OGSSL: A Semi-Supervised Classification Model Coupled With Optimal Graph Learning for EEG Emotion Recognition

Yong Peng, *Member, IEEE*, Fengzhe Jin, Wanzeng Kong, *Member, IEEE*, Feiping Nie, *Senior Member, IEEE*, Bao-Liang Lu, *Fellow, IEEE*, and Andrzej Cichocki, *Life Fellow, IEEE*

***Abstract*— Electroencephalogram (EEG) signals are generated from central nervous system which are difficult to disguise, leading to its popularity in emotion recognition. Recently, semi-supervised learning exhibits promising emotion recognition performance by involving unlabeled EEG data into model training. However, if we first build a graph to characterize the sample similarities and then perform label propagation on this graph, these two steps cannot well collaborate with each other. In this paper, we propose an Optimal Graph coupled Semi-Supervised Learning (OGSSL) model for EEG emotion recognition by unifying the adaptive graph learning and emotion recognition into a single objective. Besides, we improve the label indicator matrix of unlabeled samples in order to directly obtain their emotional states. Moreover, the key EEG frequency bands and brain regions in emotion expression are automatically recognized by the projection matrix of OGSSL. Experimental results on the SEED-IV data set demonstrate that 1) OGSSL achieves excellent average accuracies of 76.51%, 77.08% and 81.29% in three cross-session emotion recognition tasks, 2) OGSSL is competent for discriminative EEG feature selection in emotion recognition, and 3) the Gamma frequency band, the left/right temporal, prefrontal, and (central) parietal lobes are identified to be more correlated with the occurrence of emotions.**

***Index Terms*— Electroencephalogram (EEG), emotion recognition, feature selection, graph learning, semi-supervised learning.**

Manuscript received March 8, 2022; revised April 17, 2022; accepted May 11, 2022. Date of publication May 16, 2022; date of current version May 20, 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFE0116800, in part by the National Natural Science Foundation of China under Grant 61971173 and Grant U20B2074, in part by the Natural Science Foundation of Zhejiang Province under Grant LY21F030005, in part by the Fundamental Research Fund for the Provincial Universities of Zhejiang under Grant GK209907299001-008, and in part by the China Postdoctoral Science Foundation under Grant 2017M620470. *(Corresponding author: Wanzeng Kong.)*

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Shanghai Jiao Tong University under Protocol No. 2017060.

Yong Peng, Fengzhe Jin, and Wanzeng Kong are with the School of Computer Science, Hangzhou Dianzi University, Hangzhou 310018, China, and also with the Zhejiang Key Laboratory of Brain-Machine Collaborative Intelligence, Hangzhou 310018, China (e-mail: kongwanzeng@hdu.edu.cn).

Feiping Nie is with the School of Artificial Intelligence, OPtics and ElectroNics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, China.

Bao-Liang Lu is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

Andrzej Cichocki is with the Center for Computational and Data-Intensive Science and Engineering, Skolkovo Institute of Science and Technology, 121205 Moscow, Russia.

Digital Object Identifier 10.1109/TNSRE.2022.3175464

I. Introduction

EMOTION recognition is of great significance in realizing the affective brain-computer interfaces, which endeavors to make machine recognize human emotions efficiently and accurately. Generally, emotion states can be judged from four different data modalities, *i.e.*, facial expressions, speech, text and physiological signals [1]. Though it is easier to realize emotion recognition through the former three data modalities, there also exist some inherent drawbacks. For example, we can hide our true emotional states by disguising our facial expressions or changing the tone of our voices. Physiological signals, especially the EEG signals from the central nervous system, are spontaneous with the occurrence of emotions, which offer us a more reliable data source for objective emotion recognition. Recently, rapid progresses have been made in the weak signal acquisition devices and analysis techniques, leading to diverse applications of EEG such as instrument control [2], brainprint recognition [3], emotion recognition [4], fatigue detection [5] and workload identification [6].

