel multi-scale convolutional neural network architecture adopted in this study follows a parallel network design. Upon extracting the ST-C shapes, this shape is utilized as input for the parallel multi-scale CNN. The convolutional block comprises three branches, each employing convolution kernels of sizes 3×3, 5×5, and 7×7, respectively. The distinct sizes of these convolutions enable the network to encompass diverse receptive fields, thereby comprehensively capturing the spatial-temporal characteristics of connectivity and effectively reflecting the dynamic inter-actions among different brain regions. In the context of 3×3 convolutional block branches, the shapes of each layer are as follows: (4, 32, 32), (16, 32, 32), (64, 32, 32), (64, 16, 16), and (64, 1, 1). Similarly, for the 5×5 and 7×7 convolutional block branches, the shape of each layer bears a resemblance to that of the 3×3 convolutional block branch.The input undergoes two convolutional layers, each consisting of four kernels. Subsequently, the network is augmented with a rectified linear unit (ReLU) activation function to introduce

TABLE 2. The performance comparison of the state-of-art strategies on the DEAP dataset.

Authors	Method	Valence		Arousal	
		Accuracy	F1-score	Accuracy	F1-score
Li J et al.	MTL-MSRN [13]	71.92%	0.73	71.29%	0.72
A. Samavat et al.	CNN+BiLSTM [30]	72.38%	0.74	72.38%	0.74
Wang Z et al.	P-GCNN [31]	77.03%	0.75	73.31%	0.76
Gao Y et al.	EEG-GCN [32]	81.70%	0.80	81.95%	0.81
Jia Z et al	MSTGCNN [33]	90.45%	0.91	90.60%	0.91
Yang H et al.	Multi column CNN [34]	90.65%	0.90	90.01%	0.90
Wang H et al.	BDC-DE [35]	89.58%	0.91	89.58%	0.91
ours	ST-C Shapes	93.25%	0.94	93.16%	0.93

nonlinearities. A max pooling layer with a kernel size of 3 × 3 and a stride of 2 is then applied. The Flatten layer serves a vital role in transforming the multi-dimensional feature representation extracted from each frequency band into a one-dimensional feature vector.

Subsequently, the outcomes obtained from different convolution kernel branches are aggregated, combining the spatial-temporal connectivity characteristics of the four frequency bands. The fusion result is then input to the fully connected layer responsible for classifying the emotional state. This classification task involves two fully connected layers within the network architecture. The first fully connected layer, denoted as FC1, comprises 128 neurons. The output layer, referred to as FC2, is composed of two neurons aligned with the number of predicted emotion categories. To obtain the predicted probability for each emotion class, a softmax function is applied to the output of FC2. This activation function ensures that the predicted probabilities sum to 1 and provides a reliable estimate of the likelihood for each sentiment class. The final output is as follows.

$$\mathrm{P}\left(l_i/y\right) = \frac{\exp(y)}{\sum_{i=1}^{N} \exp(y)} \tag{7}$$

III. EXPERIMENT

A. DATASET AND SETTING

The DEAP dataset is a publicly available multimodal dataset designed for emotion recognition, encompassing facial videos, external physiological signals, and EEG signals [29]. It comprises EEG data from 32 participants, with an equal distribution of male and female participants. During the experiment, 40 one-minute-long videos served as emotional stimuli. Following each video, participants rated their emotional responses on a scale ranging from 1 to 9. For the emotion recognition experiment, the analysis focused solely on 32 EEG signal channels, while the remaining channels were excluded from consideration.

