



---

*Research article*

## **Semi-supervised regression with adaptive graph learning for EEG-based emotion recognition**

**Tianhui Sha<sup>1</sup>, Yikai Zhang<sup>1</sup>, Yong Peng<sup>1,2,\*</sup> and Wanzeng Kong<sup>1,2</sup>**

<sup>1</sup> School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, 310018, China

<sup>2</sup> Zhejiang Key Laboratory of Brain-Machine Collaborative Intelligence, Hangzhou, 310018, China

\* **Correspondence:** Email: [yongpeng@hdu.edu.cn](mailto:yongpeng@hdu.edu.cn).

**Abstract:** Electroencephalogram (EEG) signals are widely used in the field of emotion recognition since it is resistant to camouflage and contains abundant physiological information. However, EEG signals are non-stationary and have low signal-noise-ratio, making it more difficult to decode in comparison with data modalities such as facial expression and text. In this paper, we propose a model termed semi-supervised regression with adaptive graph learning (SRAGL) for cross-session EEG emotion recognition, which has two merits. On one hand, the emotional label information of unlabeled samples is jointly estimated with the other model variables by a semi-supervised regression in SRAGL. On the other hand, SRAGL adaptively learns a graph to depict the connections among EEG data samples which further facilitates the emotional label estimation process. From the experimental results on the SEED-IV data set, we have the following insights. 1) SRAGL achieves superior performance compared to some state-of-the-art algorithms. To be specific, the average accuracies are 78.18%, 80.55%, and 81.90% in the three cross-session emotion recognition tasks. 2) As the iteration number increases, SRAGL converges quickly and optimizes the emotion metric of EEG samples gradually, leading to a reliable similarity matrix finally. 3) Based on the learned regression projection matrix, we obtain the contribution of each EEG feature, which enables us to automatically identify critical frequency bands and brain regions in emotion recognition.

**Keywords:** adaptive graph learning; Electroencephalogram (EEG); emotion recognition; graph label propagation; semi-supervised regression

---

### **1. Introduction**

Emotion is a multifaceted psychological state, which includes not only an individual's psychological reaction to external environment or self-stimulation, but also the physiological reactions that accom-

pany this response [1]. In the past period, researchers have studied emotion recognition mainly rely on data derived from an individual's physical behaviours such as sounds from microphones, gestures and expressions from cameras, and text from websites. However, the performance of recognition methods based on the information from these non-physiological cues may be affected by the users' intentionality. When users deliberately disguise their true feelings, the accuracy of emotion recognition methods can be significantly affected [2]. To overcome this limitation, researchers have focused on using EEG signals, which are resistant to camouflage and commonly utilized in the field of emotion recognition in recent years. EEG signals are the reflection of the central nervous system activities and have become a reliable data source for objective emotion recognition [3, 4].

Recent advances in computer technology, biological science, electronic information, modern informatics, and other related fields have led to significant progresses in brain-computer interface (BCI) technology. In the meantime, affective brain-computer interface (aBCI) systems, which are applicable to human emotion state recognition, have received increasing attention in both academia and industry [5]. A typical aBCI system consists of the following several steps including signal acquisition, pre-processing, feature extraction or learning, and emotional state identification by recognition models [6]. In this paper, the study corresponds mainly to the latter two stages. In recent years, more and more machine learning models have been proposed for EEG emotion recognition [7]. For example, Li *et al.* explored robust EEG features in cross-subject sentiment recognition using support vector machines combined with the "leave-one-subject-out" strategy [8]. Dan *et al.* proposed a probabilistic cluster-promoting semi-supervised learning method which introduces a regularization term on fuzzy entropy to obtain a more generalized label affiliation function to reduce the effect of EEG noise [9]. Moreover, the experimental results also confirm its robustness.

However, there are two fundamental problems in the existing studies within the machine learning-based EEG emotion recognition research. On one hand, it is how to leverage the properties of EEG data to fit well into the emotion recognition process and improve its effectiveness. Specifically speaking, the trained model should consider the samples' local manifold structure, *i.e.*, the high-similar samples can be grouped into the same cluster and the low-similar samples can be separated. On the other hand, in addition to improving recognition accuracy, we hope to gain some additional insights related to emotional effects from the model. In specific, the learned graph similarity matrix can well depict the information of emotional states of EEG data. However, in some existing graph-based models, the learned graphs may not have been further investigated. In this paper, we use semi-supervised regression combined with adaptive graph similarity learning strategy to deal with these two problems.

Least square regression (LSR) has been widely used in pattern classification due to its efficacy and simplicity [10]. However, LSR is easily influenced by specific labeled samples, which may result in the selected features not necessarily being optimal [11]. Therefore, to improve the performance of LSR-based learning models in EEG emotion recognition, we construct a similarity matrix to explore the manifold structure of EEG data. In graph-based semi-supervised learning models, a two-stage strategy is commonly used. That is, a similarity graph is firstly constructed based on a certain distance metric and then, label propagation or other learning tasks is performed on the constructed graph [12, 13]. However, these models have some shortcomings that cannot be ignored. Firstly, it is not appropriate to construct a graph in the original space where there might be some noises. In addition, the two-stage strategy breaks the internal connection so that a flawed similarity matrix is learned, which may produce suboptimal results and accordingly reduce the model performance.

Therefore, how to effectively construct a high quality graph becomes a common challenge in many pattern classification tasks. Though the traditional KNN-based methods are simple to implement, they are not always effective because of the noise in the raw data. To alleviate the existing problems, Kang *et al.* proposed a scalable graph learning framework which used the relationship between samples and anchor points depicted by the bipartite graph to address the challenges of data noise [14]. However, the number of anchor points is a key factor in determining the quality of the learned graph. Generally, the best number of anchor points is selected by expert experience or trial-and-error, which is risk and inefficient. In order to reduce this factor's effect, it is necessary to construct graphs that are adaptive to the number of neighboring points.

Adaptive graph learning is a popular approach that combines graph construction and learning tasks into a unified framework. In recent years, many adaptive graph learning methods have been proposed. For example, a structured optimal graph based sparse feature extraction model was proposed in [15], which integrated sparse representation, local structure learning and label propagation into a unified learning framework. Besides, Lin *et al.* adopted an automatic learning scheme by taking advantage of the data properties and structure information to learn a reliable and precise graph [16]. The methods mentioned above can adaptively learn the similarity matrix, but they may not fully utilize label information in constructing the similarity matrix, which still has the room to be improved.

Inspired by [17, 18], to enhance the reliability of manifold structures and adaptive graph learning, we utilize label information in both preserving manifold structures and adapting the graph learning process. In specific, we use semi-supervised regression to obtain the emotional states of the unlabeled EEG samples, and then the obtained label information is used to construct the sample similarity graph. These two processes iterate back and forth towards the optimum. In summary, we propose a semi-supervised regression with adaptive graph learning (SRAGL) model for EEG-based emotion recognition. In comparison to previous research, the following aspects are the primary contributions of this paper.

- A novel semi-supervised regression with adaptive graph learning (SRAGL) model is proposed for cross-session EEG emotion recognition. It achieves joint optimization of semi-supervised regression, structured graph learning and label propagation in a unified framework. In addition, the objective function of SRAGL is optimized using an efficient iterative algorithm.
- SRAGL efficiently captures the underlying data connections of both unlabeled and labeled EEG samples, and dynamically updates the graph similarity matrix through LSR and label propagation to make the learned graph more reliable. Intuitively, the graph block diagonals are gradually apparent, respectively corresponding to the emotional states.
- The performance of SRAGL is verified on the public SEED-IV data set, and the results demonstrate that it outperforms several other compared methods in terms of emotion recognition accuracy. In addition, according to the learned projection matrix in SRAGL, we provide an automatic and quantitative approach to identify key EEG bands and brain regions in cross-session emotion recognition.

**Notations.** The matrices and vectors in this paper are denoted by uppercase letters in bold and lowercase letters in bold.  $\mathbf{1}_n = (1, 1, \dots, 1)^T$  is a column vector whose elements are all ones and the subscript  $n$  denotes its length. For matrix  $\mathbf{A}$ ,  $\mathbf{a}^i$  and  $\mathbf{a}_j$  stand for the  $i$ -th row vector and the  $j$ -th column vector, respectively. The  $\ell_{2,1}$ -norm of matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is described as  $\|\mathbf{A}\|_{2,1} = \sum_{i=1}^m \sqrt{\sum_{j=1}^n A_{ij}^2} =$

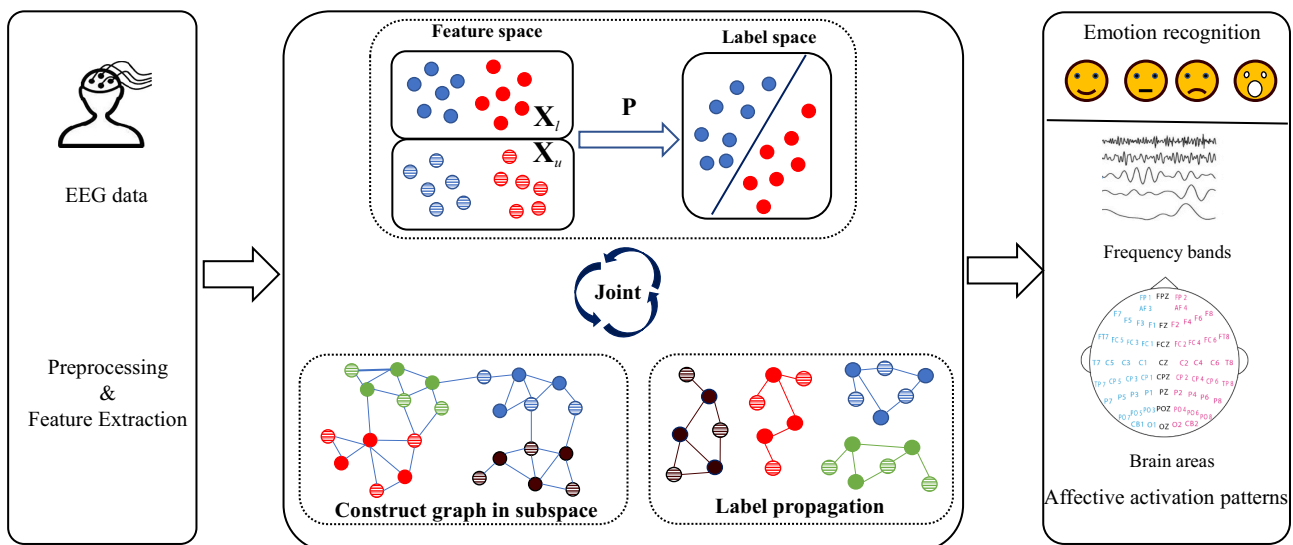
$$\sum_{i=1}^m \|\mathbf{a}^i\|_2.$$

## 2. The Proposed Method

### 2.1. Preliminary

Suppose that we have  $n$  EEG samples  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l, \mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_{l+u}] \in \mathbb{R}^{d \times n}$ , the first  $l$  samples are labeled and the remaining samples are unlabeled.  $d$  is the dimensionality of the EEG feature. We define  $\mathbf{Y} = [\mathbf{Y}_l; \mathbf{Y}_u] \in \mathbb{R}^{n \times c}$  as the label matrix for all samples, where  $c$  represents the number of emotional states and  $\mathbf{Y}_l$  stands for the known label matrix. In specific, if the sample  $\mathbf{x}_i$  belongs to the  $j$ -th class and the label vector is  $\mathbf{y}^i$ ,  $Y_{ij} = 1$  and the rest elements are zeros.