In machine learning-based emotion recognition research from EEG signals, there are two fundamental problems. One is how to dynamically model the EEG data properties to well adapt the emotion recognition process. That is, a good learning model is expected to gradually increase the similarity of samples from the same class while simultaneously decrease the similarity of samples from different classes during its optimization process. The other is whether we can obtain some meaningful insights related to the affective effects in addition to the recognition accuracy. In the present work, we move a step forward to handle both problems within the graph-based semi-supervised learning framework.

Semi-supervised learning utilizes both labeled and unlabeled samples simultaneously for better capturing the underlying data properties, which has been widely used in EEG-based

brain-computer interfaces [7], [8]. In [9], a semi-supervised emotion recognition framework was proposed in which the consistency between original and reconstructed data was maximized and the cross-entropy between input and output labels was minimized. By injecting the domain knowledge of labeled data from multiple source subjects into the customized model of the target subject, a semi-supervised joint domain adaptation model was proposed in [10]. Dan *et al.* developed a probabilistic clustering-promoting semi-supervised emotion recognition model with improved reliability in which each sample was constrained to share the same label membership value with its local weighted mean [11]. In [12], multi-modal signals were processed for emotion recognition by semi-supervised learning and neural networks. Besides, the safety of introducing unlabeled samples into learning was explicitly considered in EEG data classification [13].

Graph is an effective data structure to characterize the data relationship [14], based on which many learning tasks can be performed such as clustering [15], dimensionality reduction [16] and semi-supervised learning [8]. Traditionally, graph-based semi-supervised learning models employed a two-stage strategy that first constructs a similarity graph and then performs label propagation on this graph [17]–[19]. Obviously, the label propagation performance depends heavily on the graph quality; therefore, such two-stage strategy breaks the inner connections between graph construction and learning tasks, which easily causes the sub-optimality.

Inspired by recent advances in structured graph learning theory [20], [21], we propose an Optimal Graph coupled Semi-Supervised Learning (OGSSL) model for EEG-based emotion recognition in the paper. OGSSL is an embedded model in which the updating of graph similarity matrix is accompanied with the emotional states estimation of unlabeled EEG samples. Instead of modeling the sample similarities in original feature space, they are modeled according to their subspace representations and the projection matrix is explicitly enforced to be row sparse to enhance the feature selection ability. Besides, we constrain each row of the predicted label indicator matrix of unlabeled EEG data to satisfy the non-negativity and normalization properties. Compared to existing related works, the present work has the following contributions.

- We propose a unified model OGSSL in which the optimal graph learning and semi-supervised EEG emotion recognition are jointly completed. On one hand, OGSSL effectively captures the data characteristics by utilizing both the unlabeled and labeled EEG data. On the other hand, the label propagation in OGSSL is more accurate with the help of dynamically updating the graph, which avoids the sub-optimality caused by the two-stage strategy in graph-based semi-supervised learning.
- We can obtain the emotional state estimation results directly by checking the maximum value in each row of the predicted label matrix. Usually, the label indicator matrix in graph-based semi-supervised learning is real-valued, which requires further postprocessing step to achieve the final assignments. Specifically, the label matrix is enforced to satisfy the non-negativity and row-normalization constraints in OGSSL.
- Besides the recognition accuracy, we provide more analysis on the learned projection matrix from two aspects. One is that it shows excellent performance in selecting discriminative EEG features by enforcing its row-sparsity. The other is that it offers us an automatic and quantitative way in identifying the critical EEG frequency bands and brain regions in emotion recognition.

The remainder of this paper is structured as follows. Section II introduces the detailed model formulation and optimization of OGSSL. Emotion recognition experiments on public EEG data set and the results analysis are described in section III. Section IV discusses the connections and differences of OGSSL and some related models. In section V, we conclude the whole paper.

Notations: In this paper, the upper case and the lower case letters are used to denote matrices and vectors respectively. For matrix $\mathbf{A}$, $\mathbf{a}^i$ is used to denote its i-th row and $\mathbf{a}_j$ denotes its j-th column. a_{ij} denotes its (i, j)-th element. EEG frequency bands are denoted as *Delta*, *Theta*, *Alpha*, *Beta* and *Gamma*.