The original dataset is partitioned into training data, constituting eighty percent, and test data, comprising twenty percent. To ensure the robustness of the evaluation process, a 10-fold cross-validation technique is employed. The model was trained using the computational capabilities of a Geforce RTX 3090 GPU, with 100 epochs executed on PyTorch platforms. To prevent overfitting, a dropout rate of 0.6 was incorporated during the training process.Additionally, the optimization algorithm employed was Adam. The learning rate is set to 0.0001

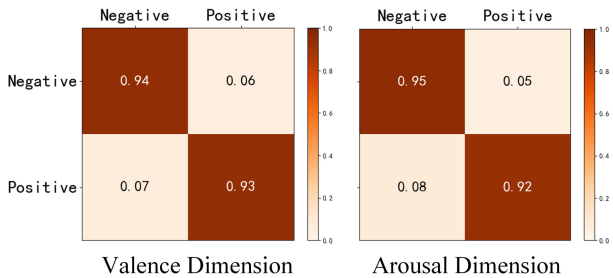

FIGURE 4. Confusion matrix of ST-C shapes model.

B. RESULT AND DISCUSSION

To assess the performance of the proposed ST-C shapes model under different emotional categories, the results obtained from multiple frequency bands are fused, and the classification results for different emotional states are calculated using the DEAP dataset. The confusion matrix, presented in FIGURE. 4, illustrates the model's performance. The results demonstrate that the proposed model achieves high accuracy for both the valence and arousal dimensions of the emotional state, with accuracy rates of 93.25% and 93.16%, respectively. Notably, the model exhibits greater accuracy in identifying negative emotions compared to positive ones. This finding aligns with the psychological perspective that negative emotions tend to have more profound psychological impacts than positive ones [36]. These findings have significant implications for various applications, such as affective computing and mental health assessment.

To further substantiate the superiority of the proposed ST-C shapes model, comprehensive comparisons with several state-of-the-art algorithms are performed. Table 2 presents this comparative analysis, comparing a range of emotion recognition methods applied to the DEAP dataset with our proposed model.

Among these methodologies, Li et al. [13] merge multi-scale residual network (MSRN) with meta-transfer learning (MTL) strategies, offering an investigation into the intricate interplay of connectivity features and emotional states. Notably, convolutional neural networks (CNN) have garnered considerable attention within emotion recognition research due to their adeptness in extracting emotional cues from EEG signals. For instance, Samavat et al. [30] synergize CNN and long short-term memory (LSTM) networks, leading to marked enhancements in emotion recognition accuracy. Similarly, literature [34] introduces a multi-column CNN structural model for bolstering emotion recognition

TABLE 3. The average accuracy on the DEAP Dataset.

Dimension	θ band	α band	β band	γ band	all
Valence	92.06%	90.52%	85.41%	86.92%	93.25%
Arousal	92.13%	90.43%	85.07%	86.73%	93.16%

performance. In recent times, GCNN have gained prominence as a valuable avenue for unraveling emotional insights from EEG signals. In this context, Wang et al. [31] adopt a GCNN grounded in phase-locked value (PLV) connections to elucidate emotion-associated functional connectivity patterns. Reference [32] leverages spatial-temporal and adaptive GCNN, while [33] introduces multi-view spatial-temporal graph convolutional networks fortified with generalization functions. These instances underscore the effectiveness of GCNN-based techniques. Moreover, Wang et al. [35] amalgamate brain-derived connectivity (BDC) features to demonstrate that sensory attributes distilled from high-density EEG signals manifest heightened recognition accuracy. Nonetheless, our proposed approach surpasses the aforementioned mainstream methods in terms of emotion recognition accuracy and F1-score. This demonstrably superior accuracy underlines its proficiency in precisely discerning and categorizing emotional states within EEG signals.

The superior performance of our proposed ST-C shapes model can be attributed to its ability to capture the complex dynamics of EEG signals by integrating both spatial and temporal connectivity features. The remaining methods solely extract diverse types of features from EEG signals without integrating them, consequently overlooking the interrelationships that exist between these features. By considering the intricate connection between brain regions in both space and time, our model provides a more comprehensive and discriminative representation of emotional processes, resulting in improved recognition accuracy.Through the comparison with existing methods, the results demonstrate the superiority of proposed method. This further underscores the important role of spatial-temporal connectivity structure among brain regions in EEG-based emotion recognition. By considering the complex dynamics of brain connectivity, our method not only achieves improved performance but also contributes to a deeper understanding of the underlying mechanisms of emotional processing in the brain.