Figure 1 illustrates the general framework of the proposed SRAGL model. It is composed of three main parts, which are the semi-supervised regression, graph construction, and semi-supervised label propagation, respectively. Based on SRAGL, our task are two folds. On one hand, we aim to accurately predict the emotional states of the target session's EEG data  $\mathbf{Y}_u$  based on the inputs  $\mathbf{X}$  and  $\mathbf{Y}_l$ . On the other hand, we need to determine the relative importance of all EEG features, and further the importance of all frequency bands and channels in a quantitative manner. This will allow us to identify the most critical features that contribute significantly to the accuracy of our predictions.



**Figure 1.** The general framework of the SRAGL model.

### 2.2. Model Formulation

In the traditional classification task, the sample's feature can be projected into the label space and the optional transformation matrix is learned by minimizing the fitting error. Accordingly, the model can be established as

$$\min_{\mathbf{P}} \|\mathbf{X}^T \mathbf{P} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{P}\|_{2,1}, \quad (2.1)$$

where  $\mathbf{P} \in \mathbb{R}^{d \times c}$  is the transformation matrix. The first term is the loss function when projecting samples' features into label space and the second term represents the regularization term which controls the row sparsity of the transformation matrix  $\mathbf{P}$ .

According to the graph theory, we use EEG samples as vertices and the connection weights between samples as edges. We define the sample similarity matrix  $\mathbf{S} \in \mathbb{R}^{n \times n}$  to describe the essential relationship between samples. Different from obtaining the similarity between two samples by calculating the Euclidean distance of their features, we expect to learn  $\mathbf{S}$  from data adaptively. Therefore, we can obtain the similarity matrix  $\mathbf{S}$  using the following formula

$$\min_{\mathbf{S} \geq \mathbf{0}, \mathbf{S}\mathbf{1}_n = \mathbf{1}_n} \sum_{i,j=1}^n \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 S_{ij}, \quad (2.2)$$

where the constraint  $\mathbf{S} \geq \mathbf{0}$  represents all elements of the similarity matrix are non-negative and  $\mathbf{S}\mathbf{1}_n = \mathbf{1}_n$  denotes that the sum of degree of similarity between the sample  $\mathbf{x}_i$  and the others is 1. From equation (2.2), it is realized that if the closer the Euclidean distance of two samples, the more similar two samples are, and the corresponding weight should be larger.

The local structure preservation method is introduced in order to thoroughly examine the local structure features of samples so that the proposed model can explore more discriminative information in EEG data. To be specific, if two samples in the original space are similar, they should maintain a similar connection when they are projected into the subspace. Therefore, we define  $\mathbf{W} \in \mathbb{R}^{d \times m}$  as the projection matrix and  $m$  is dimensionality of the subspace. Moreover, there exists a trivial solution for problem (2.2); that is, for a certain sample  $\mathbf{x}_i$ , the weight  $S_{ij}$  would be 1 if  $\mathbf{x}_j$  is its nearest neighbor. Above all, in order to avoid such non-ideal case and make data similarities more accurate, we introduce a term in equation (2.2) to shrink the elements in  $S^i$ . Besides, we use  $\mathbf{W}^T \mathbf{x}_i$  to replace  $\mathbf{x}_i$  for better exploring the local structure information. Therefore, we have

$$\begin{aligned} \min_{\mathbf{W}, \mathbf{S}} \sum_{i,j=1}^n (\|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_2^2 S_{ij} + \alpha S_{ij}^2), \\ \text{s.t. } \mathbf{S} \geq \mathbf{0}, \mathbf{S}\mathbf{1}_n = \mathbf{1}_n, \mathbf{W}^T \mathbf{W} = \mathbf{I}_m, \end{aligned} \quad (2.3)$$

where  $\alpha$  is a regularization parameter to balance the impacts of the two terms. The constraint on the projection matrix  $\mathbf{W}$  defines an orthogonal subspace.

In problem (2.3), it just evaluates data similarity in feature space and ignores label space consistency. In other words, if two EEG samples are similar, then their labels should also belong to the same emotional state as much as possible. Mathematically, the model is represented as

$$\begin{aligned} \min_{\mathbf{Y}_u \geq \mathbf{0}, \mathbf{Y}_u \mathbf{1}_c = \mathbf{1}_u} \sum_{i,j=1}^n \|\mathbf{y}^i - \mathbf{y}^j\|_2^2 S_{ij} \\ \Leftrightarrow \min_{\mathbf{Y}_u \geq \mathbf{0}, \mathbf{Y}_u \mathbf{1}_c = \mathbf{1}_u} \text{Tr}(\mathbf{Y}^T \mathbf{L}_S \mathbf{Y}), \end{aligned} \quad (2.4)$$

where  $\mathbf{L}_S \in \mathbb{R}^{n \times n}$  is the Laplacian matrix and can be calculated by  $\mathbf{L}_S = \mathbf{D} - \mathbf{S}$ .  $\mathbf{D}$  is a diagonal matrix and the value of the  $i$ -th element  $D_{ii}$  is calculated by  $\sum_{j=1}^n S_{ij}$ .  $\mathbf{y}^i$  and  $\mathbf{y}^j$  respectively correspond to the label indicator vectors of samples  $\mathbf{x}_i$  and  $\mathbf{x}_j$ . In equation (2.4), the two constraints require that for each row in  $\mathbf{Y}_u$ , all its elements are non-negative and their summation is 1. Equivalently, each element of  $\mathbf{y}^i$  means the probability of the  $i$ -th unlabeled sample belongs to a certain emotion. For example, if the predict label of an EEG sample is [0.12, 0.75, 0.08, 0.05], then this sample has a high probability to the second emotional state.

Finally, by combining equations (2.1), (2.3), and (2.4) together, the objective function of SRAGL is obtained as

$$\begin{aligned} & \min_{\mathbf{W}, \mathbf{S}, \mathbf{P}, \mathbf{Y}_u} \|\mathbf{X}^T \mathbf{P} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{P}\|_{2,1} \\ & + \beta \sum_{i,j=1}^n (\|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_2^2 S_{ij} + \alpha S_{ij}^2) + \gamma \text{tr}(\mathbf{Y}^T \mathbf{L}_S \mathbf{Y}), \\ & \text{s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I}_m, \mathbf{Y}_u \mathbf{1}_c = \mathbf{1}_u, \mathbf{Y}_u \geq \mathbf{0}, \mathbf{S} \mathbf{1}_n = \mathbf{1}_n, \mathbf{S} \geq \mathbf{0}, \end{aligned} \quad (2.5)$$

where  $\lambda, \beta, \alpha, \gamma$  are the regularization parameters.

### 2.3. Optimization

Below the optimization method to objective function (2.5) is introduced; that is, we propose to update the four variables  $\mathbf{W}$ ,  $\mathbf{S}$ ,  $\mathbf{P}$  and  $\mathbf{Y}_u$  by solving one variable and fixing others.

- **Update P.** The objective function with respect to  $\mathbf{P}$  is

$$\min_{\mathbf{P}} \|\mathbf{X}^T \mathbf{P} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{P}\|_{2,1}. \quad (2.6)$$

Because the regulation term  $\|\mathbf{P}\|_{2,1}$  is equal to  $\text{Tr}(\mathbf{P}^T \mathbf{Q} \mathbf{P})$ , we can change equation (2.6) as

$$\min_{\mathbf{P}} \|\mathbf{X}^T \mathbf{P} - \mathbf{Y}\|_F^2 + \lambda \text{Tr}(\mathbf{P}^T \mathbf{Q} \mathbf{P}), \quad (2.7)$$

where  $\mathbf{Q}$  is a  $d \times d$  diagonal matrix and the value of its  $i$ -th diagonal element is

$$Q_{ii} = \frac{1}{2 \sqrt{\|\mathbf{p}^i\|_2^2 + \varepsilon}}. \quad (2.8)$$

Here  $\varepsilon$  is a small value which is used to prevent the denominator from being 0 due to the row sparsity of  $\mathbf{P}$ . By making that function's partial derivative corresponding to problem (2.7) with respect to  $\mathbf{P}$  to 0, we can obtain

$$\mathbf{P} = (\mathbf{X} \mathbf{X}^T + \lambda \mathbf{Q})^{-1} \mathbf{X} \mathbf{Y}. \quad (2.9)$$

- **Update W.** The objective function with respect to  $\mathbf{W}$  is

$$\min_{\mathbf{W}^T \mathbf{W} = \mathbf{I}_m} \sum_{i,j=1}^n \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_2^2 S_{ij}. \quad (2.10)$$

It is equal to

$$\min_{\mathbf{W}^T \mathbf{W} = \mathbf{I}_m} \text{Tr}(\mathbf{W}^T \mathbf{X} \mathbf{L}_S \mathbf{X}^T \mathbf{W}), \quad (2.11)$$

$\mathbf{W}$  is solved by computing the eigenvectors that correspond to the  $m$  smallest eigenvalues of  $\mathbf{X} \mathbf{L}_S \mathbf{X}^T$ .

- **Update S.** The objective function with respect to  $\mathbf{S}$  is

$$\begin{aligned} & \min_{\mathbf{S}} \sum_{i,j=1}^n (\|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_2^2 S_{ij} + \alpha S_{ij}^2) + \frac{\gamma}{\beta} \text{tr}(\mathbf{Y}^T \mathbf{L}_S \mathbf{Y}). \\ & \text{s.t. } \mathbf{S} \mathbf{1}_n = \mathbf{1}_n, \mathbf{S} \geq \mathbf{0} \end{aligned} \quad (2.12)$$

Because each  $i$  of problem (2.12) is independent, we can solve  $\mathbf{S}$  in row-wise manner. For the  $i$ -th row, *i.e.*,  $\mathbf{s}^i$ , we have the following objective function

$$\min_{\mathbf{s}^i \geq \mathbf{0}, \mathbf{s}^i \mathbf{1}_n = \mathbf{1}_n} \|\mathbf{s}^i + \frac{1}{2\alpha} \mathbf{a}^i\|_2^2, \quad (2.13)$$

where  $\mathbf{a}^i \in \mathbb{R}^{1 \times n}$  is a row vector and the  $j$ -th value is  $a_{ij} = a_{ij}^x + a_{ij}^y$ . Among, we define  $a_{ij}^x = \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_2^2$  and  $a_{ij}^y = \frac{\gamma}{\beta} \|\mathbf{y}^i - \mathbf{y}^j\|_2^2$ . The Lagrangian function of the problem (2.13) is

$$L(\mathbf{s}^i, \mu, \boldsymbol{\theta}_i) = \frac{1}{2} \|\mathbf{s}^i + \frac{1}{2\alpha} \mathbf{a}^i\|_2^2 - \mu(\mathbf{s}^i \mathbf{1}_n - 1) - \mathbf{s}^i \boldsymbol{\theta}_i, \quad (2.14)$$

where  $\mu$  and  $\boldsymbol{\theta}_i$  are Lagrangian multipliers. On the basis of the KKT condition [19], we can obtain the solution of  $\mathbf{s}^i$

$$S_{ij} = \left(-\frac{1}{2\alpha} a_{ij} + \mu\right)_+, \quad (2.15)$$

where  $(g(x))_+ = \max(g(x), 0)$ .