II. Method

A. Preliminary

In semi-supervised EEG emotion recognition, we are often given an EEG data set $\mathbf{X} \in \mathbb{R}^{d\times n}$ consisting of two subsets. One subset has l labeled EEG samples $\mathbf{X}_l = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l] \in \mathbb{R}^{d\times l}$ whose label matrix is $\mathbf{Y}_l = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l]^T \in \mathbb{R}^{l\times c}$. The other subset has u unlabeled EEG samples $\mathbf{X}_u = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_u] \in \mathbb{R}^{d\times u}$, whose label matrix $\mathbf{Y}_u = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_u]^T \in \mathbb{R}^{u\times c}$ is unknown. Here, d is the EEG sample dimensionality and c represents the number of emotional states. $\mathbf{y}_i \in \mathbb{R}^c, i = 1, 2, \cdots, l$ is defined as

$$y_{ij} = \begin{cases} 1, \text{ if sample } \mathbf{x}_i \text{ belongs to the } j\text{-th state;} \\ 0, \text{ otherwise.} \end{cases} \quad (1)$$

In addition, we introduce an label matrix $\mathbf{F} = [\mathbf{F}_l; \mathbf{F}_u] \in \mathbb{R}^{n\times c}$, where $n = l + u$, $\mathbf{F}_l = \mathbf{Y}_l$ and $\mathbf{F}_u$ is the pseudo-label of the unlabeled EEG data. Our task is to estimate $\mathbf{F}_u$ as accurate as possible given $\mathbf{X} = [\mathbf{X}_l, \mathbf{X}_u]$ and $\mathbf{Y}_l$.

B. OGSSL Model Formulation

In Fig. 1, we show the general framework of OGSSL, which has three components of the discriminative subspace identification, optimal graph learning and the graph-based semi-supervised label estimation.

Considering an undirected graph whose vertices correspond to EEG samples in $\mathbf{X}$ and the edges characterize the connection weights between sample pairs. We use $\mathbf{S} \in \mathbb{R}^{n\times n}$ to denote the graph similarity matrix in which element s_{ij} depicts the connection weights of samples $\mathbf{x}_i$ and $\mathbf{x}_j$. Different from using fixed rules such as the 0-1 weighting, heat kernel weighting and dot-product weighting to measure the similarity [18], [22], we prefer to adaptively learn $\mathbf{S}$ from data. Intuitively, if $\mathbf{x}_i$ and $\mathbf{x}_j$ are similar (*i.e.*, the Euclidean distance between them is small), they are more likely to have the same affective state and therefore, the weight s_{ij} should be large. Otherwise, s_{ij} should be small to express the irrelevance between $\mathbf{x}_i$ and $\mathbf{x}_j$.