Subsequently, we delve into the essential components of the model. The presented approach involves the fusion of results obtained from four distinct frequency bands ($\theta, \alpha, \beta,$ *and* γ). In order to comprehensively assess the model's performance, a comparative analysis is conducted by evaluating the results separately for each of the four frequency bands and the combined fusion results of four frequency bands on the DEAP dataset. The frequency range for all bands in the DEAP dataset was set to 4-45 Hz. The experimental results are presented in Table 3. The results indicate that the theta frequency band yields the best performance in terms of the valence and arousal dimensions, achieving accuracies of 92.06% and 92.13%, respectively. However, when the four frequency bands are fused together, the classification results

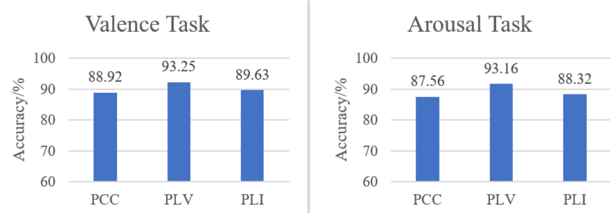

FIGURE 5. Comparison of connectivity method effects.

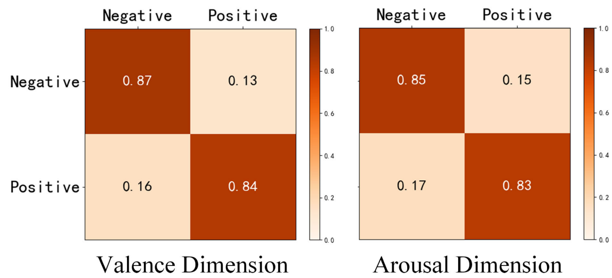

FIGURE 6. Confusion matrix of S-Connectivity model.

are significantly improved compared to using the individual bands alone. The valence and arousal dimensions reach accuracies of 93.25% and 93.16%, respectively.

Based on these experimental findings, the fusion of all four frequency bands yields a substantial improvement in the classification performance. These results highlight the effectiveness of our approach in capturing the intricate dynamics of emotional states and leveraging the combined information from multiple frequency bands to enhance the recognition accuracy.

Various connectivity methods can capture the connectivity features between EEG signal channels, including PCC, PLV, and PLI. To comprehensively assess the effectiveness of different connectivity feature extraction methods and their ability to capture spatial-temporal connectivity features, a comparative experiment on the DEAP dataset was conducted, involving the three connectivity extraction methods. The experimental results are presented in FIGURE. 5.

In terms of valence, the PLV connectivity method achieve the highest accuracy of 93.25%, surpassing PCC by 4.33% and PLI by 3.62%. Similarly, for the arousal dimension, PLV outperforms other connectivity methods with an accuracy rate of 93.16%, surpassing PCC by 5.21% and PLI by 4.84%. These findings indicate that different connectivity methods have varying impacts on the final accuracy of emotion recognition. Through the comparison of various connectivity methods, valuable insights can be obtained regarding their individual capabilities in extracting connectivity features and their suitability for emotion recognition tasks. Among the three connectivity methods, PLV exhibits a stronger ability to capture connectivity features and provides a better reflection of the interplay between brain regions. These results indirectly suggest that PLV connectivity method, which consider phase synchronization, are more closely associated with EEG signals under the influence of emotions.

TABLE 4. The comparison of feature extraction method.

Method	Valence		Arousal	
	Accuracy	F1-score	Accuracy	F1-score
S-Connectivity	86.16%	0.86	85.74%	0.86
ST-C Shapes	93.25%	0.94	93.16%	0.93

To assess the superiority of the proposed spatial-temporal connectivity feature extraction method in comparison with traditional spatial connectivity feature extraction methods in the entire emotion recognition process, experiments are performed on the DEAP dataset while keeping the other parts of the model unchanged. The traditional connectivity feature extraction method is evaluated, and the corresponding confusion matrix of the experiment is displayed in FIGURE.6. Notably, the results demonstrate a significant disparity in the recognition accuracy of emotional states between the traditional connectivity feature extraction methods and our proposed spatial-temporal connectivity feature extraction method.