- **Update  $\mathbf{Y}_u$ .** The objective function with respect to  $\mathbf{Y}_u$  is

$$\min_{\mathbf{Y}_u \geq \mathbf{0}, \mathbf{Y}_u \mathbf{1} = \mathbf{1}} \|\mathbf{X}^T \mathbf{P} - \mathbf{Y}\|_F^2 + \gamma \text{tr}(\mathbf{Y}^T \mathbf{L}_S \mathbf{Y}). \quad (2.16)$$

By completing the squared form for each  $i|_{l+1}^n$ , the problem (2.16) is changed into

$$\min_{\mathbf{y}^i \geq \mathbf{0}, \mathbf{y}^i \mathbf{1}_c = \mathbf{1}_u} \sum_{i=l+1}^n (\|\mathbf{x}_i^T \mathbf{P} - \mathbf{y}^i\|_2^2 + \gamma \|\mathbf{y}^i - \mathbf{m}\|_2^2), \quad (2.17)$$

where  $\mathbf{m} \triangleq \sum_{j=l+1, j \neq i}^n \mathbf{y}^j S_{ij} \in \mathbb{R}^{1 \times c}$ . Obviously, for each  $i$ , the above objective function can be expressed as

$$\begin{aligned} & \min_{\mathbf{y}^i \geq \mathbf{0}, \mathbf{y}^i \mathbf{1}_c = \mathbf{1}_u} \|\mathbf{x}_i^T \mathbf{P} - \mathbf{y}^i\|_2^2 + \gamma \|\mathbf{y}^i - \mathbf{m}\|_2^2 \\ & \Leftrightarrow \min_{\mathbf{y}^i \geq \mathbf{0}, \mathbf{y}^i \mathbf{1}_c = \mathbf{1}_u} (\mathbf{x}_i^T \mathbf{P} - \mathbf{y}^i)(\mathbf{x}_i^T \mathbf{P} - \mathbf{y}^i)^T + \gamma (\mathbf{y}^i - \mathbf{m})(\mathbf{y}^i - \mathbf{m})^T \\ & \Leftrightarrow \min_{\mathbf{y}^i \geq \mathbf{0}, \mathbf{y}^i \mathbf{1}_c = \mathbf{1}_u} (1 + \gamma) \mathbf{y}^i \mathbf{y}_i - 2(\mathbf{x}_i^T \mathbf{P} + \gamma \mathbf{m}) \mathbf{y}_i. \end{aligned} \quad (2.18)$$

Obviously, the above equation implies the following standard quadratic programming problem

$$\min_{\mathbf{v}^T \mathbf{1}_c = \mathbf{1}, \mathbf{v} \geq \mathbf{0}} \mathbf{v}^T \mathbf{A} \mathbf{v} - \mathbf{v}^T \mathbf{b}, \quad (2.19)$$

where  $\mathbf{v} \triangleq \mathbf{y}_i \in \mathbb{R}^{c \times 1}$ ,  $\mathbf{A} \triangleq (1 + \gamma) \mathbf{I}_c \in \mathbb{R}^{c \times c}$ ,  $\mathbf{b} \triangleq 2(\mathbf{P}^T \mathbf{x}_i + \gamma \mathbf{m}) \in \mathbb{R}^{c \times 1}$ . It can be efficiently solved by the method proposed in [20].

To sum up, the optimization procedure is presented in Algorithm 1. Below, we analyze the time complexity of our proposed SRAGL by using the big  $\mathcal{O}$  notation. For each iteration, the time is mainly consumed in updating the graph adjacency matrix  $\mathbf{S}$ , the label matrix of the unlabeled samples  $\mathbf{Y}_u$  and the projection matrix  $\mathbf{W}$ . To be specific, the complexity of updating  $\mathbf{s}^i$  is  $\mathcal{O}(dm)$  and therefore we need  $\mathcal{O}(dmn)$  complexity to update  $\mathbf{S}$ . Then, we need  $\mathcal{O}(n^2d)$  to update  $\mathbf{W}$  and  $\mathcal{O}(uc)$  to obtain  $\mathbf{Y}_u$ . Suppose that the number of iterations is  $t$ , and the total computational complexity of the SRAGL model learning is  $\mathcal{O}(t(n^2d + uc + dmn))$ .

---

**Algorithm 1** The overall procedure of SRAGL
 

---

**Input:** The data matrix  $\mathbf{X} \in \mathbb{R}^{d \times n}$ , the label matrix  $\mathbf{Y}_l \in \mathbb{R}^{l \times c}$ , the regularization parameters  $\alpha, \beta, \gamma, \lambda$ ;

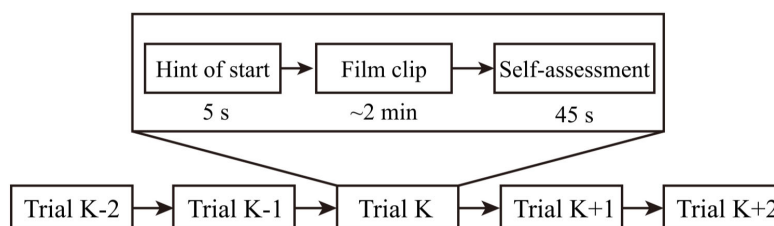
**Output:** The predicted label  $\mathbf{Y}_u \in \mathbb{R}^{u \times c}$ ;

- 1: Initialize the label matrix  $\mathbf{Y}_u = \frac{1}{c} \mathbf{I}_{u \times c}$  and  $\mathbf{Q} \in \mathbb{R}^{d \times d}$  as the identity matrix;
  - 2: Initialize  $\mathbf{S} \in \mathbb{R}^{n \times n}$  and calculate the corresponding Laplace matrix  $\mathbf{L}_S = \mathbf{D} - \mathbf{S} \in \mathbb{R}^{n \times n}$  where  $\mathbf{D}$  is the degree matrix;
  - 3: **while** not converge **do**
  - 4:   Update  $\mathbf{P}$  by updating rule (2.9);
  - 5:   Update  $\mathbf{W}$  by computing the eigenvectors corresponding to the  $m$  smallest eigenvalues of  $\mathbf{X} \mathbf{L}_S \mathbf{X}^T$ ;
  - 6:   For  $i = 1, 2, \dots, n$ , update the  $i$ -th row of  $\mathbf{S}$  by solving equation (2.15);
  - 7:   For  $i = l + 1, l + 2, \dots, n$ , update the  $i$ -th row of  $\mathbf{Y}_u$  by solving the optimization problem (2.19);
  - 8: **end while**
- 

### 3. Experiments

#### 3.1. Data Set

SEED-IV [21] is an emotional EEG data set that was collected in response to 72 movie clips, each of which intended to induce one of four emotional states (*i.e.*, happiness, sadness, fear, and neutrality). This data set was developed by Shanghai Jiao Tong University, and each subject participated in three sessions. In total, there were 15 subjects who contributed EEG data to this data set. During each session, the participants watched a total of 24 movie clips. Within each session, 6 of the 24 clips were intended to induce a particular emotional state. In Figure 2, it describes the procedure of each trial. EEG signals were collected using a 62-channel ESI NeuroScan System while the subjects were watching the movie clips. Then, the raw EEG data were downsampled to 200 Hz and processed by an 1-50 Hz bandpass filter. In their experiments, differential entropy features, extracted from five frequency bands, *Delta* (1-3 Hz), *Theta* (4-7 Hz), *Alpha* (8-13 Hz), *Beta* (14-30 Hz) and *Gamma* (31-50 Hz), were used for emotion recognition. Previous works have shown that differential entropy features perform well for emotion recognition task [22, 23]. Therefore, the dimensionality of the extracted EEG signal features is 310 (*i.e.*, 62 channels multiplied by 5 frequency bands). Due to the time durations of the movie clips in each session are slightly different, the numbers of EEG samples in the three sessions are accordingly different. To be specific, there are 851, 832 and 822 EEG samples in these three sessions.



**Figure 2.** The experimental setup of each trial in SEED-IV [21].



### 3.2. Experimental Setup

In this section, we compare the performance of SRAGL with some related learning models on the SEED-IV data set, which are listed below.

- 1) Feature Selection with Orthogonal Regression (FSOR) [24], which can be used for high-dimensional data feature selection in supervised learning tasks.
- 2) Discriminative LSR (DLSR) [25] which introduces the  $\epsilon$ -dragging technique into LSR and obtains stronger discriminative power and generalization ability.
- 3) Efficient Anchor Graph Regularization (EAGR) [26], which optimizes the performance of anchor graph by introducing anchor points, and leads to more accurate classification results.
- 4) Rescaled Linear Square Regression (RLSR) [27], which proposes an adaptive weight assignment strategy to quantitatively learn the contributions of different features and explicitly illustrates the underlying rationale of applying  $\ell_{2,1}$ -norm for feature selection.
- 5) RLSR with Orthogonal constraint (ORLSR), which adds orthogonal constraints onto the projection matrix in RLSR in order to better maintain the structure information of data.
- 6) Sparse Discriminative Semi-supervised Feature Selection (SDSSFS) [28], which can be considered as the extension of both RLSR and DLSR. That is, the  $\epsilon$ -dragging technique is extended into semi-supervised paradigm on the basis of the joint emotional state estimation of unlabeled EEG samples.
- 7) Semi-supervised Feature Selection via Adaptive Structure Learning and Constrained Graph Learning (ASLCGLFS) [29], which is a semi-supervised feature selection method based on adaptive structure learning and constraint graph learning. Its central idea is to explore the discriminative information of labeled data and the structure information of unlabeled data to improve the learning performance.

FSOR has no parameters to be adjusted. For DLSR, RLSR, SDSSFS and ORLSR, the involved regularization parameter is adjusted in  $\{2^{-20}, 2^{-19}, \dots, 2^{20}\}$ . For EAGR, the parameters  $(\lambda, \gamma)$  are tuned from  $\{10^{-3}, 10^{-2}, \dots, 10^3\}$  and anchor number is adjusted in  $\{10, 20, \dots, 100\}$ . For ASLCGLFS, the parameters  $(\alpha, \lambda, \beta)$  are tuned from  $\{10^{-3}, 10^{-2}, \dots, 10^3\}$ . In SRAGL, we have four parameters, *i.e.*,  $\lambda, \alpha, \beta, \gamma$ , to be tuned from  $\{10^{-3}, 10^{-2}, \dots, 10^3\}$  and the subspace dimension  $m$  is adjusted in  $\{30, 50, 70, 100\}$ .