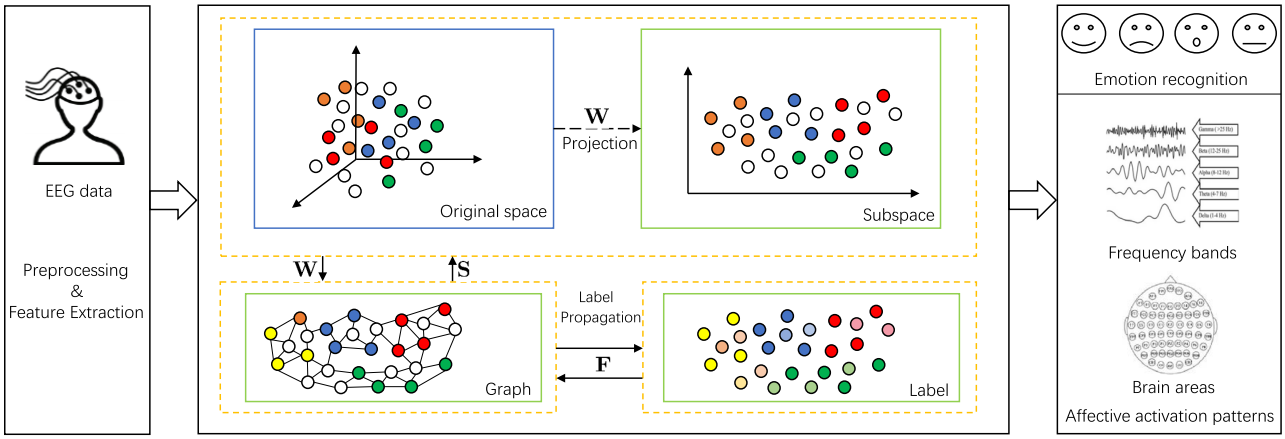


Fig. 1. The general framework of the OGSSL model.

This principle can be implemented by the following objective function

$$\min_{s_{ij}\geq 0, \mathbf{s}^i\mathbf{1}=1} \sum_{i,j=1}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 s_{ij}, \tag{2}$$

where $\mathbf{s}^i \in \mathbb{R}^{1\times n}$ is a row vector with its j-th element as s_{ij}. In (2), the first constraint denotes the non-negativity of the similarity matrix. The second one assumes that the summation of connection probabilities between $\mathbf{x}_i$ and all the other samples is one. Problem (2) might have a trivial solution that only one element in $\mathbf{s}^i$ is one and all the others are zeros [20], [23]. To avoid this, an additional term is incorporated into (2) to shrink the elements in $\mathbf{s}^i$. Then, we have

$$\min_{s_{ij}\geq 0, \mathbf{s}^i\mathbf{1}=1} \sum_{i,j=1}^{n} \left(\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2\right). \tag{3}$$

Because EEG data is non-stationary, which is prone to have redundant and noisy features [1]. This phenomenon is common and challenging in EEG-based emotion recognition task. Intuitively, we can project EEG data into a subspace where the data similarities can be modeled more accurately [24]. Assume that the subspace is identified by projection matrix $\mathbf{W} \in \mathbb{R}^{d\times m}$, where m is the subspace dimensionality. We enable $\mathbf{W}$ to have the ability of selecting discriminative features by enforcing it to be row sparse. Mathematically, we replace $\mathbf{x}_i$ in (3) by $\mathbf{W}^T\mathbf{x}_i$ (similarly, replace $\mathbf{x}_j$ by $\mathbf{W}^T\mathbf{x}_j$) and incorporate $\|\mathbf{W}\|_{2,1}$ as a regularizer; then, we have

$$\begin{aligned} &\min_{\mathbf{S},\mathbf{W}} \sum_{i,j=1}^{n} \left(\|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2\right) + \beta\|\mathbf{W}\|_{2,1}, \\ &s.t.\ s_{ij} \geq 0, \mathbf{s}^i\mathbf{1} = 1, \mathbf{W}^T\mathbf{S}_t\mathbf{W} = \mathbf{I}, \end{aligned} \tag{4}$$

where $\mathbf{S}_t = \mathbf{X}\mathbf{H}\mathbf{X}^T$ is the total-scatter of data, $\mathbf{H} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ is the centering matrix. The third constraint in (4) ensures that the subspace data should be statistically uncorrelated.

However, problem (4) considers only the data similarities in feature space while neglects the consistency of label space. We use $\mathbf{f}^i \in \mathbb{R}^{1\times c}$ and $\mathbf{f}^j$ are label indicator vectors of samples $\mathbf{x}_i$ and $\mathbf{x}_j$. According to the local invariance idea that similar samples generate similar labels, if s_{ij} is large which means that $\mathbf{x}_i$ and $\mathbf{x}_j$ are similar, then $\mathbf{f}^i$ should be similar to $\mathbf{f}^j$. Mathematically, we have

$$\min_{\mathbf{F}_l=\mathbf{Y}_l, \mathbf{F}_u\geq 0, \mathbf{F}_u\mathbf{1}=1} \sum_{i,j=1}^{n} \|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij} = \mathrm{Tr}(\mathbf{F}^T\mathbf{L}\mathbf{F}), \tag{5}$$