Table 4 presents a comparative analysis of two model variants in terms of accuracy and F1-score: the S-Connectivity model (which does not account for the temporal interaction of connectivity features between channels) and the ST-C Shapes model (which utilizes the proposed ST-C shapes). The results depicted in Table 4 indicate that the ST-C Shapes model outperforms the S-Connectivity model in terms of recognition accuracy, with a higher F1-score index.

From the results shown in the Table 4, it can be observed that the recognition accuracy of ST-C Shapes is higher than that of S-Connectivity. This finding provides evidence that the temporal interaction of connectivity features between channels plays a crucial role in emotion classification. The superiority of ST-C Shapes further highlights the significance of considering the dynamic nature of connectivity patterns in EEG-based emotion recognition.By capturing the temporal dynamics of connectivity features, the proposed spatial-temporal connectivity shape model enhances the model's ability to discriminate and classify emotions accurately. These results support the effectiveness of our approach in capturing the complex dynamics of emotional processes and improving the performance of emotion recognition systems.

IV. CONCLUSION

A comprehensive framework for EEG-based emotion recognition that leverages ST-C shapes has been presented. By quantifying the connectivity strengths between EEG channels within and across time intervals, the proposed model captures the intricate dynamics of emotional states. The extensive experimentation conducted on the DEAP dataset has not only validated the efficacy of our approach but also showcased its superiority over existing methods. The achieved average classification accuracies of 93.25% and 93.16% for valence and arousal dimensions, respectively, demonstrate the model's remarkable potential.

The incorporation of spatial-temporal connectivity information in EEG-based emotion recognition reinforces the notion that emotions are intricately woven into the dynamics of neural connectivity. As such, this approach advances our understanding of the underlying neural mechanisms governing emotional processes. The significance of spatial-temporal connectivity shapes goes beyond their classification accuracy enhancement; they offer a window into the complexities of the human mind, shedding light on how emotions manifest in the brain's intricate networks.

However, certain limitations warrant consideration. While our model demonstrates substantial improvements, the inherent complexity of emotions might not be fully encapsulated by EEG signals alone. Exploring the integration of multiple modalities and more sophisticated neural architectures could potentially enhance the model's robustness and generalizability. Additionally, conducting experiments on larger and more diverse datasets could further validate the model's performance across a broader spectrum of emotions and individuals.

Looking ahead, the avenue of research remains expansive. Exploring advanced connectivity measures could yield deeper insights into the intricate neural interactions associated with emotions. Furthermore, fusing EEG signals with other physiological signals, such as facial expressions or heart rate variability, holds promise for achieving a more comprehensive and accurate emotion recognition system. Investigating the compatibility of our feature extraction method with a diverse range of neural network architectures is an exciting frontier.

REFERENCES

[1] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cognition Emotion*, vol. 23, no. 2, pp. 209–237, Feb. 2009.

[2] U. Retkoceri, "Remembering emotions," *Biol. Philosophy*, vol. 37, no. 1, p. 5, Feb. 2022.

[3] A. Al-Nafjan, M. Hosny, Y. Al-Ohali, and A. Al-Wabil, "Review and classification of emotion recognition based on EEG brain–computer interface system research: A systematic review," *Appl. Sci.*, vol. 7, no. 12, p. 1239, Dec. 2017.

[4] L. Xin, S. Xiao-Qi, Q. Xiao-Ying, and S. Xiao-Feng, "Relevance vector machine based EEG emotion recognition," in *Proc. 6th Int. Conf. Instrum. Meas., Comput., Commun. Control (IMCCC)*, 2016, pp. 293–297.

[5] B. Fu, C. Gu, M. Fu, Y. Xia, and Y. Liu, "A novel feature fusion network for multimodal emotion recognition from EEG and eye movement signals," *Frontiers Neurosci.*, vol. 17, Aug. 2023, Art. no. 1234162.