In our experiment, we perform the cross-session EEG emotion recognition. More specifically, for ‘session2→session3’ classification task, the EEG samples from the third session are unlabeled while those from the second session are labeled. Then, we aim to estimate the emotional states of these unlabeled EEG samples as accurately as possible, within the transductive semi-supervised learning framework. For a certain sample, if it is from the unlabeled session, the elements in its label indicator vector (*i.e.*, the corresponding row in  $\mathbf{Y}_u$ ) are all initialized to  $1/c$ , meaning that we have no prior knowledge to its emotional state. When the criterion  $\frac{\|\text{obj}^{(i+1)} - \text{obj}^{(i)}\|_2}{\|\text{obj}^{(i)}\|_2} \leq 1e - 4$  is satisfied or the maximum number of iterations (*i.e.*, 30) is reached, we terminate the iteration process. Here,  $\text{obj}^{(i)}$  indicates the objective function value in the  $i$ -th iteration. It should be noted that the value of 30 for maximum iterations is typically much larger than the actual number of iterations required to achieve convergence with the algorithm in this experiment.

### 3.3. Recognition Results and Analysis

In Tables 1, 2, and 3, the recognition results of all comparison models and our SRAGL model for cross-session EEG emotion classification tasks are shown, where the bolded result(s) in each row indicate the best accuracy among all the compared models. By comparing and analyzing the recognition accuracy of these eight models, the following points can be realized.

**Table 1.** Comparison of the recognition results (%) on the ‘session1 → session2’ task.

subject	a1	a2	a3	a4	a5	a6	a7	a8
sub1	44.71	55.33	46.15	55.65	58.05	53.61	64.18	<b>65.26</b>
sub2	82.81	89.06	84.86	89.18	90.14	91.23	<b>91.71</b>	88.94
sub3	71.51	61.06	66.59	69.71	71.63	67.79	78.73	<b>79.81</b>
sub4	60.22	61.90	70.55	68.39	62.62	70.55	<b>77.52</b>	71.27
sub5	62.14	67.79	60.70	67.67	67.07	68.39	65.02	<b>68.99</b>
sub6	77.52	70.55	64.42	71.03	<b>86.66</b>	70.55	81.25	81.37
sub7	74.40	75.48	72.60	80.77	78.73	80.77	<b>91.59</b>	82.93
sub8	71.88	62.86	73.92	69.95	74.28	70.43	75.24	<b>81.13</b>
sub9	66.83	67.07	77.28	78.73	74.04	79.09	72.00	<b>82.21</b>
sub10	54.81	68.27	65.75	53.85	62.38	55.89	69.95	<b>73.44</b>
sub11	64.18	49.28	51.21	52.04	65.38	52.04	56.61	<b>69.59</b>
sub12	63.58	55.17	67.43	53.13	74.16	70.67	65.99	<b>74.16</b>
sub13	65.99	67.55	66.83	68.63	71.88	69.59	66.47	<b>72.60</b>
sub14	76.92	68.15	74.76	76.92	76.68	74.76	68.15	<b>83.65</b>
sub15	94.23	93.63	91.71	87.14	97.36	97.36	96.23	<b>97.36</b>
Avg.	68.78	67.54	69.18	69.52	74.07	71.51	74.71	<b>78.18</b>

a1: FSOR, a2: DLSR, a3: EAGR, a4: RLSR, a5: ORLSR, a6: SDSSFS, a7: ASLCGLFS, a8: SRAGL.

- 1) The proposed SRAGL model achieves the highest recognition accuracy on most subjects in the cross-session emotion recognition tasks, and the average recognition accuracies of all the three cross-session tasks are higher than the remaining models, *i.e.*, 78.18%, 80.55%, and 81.99%, respectively. Compared to the second best model ASLCGLFS, SRAGL improves by 3.47%, 6.63%, 3.85% in the three cross-session tasks. Overall, the proposed SRAGL model is verified to be effective and the adaptive graph learning-based similarity matrix facilitates the improvement of emotion recognition performance.
- 2) Though EAGR, ASLCGLFS, and SRAGL are all graph-based models, the recognition results of EAGR are generally worse than those obtained by ASLCGLFS and SRAGL. To be specific, EAGR only obtained 69.18%, 69.47%, and 72.58% average recognition accuracies in three cross-session tasks. The reason accounting for it is that the similarity matrix constructed by EAGR is not powerful enough to depict the desirable data connections. EAGR selects a few representative samples, and then uses them to represent all sample points. Though the computational complexity of this anchor point based model is greatly reduced, recognition results and the quality of the

**Table 2.** Comparison of the recognition results (%) on the ‘session1 → session3’ task.

subject	a1	a2	a3	a4	a5	a6	a7	a8
sub1	<b>75.55</b>	74.45	68.37	70.68	74.21	70.19	64.72	73.48
sub2	79.93	79.68	83.33	89.29	87.47	84.43	83.33	<b>93.43</b>
sub3	40.88	43.55	60.10	48.78	53.89	45.50	59.73	<b>67.03</b>
sub4	61.68	65.69	64.72	71.17	69.59	79.32	<b>80.29</b>	80.05
sub5	68.00	68.73	70.92	58.39	72.75	74.57	<b>75.43</b>	65.21
sub6	79.08	79.56	69.46	<b>83.45</b>	76.03	82.24	77.98	81.39
sub7	89.05	88.69	79.56	88.44	91.12	84.91	87.10	<b>96.35</b>
sub8	<b>88.20</b>	80.17	78.47	80.78	82.36	79.68	76.28	85.04
sub9	50.73	54.74	58.76	62.77	65.57	63.50	68.00	<b>79.56</b>
sub10	58.27	48.91	63.75	49.64	68.37	49.64	64.23	<b>75.30</b>
sub11	78.35	71.90	62.65	71.17	73.36	70.68	69.22	<b>80.66</b>
sub12	54.01	57.06	56.57	65.45	68.13	66.79	66.67	<b>74.94</b>
sub13	58.64	54.99	60.58	62.41	75.91	61.31	73.11	<b>73.11</b>
sub14	85.40	80.17	83.21	82.85	87.59	82.73	78.22	<b>89.66</b>
sub15	86.37	83.82	81.63	85.40	86.37	85.40	84.43	<b>93.07</b>
Avg.	70.28	68.81	69.47	71.38	75.51	72.06	73.92	<b>80.55</b>

**Table 3.** Comparison of the recognition results (%) on the ‘session2 → session3’ task.

subject	a1	a2	a3	a4	a5	a6	a7	a8
sub1	57.79	61.92	67.52	62.65	68.86	61.68	64.60	<b>72.63</b>
sub2	88.56	85.59	86.98	83.21	<b>90.63</b>	86.74	86.13	88.81
sub3	70.80	67.03	65.94	67.27	68.98	68.98	71.78	<b>76.64</b>
sub4	80.78	75.91	79.08	80.17	77.49	<b>89.54</b>	78.83	83.45
sub5	83.70	76.76	83.45	72.87	<b>83.58</b>	77.98	80.54	83.21
sub6	78.95	78.35	66.55	83.70	85.16	<b>92.70</b>	83.33	<b>92.70</b>
sub7	81.87	81.14	80.05	88.56	81.14	89.78	89.05	<b>90.02</b>
sub8	73.48	74.70	77.37	82.97	79.56	80.41	76.52	<b>84.06</b>
sub9	59.85	52.07	59.73	61.80	74.45	58.39	69.83	<b>77.25</b>
sub10	70.68	67.03	66.18	78.35	69.71	78.35	74.57	<b>78.59</b>
sub11	58.88	56.20	67.64	59.49	66.67	73.60	<b>76.16</b>	70.68
sub12	78.71	66.30	57.54	63.63	74.57	70.32	<b>80.17</b>	77.86
sub13	54.14	51.70	62.41	64.48	68.00	62.17	61.80	<b>72.63</b>
sub14	84.67	86.86	82.85	87.10	90.15	87.10	89.17	<b>91.00</b>
sub15	83.94	85.64	85.40	89.90	89.54	<b>93.07</b>	89.90	90.27
Avg.	73.79	71.17	72.58	75.08	77.90	78.05	78.16	<b>81.99</b>

constructed graph depend severely on the selected anchor points. In contrast, the graph learning in ASLCGLFS and SRAGL is by projecting the original data into the subspace to suppress the noisy and unwanted data components, and then constructing the graph similarity matrix based on the distance of the projected data.

- 3) The remaining comparison models can be considered as variations of LSR. Among them, RLSR outperforms DLSR and FSOR due to its ability of adaptively learning the significance of different features through a feature self-weighting variable. By comparing the recognition results obtained by RLSR, ORLSR, and SDSSFS, the latter two models are superior to RLSR. ORLSR additionally includes the orthogonal constraint to the transformation matrix in RLSR, which on one hand can better remove the data redundancy information and on the other hand can retain the EEG data metric. SDSSFS employs the label dragging strategy to correct the labels of the originally misclassified samples, thus improving the recognition accuracy.