where $\mathbf{L} = \mathbf{D} - \mathbf{S}$ is the graph Laplacian matrix, $\mathbf{D}$ is a diagonal matrix and the i-th diagonal element is defined as $d_{ii} = \sum_j s_{ij}$. The constraints $\mathbf{F}_u \geq \mathbf{0}$ and $\mathbf{F}_u\mathbf{1} = \mathbf{1}$ are equivalent to $\mathbf{f}^i \geq \mathbf{0}$ and $\mathbf{f}^i\mathbf{1} = 1$ ($i = l+1, \cdots, n$). Then, each element in $\mathbf{f}^i$ can be viewed as the probability of $\mathbf{x}_i$ belonging to a certain class. For example, if $\mathbf{f}^i = [0.09, 0.82, 0.05, 0.04]$, the probabilities of the i-th sample to the four emotional states are 0.09, 0.82, 0.05 and 0.04, respectively. Accordingly, this sample should be marked with the second emotional state.

Finally, by combining (4) and (5), we formulate the objective function of OGSSL as

$$\begin{aligned} &\min_{\mathbf{S},\mathbf{W},\mathbf{F}} \sum_{i,j=1}^{n} \left(\|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2\right) \\ &\qquad + \beta\|\mathbf{W}\|_{2,1} + \gamma\,\mathrm{Tr}(\mathbf{F}^T\mathbf{L}\mathbf{F}), \\ &s.t.\ \mathbf{W}^T\mathbf{S}_t\mathbf{W} = \mathbf{I}, \\ &s.t.\ \mathbf{S} \geq \mathbf{0}, \mathbf{S}\mathbf{1} = \mathbf{1}, \mathbf{F}_l = \mathbf{Y}_l, \mathbf{F}_u \geq \mathbf{0}, \mathbf{F}_u\mathbf{1} = \mathbf{1}, \end{aligned} \tag{6}$$

where α, β and γ represent three regularization parameters.

C. OGSSL Model Optimization

Below we derive the updating rules to $\mathbf{S}$, $\mathbf{W}$ and $\mathbf{F}_u$. Basically, we update one variable by fixing the two others.

■ Update $\mathbf{F}_u$ by fixing $\mathbf{W}$ and $\mathbf{S}$. The corresponding objective function is

$$\begin{aligned} &\min \sum_{i=l+1}^{n} \sum_{j=l+1}^{n} \|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij}, \\ &s.t.\ \mathbf{f}^i \geq \mathbf{0}, \mathbf{f}^i\mathbf{1} = 1, \end{aligned} \tag{7}$$

where $\mathbf{f}^i \in \mathbb{R}^{1\times c}$ is the i-th row of $\mathbf{F}_u$. Obviously, the above objective can be decoupled into the following form for each i. That is

$$\min_{\mathbf{f}^i\geq\mathbf{0}, \mathbf{f}^i\mathbf{1}=1} \sum_{j=l+1}^{n} \|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij} = \sum_{j=l+1, j\neq i}^{n} \|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij}. \tag{8}$$

By completing the square form w.r.t. $\mathbf{f}^i$, (8) is equivalent to

$$\min_{\mathbf{f}_i \geq \mathbf{0}, \mathbf{f}^i \mathbf{1}=1} \|\mathbf{f}^i - \mathbf{m}\|_2^2, \tag{9}$$

where $\mathbf{m} = \sum_{j=l+1, j\neq i}^{n} s_{ij}\mathbf{f}^j$. It defines an Euclidean projection on a simplex which can be solved by the Lagrangian multiplier method combined with Newton's method [25].