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

[7] Y. Yang, Q. Wu, Y. Fu, and X. Chen, "Continuous convolutional neural network with 3D input for EEG-based emotion recognition," in *Proc. Neural Inf. Process., 25th Int. Conf. (ICONIP)*. Siem Reap, Cambodia: Springer, Dec. 2018, pp. 433–443.

[8] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, Jul. 2022.

[9] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, and X. Chen, "EEG-based emotion recognition via channel-wise attention and self attention," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 382–393, Jan. 2023.

[10] H. Cui, A. Liu, X. Zhang, X. Chen, K. Wang, and X. Chen, "EEG-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network," *Knowl.-Based Syst.*, vol. 205, Oct. 2020, Art. no. 106243.

[11] X. Wu, W.-L. Zheng, and B.-L. Lu, "Identifying functional brain connectivity patterns for EEG-based emotion recognition," in *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Mar. 2019, pp. 235–238.

[12] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Trans. Affect. Comput.*, vol. 11, no. 3, pp. 532–541, Jul. 2020.
[13] J. Li, H. Hua, Z. Xu, L. Shu, X. Xu, F. Kuang, and S. Wu, "Cross-subject EEG emotion recognition combined with connectivity features and meta-transfer learning," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105519.
[14] G. Xiao, M. Shi, M. Ye, B. Xu, Z. Chen, and Q. Ren, "4D attention-based neural network for EEG emotion recognition," *Cognit. Neurodynamics*, vol. 16, pp. 805–818, Jan. 2022.
[15] T. Li, B. Fu, Z. Wu, and Y. Liu, "EEG-based emotion recognition using spatial–temporal-connective features via multi-scale CNN," *IEEE Access*, vol. 11, pp. 41859–41867, 2023.
[16] F. Bahari and A. Janghorbani, "EEG-based emotion recognition using recurrence plot analysis and K nearest neighbor classifier," in *Proc. 20th Iranian Conf. Biomed. Eng. (ICBME)*, Dec. 2013, pp. 228–233.
[17] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, "A novel bi-hemispheric discrepancy model for EEG emotion recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 2, pp. 354–367, Jun. 2021.
[18] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–temporal recurrent neural network for emotion recognition," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 839–847, Mar. 2019.
[19] X. Wang, Y. Ma, J. Cammon, F. Fang, Y. Gao, and Y. Zhang, "Self-supervised EEG emotion recognition models based on CNN," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1952–1962, 2023.
[20] X. Liu, T. Li, C. Tang, T. Xu, P. Chen, A. Bezerianos, and H. Wang, "Emotion recognition and dynamic functional connectivity analysis based on EEG," *IEEE Access*, vol. 7, pp. 143293–143302, 2019.
[21] N. Liu, Y. Fang, L. Li, L. Hou, F. Yang, and Y. Guo, "Multiple feature fusion for automatic emotion recognition using EEG signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 896–900.
[22] P. Li, H. Liu, Y. Si, C. Li, F. Li, X. Zhu, X. Huang, Y. Zeng, D. Yao, Y. Zhang, and P. Xu, "EEG based emotion recognition by combining functional connectivity network and local activations," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2869–2881, Oct. 2019.
[23] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Human Brain Mapping*, vol. 28, no. 11, pp. 1178–1193, 2007.
[24] C. Chen, Z. Li, F. Wan, L. Xu, A. Bezerianos, and H. Wang, "Fusing frequency-domain features and brain connectivity features for cross-subject emotion recognition," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–15, 2022.
[25] J. Cai, R. Xiao, W. Cui, S. Zhang, and G. Liu, "Application of electroencephalography-based machine learning in emotion recognition: A review," *Frontiers Syst. Neurosci.*, vol. 15, Nov. 2021, Art. no. 729707.
[26] J. Cheng, M. Chen, C. Li, Y. Liu, R. Song, A. Liu, and X. Chen, "Emotion recognition from multi-channel EEG via deep forest," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 453–464, Feb. 2021.
[27] Y. Luo, C. Wu, and C. Lv, "Cascaded convolutional recurrent neural networks for EEG emotion recognition based on temporal–frequency–spatial features," *Appl. Sci.*, vol. 13, no. 11, p. 6761, Jun. 2023.
[28] M. A. Maria, M. A. H. Akhand, A. B. M. A. Hossain, Md. A. S. Kamal, and K. Yamada, "A comparative study on prominent connectivity features for emotion recognition from EEG," *IEEE Access*, vol. 11, pp. 37809–37831, 2023.
[29] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, and B.-L. Lu, "Data augmentation for enhancing EEG-based emotion recognition with deep generative models," *J. Neural Eng.*, vol. 17, no. 5, Oct. 2020, Art. no. 056021.
[30] A. Samavat, E. Khalili, B. Ayati, and M. Ayati, "Deep learning model with adaptive regularization for EEG-based emotion recognition using temporal and frequency features," *IEEE Access*, vol. 10, pp. 24520–24527, 2022.
[31] Z. Wang, Y. Tong, and X. Heng, "Phase-locking value based graph convolutional neural networks for emotion recognition," *IEEE Access*, vol. 7, pp. 93711–93722, 2019.
[32] Y. Gao, X. Fu, T. Ouyang, and Y. Wang, "EEG-GCN: Spatio-temporal and self-adaptive graph convolutional networks for single and multi-view EEG-based emotion recognition," *IEEE Signal Process. Lett.*, vol. 29, pp. 1574–1578, 2022.
[33] Z. Jia, Y. Lin, J. Wang, X. Ning, Y. He, R. Zhou, Y. Zhou, and L. H. Lehman, "Multi-view spatial–temporal graph convolutional networks with domain generalization for sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1977–1986, 2021.
[34] H. Yang, J. Han, and K. Min, "A multi-column CNN model for emotion recognition from EEG signals," *Sensors*, vol. 19, no. 21, p. 4736, Oct. 2019.
[35] H. Wang, X. Wu, and L. Yao, "Identifying cortical brain directed connectivity networks from high-density EEG for emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1489–1500, Jul. 2022.
[36] R. F. Baumeister, E. Bratslavsky, C. Finkenauer, and K. D. Vohs, "Bad is stronger than good," *Rev. Gen. Psychol.*, vol. 5, no. 4, pp. 323–370, Dec. 2001.