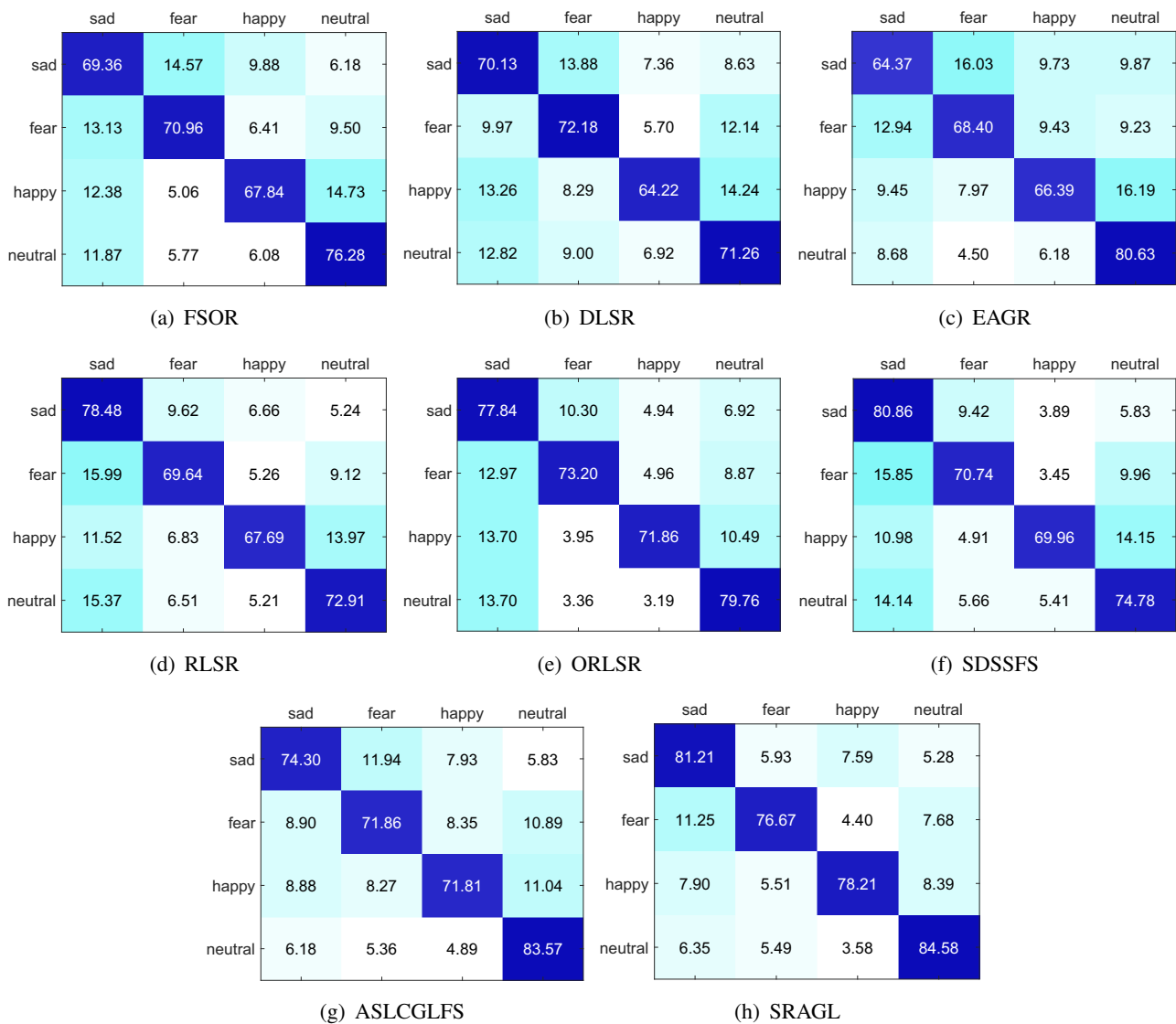
In Table 4, it shows the analysis of the variance (ANOVA) between SRAGL and each of the other models [30]. The ANOVA is a popular statistical method used to determine whether there are significant differences between two or more groups. In our experiment, we adopt a standard one-way ANOVA. Specifically, the performance accuracy is the dependent variable and the algorithm is the independent variable. The null hypothesis of ANOVA is that there is no difference between the group means. When performing ANOVA, it is usually necessary to set a significance level  $\alpha$ , which means the probability of incorrectly rejecting the null hypothesis. If  $\alpha$  is set at 0.05 and p-value is less than 0.05, it means that there are differences between the models. If  $\alpha$  is set to 0.01 and p-value is less than 0.01, the models are considered to be significantly different from each other. From Table 4, F-value and p-value are got by comparing the recognition results obtained by SRAGL (i.e., a8) and the other seven methods (i.e., a1-a7) respectively. These results demonstrate that SRAGL statistically outperforms the other seven models in terms of recognition accuracy.

**Table 4.** The analysis of the variance (\*\*p-value < 0.01, \*p-value < 0.05).

ANOVA	a1	a2	a3	a4	a5	a6	a7
F-value	16.41	24.42	24.94	14.09	5.24	8.07	5.98
p-value	1.095e-04**	3.67e-06**	2.97e-06**	3.12e-04**	0.025*	0.0056**	0.017*

In Figure 3, the confusion matrix is used to represent the accuracy of each model for four emotion states. From the figure, we can obtain more information from the recognition results, which includes 1) the eight models' average recognition accuracy in each emotion state. For example, the average recognition accuracy for the neutral emotional state using SRAGL was 84.58%, which was the highest recognition rate among the four emotional states. Conversely, the fear state had the lowest average recognition accuracy at 76.67%. 2) the accuracy improvement of SRAGL compared to the other models in each state. For example, SRAGL improved ASLCGLFS by 6.91%, 4.81%, 6.40%, and 1.01% for identifying the four states. 3) the proportion of each state that is misclassified into the other states. For example, SRAGL misclassified the EEG samples belonging to the fear state as sad, happy and neutral with 11.25%, 4.40% and 7.68%, respectively.

In SRAGL, we learn an optimal graph according to iteratively updating the graph adjacency matrix which describes the relationship between the subspace represented EEG samples. Based on the struc-

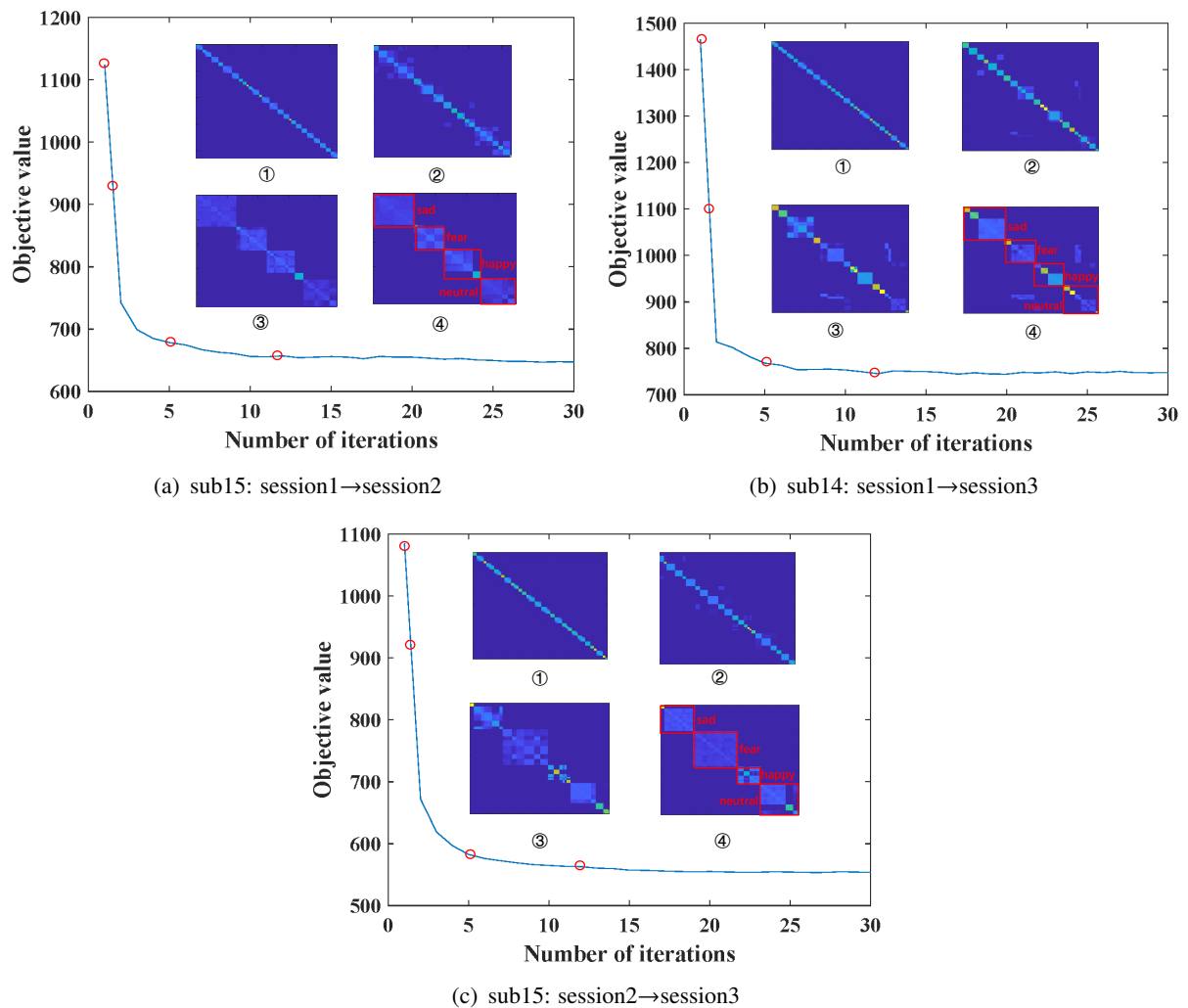


**Figure 3.** The recognition results are organized by confusion matrices, in which both the horizontal and vertical axes correspond to the four emotional states.

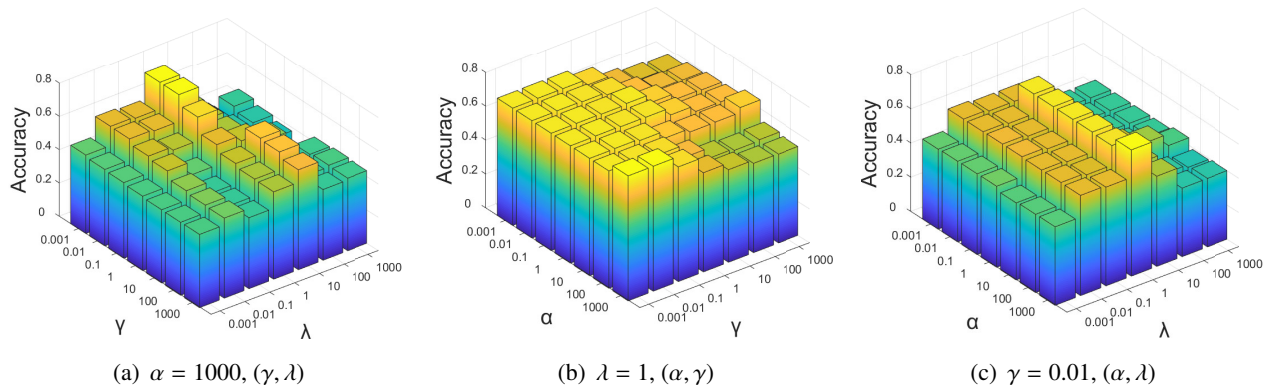
tered graph learning theory [31, 32], the number of connected sub-components (*i.e.*, diagonal blocks in visualized graph similarity matrix) in an ideal graph should be equal to the number of classes in the data set. In the SEED-IV data set, there are four emotional states (*i.e.*,  $c = 4$ ); therefore, the ideal graph adjacency matrix should have four diagonal blocks, respectively corresponding to the sad, fear, happy, and neutral states. In Figure 4, we visualize the learned graphs in terms of different iterations, *i.e.*, when the numbers of iterations are 1, 2, 5 and 12, respectively. From figures, we can see that as SRAGL iterates, the learned graph for the SEED-IV data set has four diagonal blocks, each of which corresponds to one of the four emotional states. This result demonstrates that the adjacency matrix of the learned graph can represent the relationship between EEG samples effectively. Take Figure 4(a) as an example, we can clearly observe that the four diagonal blocks gradually become more visible as the number of iterations increases, which means the connection values within the block become larger while the connection values between the blocks become smaller. In other words, EEG samples belonging to the same emotional state should be connected to each other while EEG samples belonging to different states should be unconnected. Based on the learned graph adjacency matrix, the recognition accuracy of this case is 97.36%. Similarly, other cases also illustrate the above facts. However, since the EEG samples have inter-session variabilities to some extent, it is not guaranteed that the learned graph adjacency matrix has the exact four diagonal blocks in all cases. In addition, this figure shows us that the objective function of SRAGL exhibits good convergence, with the objective function value decreasing quickly as the number of iterations increases.