■ Update $\mathbf{W}$ by fixing $\mathbf{S}$ and $\mathbf{F}$. The sub-objective function with respect to variable $\mathbf{W}$ is

$$\min_{\mathbf{W}\mathbf{S}_t\mathbf{W}=\mathbf{I}} \sum_{i,j=1}^{n} \|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2 s_{ij} + \beta\|\mathbf{W}\|_{2,1}, \tag{10}$$

which is equivalent to

$$\min_{\mathbf{W}\mathbf{S}_t\mathbf{W}=\mathbf{I}} \mathrm{Tr}(\mathbf{W}^T\mathbf{X}\mathbf{L}_\mathbf{S}\mathbf{X}^T\mathbf{W}) + \beta\mathrm{Tr}(\mathbf{W}^T\mathbf{Q}\mathbf{W}). \tag{11}$$

Here $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is a diagonal matrix and its i-th diagonal element is defined as

$$q_{ii} = \frac{1}{2\|\mathbf{w}^i\|_2}.^1 \tag{12}$$

The Lagrangian function of (11) is

$$\mathcal{L}(\mathbf{W}) = \mathrm{Tr}(\mathbf{W}^T(\mathbf{X}\mathbf{L}_\mathbf{S}\mathbf{X}^T + \beta\mathbf{Q})\mathbf{W}) - \mathrm{Tr}(\mathbf{\Lambda}(\mathbf{W}^T\mathbf{S}_t\mathbf{W} - \mathbf{I})). \tag{13}$$

By taking the derivative of $\mathcal{L}(\mathbf{W})$ w.r.t. $\mathbf{W}$, and setting it to zero, we have

$$(\mathbf{X}\mathbf{L}_\mathbf{S}\mathbf{X}^T + \beta\mathbf{Q})\mathbf{W} - \mathbf{S}_t\mathbf{W}\mathbf{\Lambda} = \mathbf{0}. \tag{14}$$

Then, the optimal solution $\mathbf{W}$ to (11) consists of the m eigenvectors of $\mathbf{S}_t^{-1}(\mathbf{X}\mathbf{L}_\mathbf{S}\mathbf{X} + \beta\mathbf{Q})$ corresponding to its m smallest eigenvalues. To guarantee that $\mathbf{S}_t$ is invertible, we assume that the null space of $\mathbf{X}$ has been removed.

■ Update $\mathbf{S}$ by fixing $\mathbf{W}$ and $\mathbf{F}$. The objective function is

$$\begin{aligned}\min_{\mathbf{S}} & \sum_{i,j=1}^{n} \left(\|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2\right) \\ & + \gamma \sum_{i,j=1}^{n} \|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij}, \\ s.t.\ & \forall i, \mathbf{s}_i \geq \mathbf{0}, \mathbf{s}^i\mathbf{1} = 1.\end{aligned} \tag{15}$$

Because problem (15) can be decoupled for each i, we propose to solve $\mathbf{S}$ in row-wise manner, namely

$$\begin{aligned}\min_{\mathbf{S}} & \sum_{j=1}^{n} \left(\|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2 + \gamma\|\mathbf{f}^i - \mathbf{f}^j\|_2^2 s_{ij}\right) \\ s.t.\ & \mathbf{s}_i \geq \mathbf{0}, \mathbf{s}^i\mathbf{1} = 1.\end{aligned} \tag{16}$$

We denote $d_{ij}^{wx} = \|\mathbf{W}^T\mathbf{x}_i - \mathbf{W}^T\mathbf{x}_j\|_2^2$, $d_{ij}^f = \|\mathbf{f}^i - \mathbf{f}^j\|_2^2$, and then $\mathbf{d}^i \in \mathbb{R}^{1\times n}$ as a vector whose j-th element is $d_{ij} = d_{ij}^{wx} + \gamma d_{ij}^f$. Then, the vector form of problem (16) is

$$\min_{\mathbf{s}^i \geq \mathbf{0}, \mathbf{s}^i\mathbf{1}=1} \left\|\mathbf{s}^i + \frac{1}{2\alpha}\mathbf{d}^i\right\|_2^2. \tag{17}$$

[1]In experiments, we usually use $q_{ii} = \frac{1}{2\sqrt{\mathbf{w}^i(\mathbf{w}^i)^T + \varsigma}}$ instead to handle the case when the denominator approaches zero. ς is a small constant.