WENHAO CHU received the bachelor's degree in electronic information from the Nanjing University of Technology. He is currently pursuing the master's degree in electronic information with the School of Automation, Qingdao University. His research interests include deep learning and pattern recognition.

BAOLE FU received the bachelor's degree in automation from the Qilu University of Technology, in 2022. He is currently pursuing the master's degree with Qingdao University.

YUXIAO XIA is currently pursuing the master's degree with Qingdao University. Her research interests include human–computer interaction, computer vision, and the recognition of emotion.

YINHUA LIU received the Ph.D. degree in mechanical engineering from Sophia University, Japan, in March 2013.

He joined Qingdao University, in August 2017. He has participated in a number of key research and development projects of the Ministry of Science and Technology. He has published more than ten papers in international and domestic journals and conferences, and has obtained or accepted more than 30 patents in the State Intellectual Property Office. His research interests include intelligent wearable's, offshore photovoltaic, deep learning, and artificial intelligence. In the field of offshore photovoltaic, intelligent parking, smart city and other fields, they have developed a number of industry-university-research innovation projects, and won a number of innovation and entrepreneurship awards. They have accumulated rich project resources and technical teams, among which more than 30 team research and development personnel are specialized in software, machinery, electromechanical, embedded, design and other fields, with interdisciplinary ability and experience. He has a number of product research and development experience and project management experience, won the outstanding mentor of RoboMaster University Single competition, the first prize of Shandong New Generation Mobile Internet Innovation Application skills Competition, and many other awards.