Below, we show how the emotion recognition performance of SRAGL is influenced by the regularization parameters, *i.e.*,  $\lambda, \beta, \alpha, \gamma$ , by some example cases. In SRAGL, the parameters  $\alpha, \beta$  correspond to the graph adjacency matrix learning,  $\gamma$  corresponds to graph label propagation and  $\lambda$  controls the sparsity of the transformation matrix  $\mathbf{P}$ . It is found experimentally that the performance of SRAGL is insensitive to parameter  $\beta$ . In Figure 5, we fix  $\beta$  as 1 and take the case of 'sub1: session1→session3' as an example to investigate how the recognition results of SRAGL varies in response to different settings of  $\alpha, \lambda$ , and  $\gamma$ . From this figure, we observe that the parameter  $\lambda$  is the most sensitive one to SRAGL. In model training, we can set  $\lambda$  within  $\{0.1, 1, 10\}$  to make SRAGL achieve good recognition results. In contrast, SRAGL is not very sensitive to the parameters  $\alpha$  and  $\gamma$  due to  $\alpha$  mainly affects the learning of the graph adjacency matrix and thus implicitly influences the label matrix of the unlabeled EEG samples.

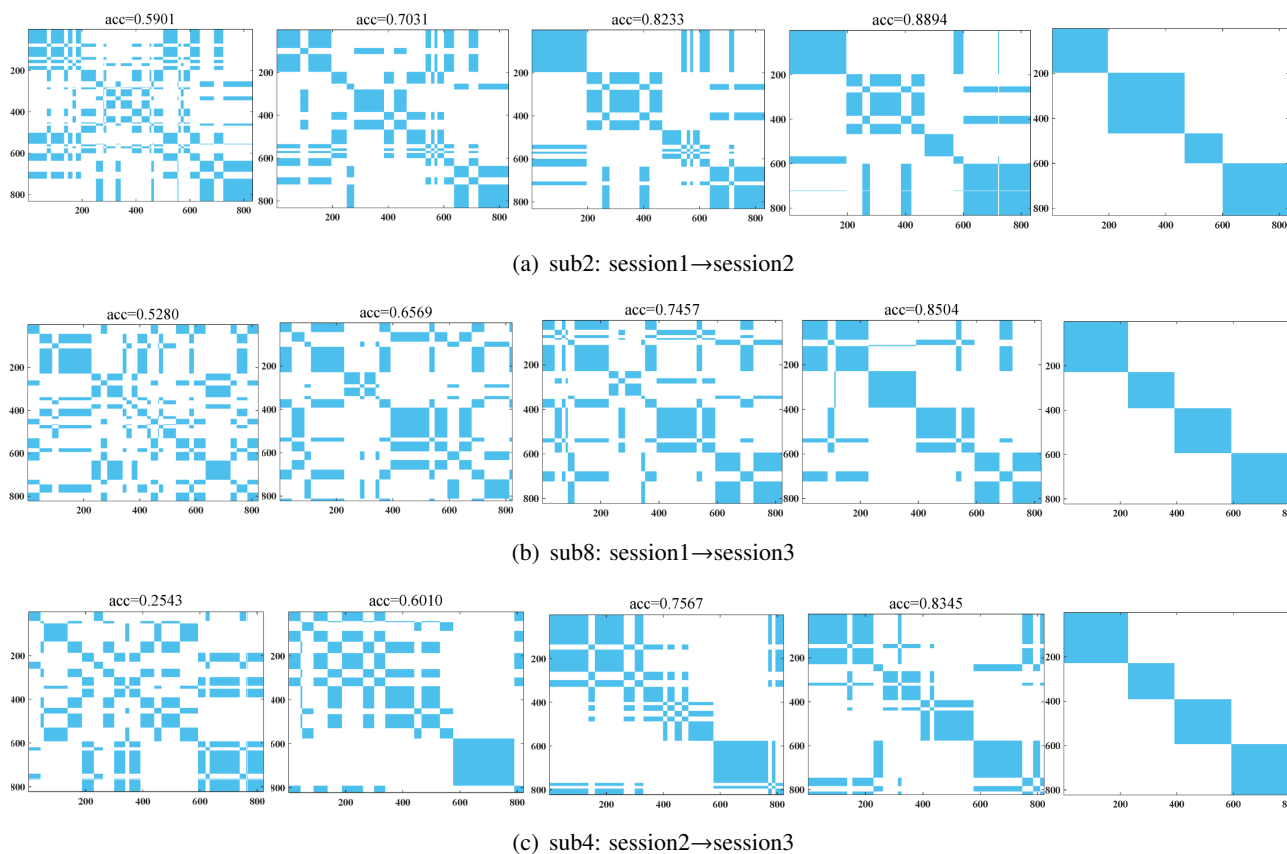
In Figure 6, three cases were used to show the iterative optimization process of SRAGL according to the learned label indicator matrix of unlabeled EEG samples, which was proposed in [33]. More specifically, with the number of iterations increasing, we reconstruct the graph adjacency matrix at different iterations by  $\mathbf{Y}_u \mathbf{Y}_u^T$ . From this figure, we observe that the correlations for the intra-class EEG samples are increasing over time, whereas the relationships for inter-class EEG samples are gradually decreasing. This is consistent with our expectation that SRAGL gradually depicts the underlying semantic information of EEG samples whilst suppressing the interference of different sessions and trials. In addition, the more accurate the label matrix of the unlabeled EEG samples was predicted, the clearer the four diagonal blocks in the reconstructed graph become. Taking Figure 6(a) 'sub2: session1→session2' as an example, four diagonal blocks correspond to four emotion states in the SEED-IV data set become more and more visible as the accuracy increases which illustrates the effectiveness of the SRAGL model optimization. Meanwhile, we show the ground truth graph as a comparison in the last column.



**Figure 4.** Three example cross-session emotion recognition cases were used to illustrate the graph learning process in SRAGL. As the number of iterations increases, the connection weights become larger for the samples belonging to the same emotional state and smaller for the ones from different states. Equivalently, the diagonal blocks corresponding to the four emotional states become more and more apparent.



**Figure 5.** Examples to show the parameter sensitivity.

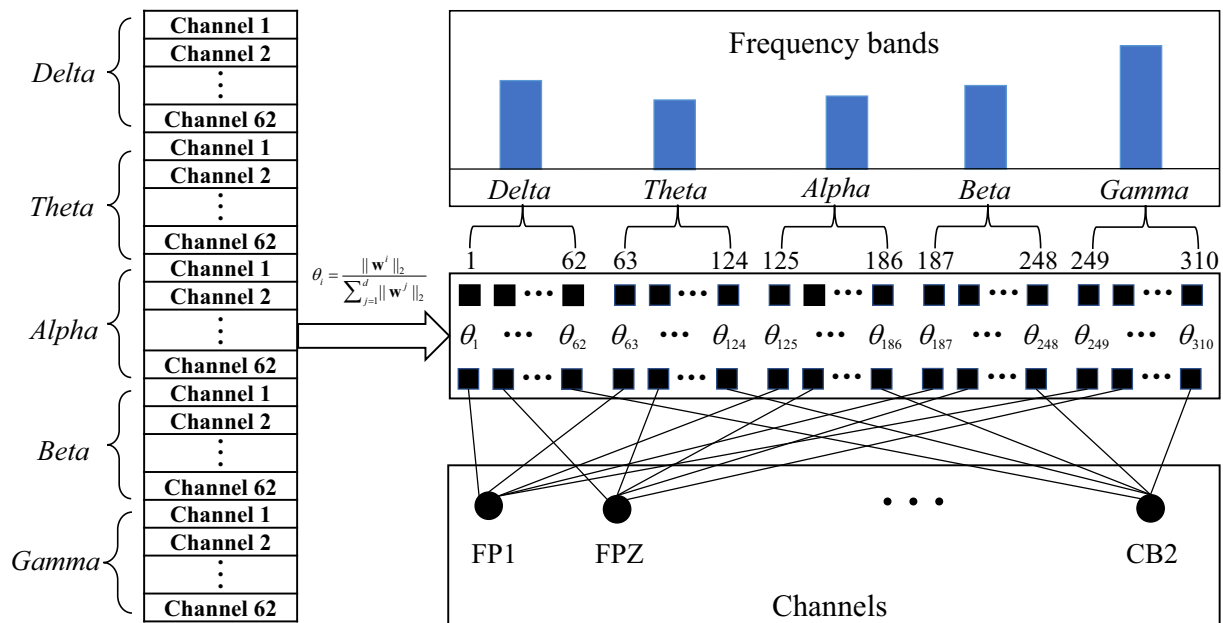


**Figure 6.** The iterative graph reconstruction process by the label matrix predicted by SRAGL.



### 3.4. EEG Emotional Activation Pattern Mining

EEG data is typically characterized as multi-frequency and multi-channel signals, and DE features are extracted from different frequency bands and channels. In this section, we quantify the contributions of EEG features to cross-session emotion recognition tasks to investigate key frequency bands and channels. As shown in Figure 7, the coupling relationship between each EEG feature dimension and each frequency band (channel) is illustrated.



**Figure 7.** Framework of emotional activation patterns analysis.

Inspired by the  $\ell_{2,1}$ -norm based feature ranking theory [34, 27], we calculate the normalized  $\ell_2$ -norm of each row of the projection matrix  $\mathbf{W}$  to determine the contribution of each EEG feature (*i.e.*,  $\theta_i$ ). In mathematics, it is expressed by

$$\theta_i = \frac{\|\mathbf{w}^i\|_2}{\sum_{j=1}^d \|\mathbf{w}^j\|_2}. \quad (3.1)$$

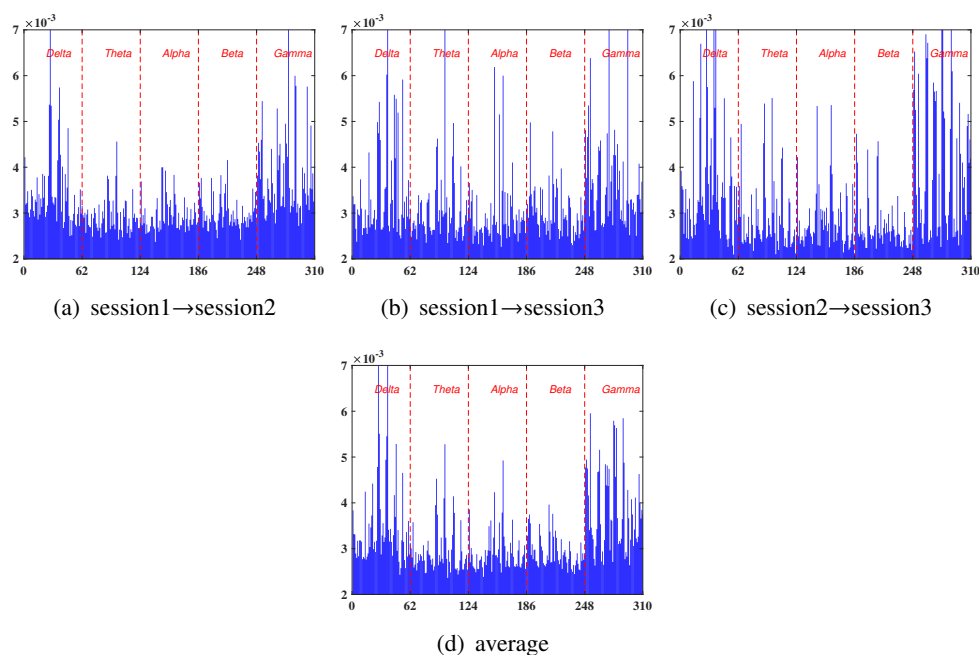
It is obvious that the larger  $\theta_i$  is, the greater contribution of the  $i$ -th EEG feature to the emotion classification tasks. For each subject, we calculate the contribution of each feature dimension. Due to the inter-subject variability, there may be variations for the learned feature importance values of different subjects. To mitigate the impact of individual variability and improve the reliability of the identified important features by SRAGL, we average the feature contributions across all the subjects. As a result, we obtain the contributions of the 310 feature dimensions, shown in Figure 8. Suppose that we have  $a$  EEG bands and  $b$  channels in total, based on the established relationship between EEG features and their correlated frequency band (channel) [35], we can calculate the contribution of the  $i$ -th frequency band by

$$p_i = \theta_{(i-1)*a+1} + \theta_{(i-1)*a+2} + \dots + \theta_{i*a}. \quad (3.2)$$

Similarly, the contribution of the  $j$ -th EEG channel is

$$q_j = \theta_j + \theta_{j+b} + \dots + \theta_{j+(a-1)*b}. \quad (3.3)$$

In SEED-IV, the feature dimension is 310, the number of frequency bands is 5 (*i.e.*,  $a=5$ ) and the number of EEG channel is 62 (*i.e.*,  $b=62$ ). According to equation (3.2), Figure 9 displays the relative importance of different frequency bands, represented as bar charts. From this figure, we observe that *Gamma* is the most significant frequency band for cross-session EEG emotion recognition, while the second one is *Delta*. Similarly, according to equation (3.3), we obtained the contribution values of the 62 EEG channels and represented their relative importance using brain topography maps in Figure 10. As a result, we conclude that prefrontal, left/right temporal, especially the (central) parietal lobes contributed more to emotion recognition, which is consistent with the existing studies [21]. In Figure 11, we show the 10 most important channels in cross-session emotion recognition, which are CZ, CPZ, CP2, P2, C2, FP1, CP1, POZ, FP2, and FC6. The emotional activation pattern analysis method utilized above can be extended to any EEG data set, regardless of the number of bands or channels.

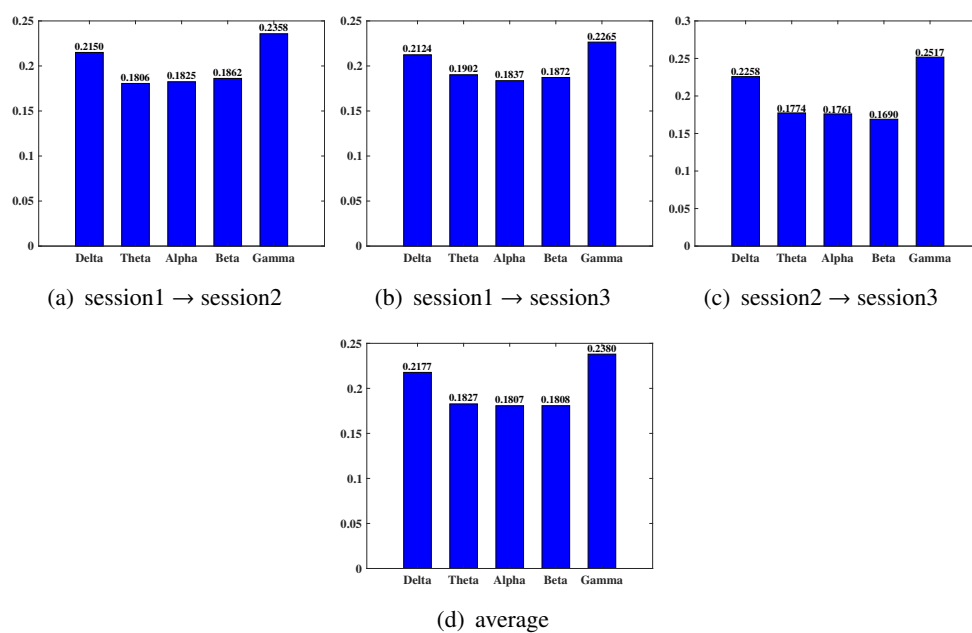


**Figure 8.** The learned feature importance of the 310 EEG feature dimensions.

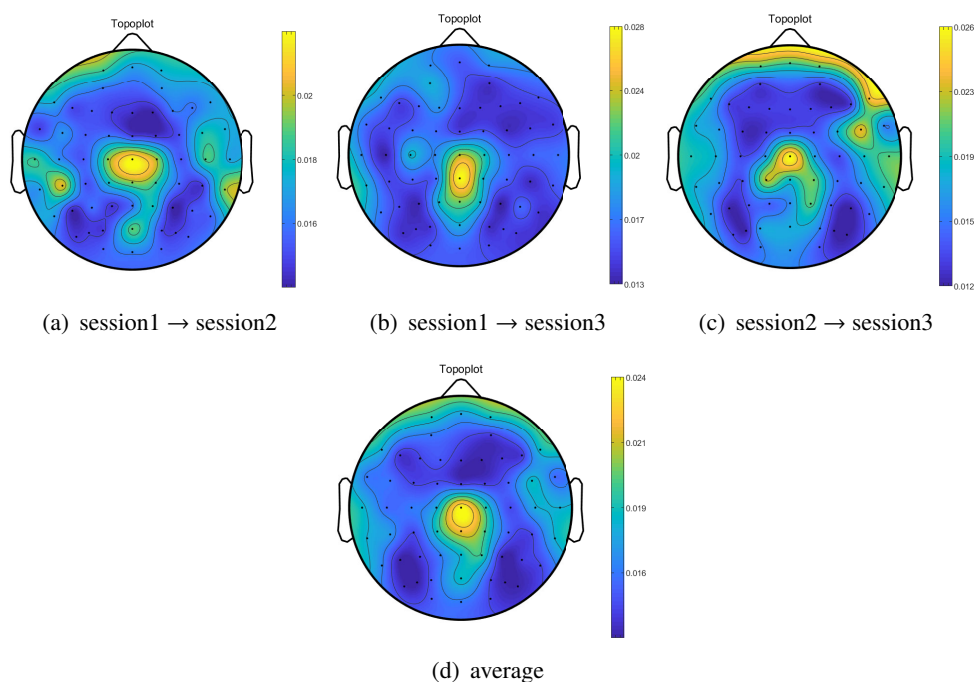
### 3.5. Discussion

To highlight the significance of this work, this section provides thorough discussions by putting the emphasis on the connections and differences between SRAGL and some of our previous studies including the Self-weighted Semi-supervised Classification (SWSC) [36], the retargeted semi-supervised regression with robust weights (RSRRW) [37], the semi-supervised Joint Sample and Feature importance Evaluation (sJSFE) [38], the Joint label-Common and label-Specific Features Exploration (JCSFE) [39], and the Optimal Graph coupled Semi-Supervised Learning (OGSSL) [40]. The connections are first analyzed, followed by the differences.

Generally, all of these models are proposed for EEG-based emotion recognition under the semi-supervised classification paradigm. On the experimental setup, EEG data from different sessions was utilized in the comparative studies for their performance evaluation, *i.e.*, cross-session emotion recogni-

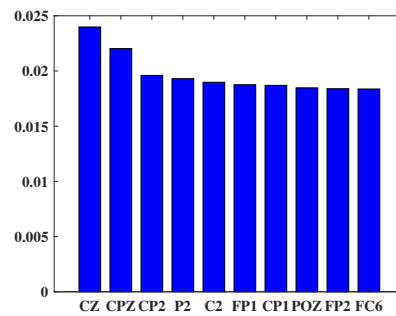


**Figure 9.** The importance analysis of the five frequency bands.



**Figure 10.** The importance of the 62 EEG channels using brain topography maps.

tion. Besides, the important EEG frequency bands and channels in emotion recognition were identified according to the explicit/implicit feature importance exploration and the correspondence between EEG features and the associated frequency bands (channels).



**Figure 11.** Top 10 important EEG channels.

On the differences between SRAGL and the previously proposed ones, we have the following two understandings.

- 1) From the perspective of model formulation, SWSC is motivated by explicitly introducing a feature weighting variable to characterize the different contributions of different features in emotion recognition while JCSFE is a more problem-driven model and puts more focus on the common activation patterns across different emotional states and the specific patterns associated with a certain state. Both RSRRW and sJSFE mainly pay attention to the model robustness by respectively introducing discrete and continuous weights to characterize the different contributions of different samples in emotion recognition, which do not involve the graph data structure in their model formulations. Though graph regularization terms were used in both SWSC and JCSFE, the graph similarity matrices were directly calculated according to fixed rules such as the ‘0-1’ and ‘heatkernel’ weighting schemes. In SRAGL, the graph similarity matrix is dynamically learned within the iterative model optimization process, which aims to enforce the label consistency of unlabeled EEG samples. In the experiments, the effectiveness of the graph learning technique in SRAGL is intuitively analyzed from two aspects. On one aspect, we showed how the learned graph similarity matrix depicts the semantic information of EEG samples; that is, each of four diagonal blocks in Figure 4 corresponds to one of the four emotional states. On the other aspect, we showed how the learned graph facilitates the estimation of the label indication matrix of the unlabeled samples in Figure 6.
- 2) Both OGSSL and SRAGL involve the dynamic graph learning strategy in their model objective functions. The main difference between them is how the label indicator matrix of unlabeled samples is estimated. In OGSSL, the estimation of the label indicator matrix corresponds to a label-propagation process over the learned graph. However, in SRAGL, the label matrix is obtained by learning a projection matrix to bridge the data space and the label space. Obviously, different optimization strategies were employed in respective models to handle the updating of the label indicator matrix. In terms of the recognition accuracy, the proposed SRAGL model in the present work improves OGSSL by 1.67%, 3.47%, and 0.70% corresponding to three cross-session tasks respectively.

#### 4. Conclusions

In this paper, a semi-supervised regression with adaptive graph learning model is proposed for EEG emotion recognition, which is termed SRAGL. In SRAGL, semi-supervised regression, adaptive graph

structure learning and label propagation are seamlessly integrated together. SRAGL updates the graph similarity matrix gradually by leveraging label information from both labeled and unlabeled samples, which is jointly estimated by LSR, ultimately resulting in the optimal structured graph. Extensive results on SEED-IV indicated that the learned graph adjacency matrix effectively improves the performance of emotion recognition. Furthermore, by the established relationship between EEG features and their correlated frequency band (channel), and utilizing the learned projection matrix, we calculated the contribution of each EEG feature and analyzed the emotional activation patterns. Our findings using SRAGL suggest that the *Gamma* band is the most significant one, and that the central parietal lobe is more closely related to emotional expression. Based on these findings, it is expected to customize the hardware and computing models for emotion recognition applications in the future. For example, we can minimize the number of electrodes required in wearable devices such as caps for EEG data collection. In addition, we explore the possibility of utilizing only *Gamma* band to reduce computational requirements. These conclusions shed new light on emotional processing for human-computer interaction and provide valuable insights for future research in this field.

### Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (No. 61971173) and the Natural Science Foundation of Zhejiang Province (No. LY21F030005).

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. R. Adolphs, D. J. Anderson, The neuroscience of emotion, In *The Neuroscience of Emotion*, Princeton University Press, 2018. <https://doi.org/10.23943/9781400889914>
2. Z. Halim, M. Rehan, On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning, *Inform. Fusion*, **53** (2020), 66–79. <https://doi.org/10.1016/j.inffus.2019.06.006>
3. H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, B. Hu, Feature-level fusion approaches based on multimodal EEG data for depression recognition, *Inform. Fusion*, **59** (2020), 2127–138. <https://doi.org/10.1016/j.inffus.2020.01.008>
4. D. Xu, X. Qin, X. Dong, X. Cui, Emotion recognition of EEG signals based on variational mode decomposition and weighted cascade forest, *Math. Biosci. Eng.*, **20** (2023), 2566–2587. <https://doi.org/10.3934/mbe.2023120>
5. J. Xue, J. Wang, S. Hu, N. Bi, Z. Lv, OVPD: odor-video elicited physiological signal database for emotion recognition, *IEEE Trans. Instrum. Meas.*, **71** (2022), 1–12. <https://doi.org/10.1109/TIM.2022.3149116>
6. N. Suhaimi, J. Mountstephens, J. Teo, EEG-based emotion recognition: A state-of-the-art review of current trends and opportunities, *Comput. Intel. Neurosc.*, **2020** (2020), 1–19. <https://doi.org/10.1155/2020/8875426>

7. Y. Ou, S. Sun, H. Gan, R. Zhou, Z. Yang, An improved self-supervised learning for EEG classification, *Math. Biosci. Eng.*, **19** (2022), 6907–6922. <https://doi.org/10.3934/mbe.2022325>
8. X. Li, D. Song, P. Zhang, Y. Zhang, Y. Hou, B. Hu, Exploring EEG features in cross-subject emotion recognition, *Front. Neurosci.*, **12** (2018), 162. <https://doi.org/10.3389/fnins.2018.00162>
9. Y. Dan, J. Tao, J. Fu, D. Zhou, Possibilistic clustering-promoting semi-supervised learning for EEG-based emotion recognition, *Front. Neurosci.*, **15** (2021), 690044. <https://doi.org/10.3389/fnins.2021.690044>
10. X. Chen, L. Song, Y. Hou, G. Shao, Efficient semi-supervised feature selection for VHR remote sensing images, In *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, (2016), 1500–1503. <https://doi.org/10.1109/IGARSS.2016.7729383>
11. B. Tang, L. Zhang, Local preserving logistic i-relief for semi-supervised feature selection, *Neurocomputing*, **399** (2020), 48–64. <https://doi.org/10.1016/j.neucom.2020.02.098>
12. H. Gan, Z. Li, W. Wu, Z. Luo, R. Huang, Safety-aware graph-based semi-supervised learning, *Expert Syst. Appl.*, **107** (2018), 243–254. <https://doi.org/10.1016/j.eswa.2018.04.031>
13. Y. Peng, W. Kong, F. Qin, F. Nie, Manifold adaptive kernelized low-rank representation for semisupervised image classification, *Complexity*, **2018** (2018), 1–12. <https://doi.org/10.1155/2018/2857594>
14. Z. Kang, Z. Lin, X. Zhu, W. Xu, Structured graph learning for scalable subspace clustering: From single view to multiview, *IEEE Trans. Cybern.*, **52** (2021), 8976–8986. <https://doi.org/10.1109/TCYB.2021.3061660>
15. Z. Liu, Z. Lai, W. Ou, K. Zhang, R. Zheng, Structured optimal graph based sparse feature extraction for semi-supervised learning, *Signal Process.*, **170** (2020), 107456. <https://doi.org/10.1016/j.sigpro.2020.107456>
16. Z. Lin, Z. Kang, L. Zhang, L. Tian, Multi-view attributed graph clustering, *IEEE Trans. Knowl. Data Eng.*, **35** (2021), 1872–1880. <https://doi.org/10.1016/10.1109/TKDE.2021.3101227>
17. F. Nie, Z. Wang, R. Wang, X. Li, Adaptive local embedding learning for semi-supervised dimensionality reduction, *IEEE Trans. Knowl. Data En.*, **34** (2021), 4609–4621. <https://doi.org/10.1109/TKDE.2021.3049371>
18. X. Chen, R. Chen, Q. Wu, F. Nie, M. Yang, R. Mao, Semisupervised feature selection via structured manifold learning, *IEEE Trans. Cybern.*, **52** (2021), 5756–5766. <https://doi.org/10.1109/TCYB.2021.3052847>
19. G. Haeser, M. Schuverdt, On approximate KKT condition and its extension to continuous variational inequalities, *J. Optimiz. Theory App.*, **149** (2011), 528–539. <https://doi.org/10.1007/s10957-011-9802-x>
20. Y. Peng, X. Zhu, F. Nie, W. Kong, Y. Ge, Fuzzy graph clustering, *Inf. Sci.*, **571** (2021), 38–49. <https://doi.org/10.1016/j.ins.2021.04.058>
21. W. Zheng, W. Liu, Y. Lu, B. Lu, A. Cichocki, Emotionmeter: A multimodal framework for recognizing human emotions, *IEEE Trans. Cybern.*, **49** (2019), 1110–1122. <https://doi.org/10.1109/TCYB.2018.2797176>

22. R. Duan, J. Zhu, B. Lu, Differential entropy feature for EEG-based emotion classification, In *Proc. Int. IEEE/EMBS Conf. Neural Eng.*, (2013), 81–84. <https://doi.org/10.1109/NER.2013.6695876>
23. L. Shi, Y. Jiao, B. Lu, Differential entropy feature for EEG-based vigilance estimation, In *Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, (2013), 6627–6630. <https://doi.org/10.1109/EMBC.2013.6611075>
24. X. Wu, X. Xu, J. Liu, H. Wang, B. Hu, F. Nie, Supervised feature selection with orthogonal regression and feature weighting, *IEEE Trans. Neural Netw. Learn. Syst.*, **32** (2020), 1831–1838. <https://doi.org/10.1109/TNNLS.2020.2991336>
25. S. Xiang, F. Nie, G. Meng, C. Pan, C. Zhang, Discriminative least squares regression for multiclass classification and feature selection, *IEEE Trans. Neural Netw. Learn. Syst.*, **23** (2012), 1738–1754. <https://doi.org/10.1109/TNNLS.2012.2212721>
26. M. Wang, W. Fu, S. Hao, D. Tao, X. Wu, Scalable semi-supervised learning by efficient anchor graph regularization, *IEEE Trans. Knowl. Data Eng.*, **28** (2016), 1864–1877. <https://doi.org/10.1109/TKDE.2016.2535367>
27. X. Chen, G. Yuan, F. Nie, J. Huang, Semi-supervised feature selection via rescaled linear regression, In *Proc. Int. J. Conf. Artif. Intell.*, **2017** (2017), 1525–1531. <https://doi.org/10.24963/ijcai.2017/211>
28. C. Wang, X. Chen, G. Yuan, F. Nie, M. Yang, Semi-supervised feature selection with sparse discriminative least squares regression, *IEEE Trans. Cybern.*, **52** (2022), 8413–8424. <https://doi.org/10.1109/TCYB.2021.3060804>
29. J. Lai, H. Chen, W. Li, T. Li, J. Wan, Semi-supervised feature selection via adaptive structure learning and constrained graph learning, *Knowl. Based Syst.*, **251** (2022), 109243. <https://doi.org/10.1016/j.knosys.2022.109243>
30. Z. Ma, Z. Xie, T. Qiu, J. Cheng, Driving event-related potential-based speller by localized posterior activities: An offline study, *Math. Biosci. Eng.*, **17** (2020) 789–801. <https://doi.org/10.3934/mbe.2020041>
31. F. Nie, X. Wang, H. Huang, Clustering and projected clustering with adaptive neighbors, In *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, (2014) 977–986. <https://doi.org/10.1145/2623330.2623726>
32. F. Nie, X. Wang, M. Jordan, H. Huang, The constrained laplacian rank algorithm for graph-based clustering, In *Proc. AAAI Conf. Artif. Intell.*, (2016), 1969–1976. <https://doi.org/10.1609/aaai.v30i1.10302>
33. J. Han, K. Xiong, F. Nie, Orthogonal and nonnegative graph reconstruction for large scale clustering, In *Proc. Int. J. Conf. Artif. Intell.*, (2017), 1809–1815. <https://doi.org/10.24963/ijcai.2017/251>
34. F. Nie, H. Huang, X. Cai, C. Ding, Efficient and robust feature selection via joint  $\ell_{2,1}$ -norms minimization, In *International Conference on Neural Information Processing Systems.*, **23** (2010), 1813–1821. <https://doi.org/10.24963/ijcai.2017/251>
35. Y. Peng, F. Qin, W. Kong, Y. Ge, F. Nie, A. Cichocki, GFIL: A unified framework for the importance analysis of features, frequency bands and channels in EEG-based emotion recognition, *IEEE Trans. Cogn. Develop. Syst.*, **14** (2022), 935–947. <https://doi.org/10.1109/TCDS.2021.3082803>

36. Y. Peng, W. Kong, F. Qin, F. Nie, J. Fang, B. Lu, A. Cichocki, Self-weighted semi-supervised classification for joint EEG-based emotion recognition and affective activation patterns mining, *IEEE Trans. Instrum. Meas.*, **70** (2021), 1–11. <https://doi.org/10.1109/TIM.2021.3124056>
37. Z. Chen, S. Duan, Y. Peng, EEG-based emotion recognition by retargeted semi-supervised regression with robust weights, *Systems*, **10** (2022), 236. <https://doi.org/10.3390/systems10060236>
38. X. Li, F. Shen, Y. Peng, W. Kong, B. Lu, Efficient sample and feature importance mining in semi-supervised EEG emotion recognition, *IEEE Trans. Circuits Syst. II, Exp. Briefs*, **69** (2022), 3349–3353. <https://doi.org/10.1109/TCSII.2022.3163141>
39. Y. Peng, H. Liu, J. Li, J. Huang, B. Lu, W. Kong, Cross-session emotion recognition by joint label-common and label-specific EEG features exploration, *IEEE Trans. Neur. Syst. Reh. Eng.*, **31** (2022), 759–768. <https://doi.org/10.1109/TNSRE.2022.3233109>
40. Y. Peng, F. Jin, W. Kong, F. Nie, B. Lu, A. Cichocki, Ogssl: A semi-supervised classification model coupled with optimal graph learning for EEG emotion recognition, *IEEE Trans. Neur. Syst. Reh. Eng.*, **30** (2022), 1288–1297. <https://doi.org/10.1109/TNSRE.2022.3175464>



AIMS Press

© 2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)