Obviously, problem (17) shares the identical form with problem (9), which can also be solved by the method in [25].

As a summary, we collect the above optimization steps in Algorithm 1. Obviously, the complexity in optimizing the OGSSL objective function is mainly from the loop in Algorithm 1. Below we analyze its computational complexity by using the big $\mathcal{O}$ notation. We need $\mathcal{O}(uc)$ time to obtain $\mathbf{F}_u$. When updating $\mathbf{W}$, the complexity is $\mathcal{O}(n^2d + nd^2 + d^3)$. For each $i \in [1, n]$, we need $\mathcal{O}(dm)$ time to update $\mathbf{s}^i$. Therefore, we need $\mathcal{O}(nmd)$ to obtain $\mathbf{S}$. The complexity of updating $\mathbf{Q}$ is $\mathcal{O}(dm)$. Assuming that the number of iterations is t, the computational complexity of optimizing OGSSL is $\mathcal{O}(t(uc + n^2d + nd^2 + d^3 + ndm + dm))$. Considering that the usual case is $n \approx u > d > m \gg c$ in semi-supervised learning, the overall complexity of OGSSL is $\mathcal{O}(tn^2d)$.

Algorithm 1 The Optimization to OGSSL Objective Function

Input: Labeled EEG data $\mathbf{X}_l \in \mathbb{R}^{d\times l}$ and its corresponding label matrix $\mathbf{Y}_l \in \mathbb{R}^{l\times c}$, unlabeled EEG data $\mathbf{X}_u \in \mathbb{R}^{d\times u}$, parameters α, β and γ;

Output: The estimated label $\mathbf{F}_u \in \mathbb{R}^{u\times c}$.

1: Initialize the graph similarity matrix $\mathbf{S}$ by solving problem $\min_{\forall i, \mathbf{s}_i \geq \mathbf{0}, \mathbf{s}^i\mathbf{1}=1} \sum_{i,j=1}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 s_{ij} + \alpha s_{ij}^2$;
2: Initialize $\mathbf{F}_u = \frac{1}{c}\mathbf{1}\mathbf{1}^T \in \mathbb{R}^{u\times c}$ and $\mathbf{Q} \in \mathbb{R}^{d\times d}$ as an identity matrix;
3: **while** not converged **do**
4: Update $\mathbf{L}_\mathbf{S} = \mathbf{D}_\mathbf{S} - \frac{\mathbf{S}^T + \mathbf{S}}{2}$, where $\mathbf{D}_\mathbf{S}$ is a diagonal matrix with the i-th diagonal element as $\sum_j (s_{ij} + s_{ji})/2$;
5: For each $i|_{i=l+1}^{n}$, update the i-th row of $\mathbf{F}_u$ by solving problem (9);
6: Update $\mathbf{W}$, whose columns are the m eigenvectors of $\mathbf{S}_t^{-1}(\mathbf{X}\mathbf{L}_\mathbf{S}\mathbf{X}^T + \beta\mathbf{Q})$ corresponding to the m smallest eigenvalues;
7: Update $\mathbf{Q}$ by equation (12);
8: For each $i|_{i=1}^{n}$, update the i-th row of $\mathbf{S}$ by solving problem (17);
9: **end while**

III. Experiments

In the section, we aim to answer the following three questions by experiments. 1) How much performance improvement can be achieved by our proposed OGSSL model in comparison with existing models? 2) What is the performance of OGSSL in feature selection? 3) What can we learn about the insights to emotion expression besides the recognition accuracy? The source code for implementing OGSSL is available from https://github.com/SunseaIU/OGSSL.

A. Data Description

We choose the SEED-IV data set in the following experiments [26], which is a public emotional EEG data set provided by Shanghai Jiao Tong University https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html. The basic properties of the SEED-IV data set are stated as follows. A total of 15 healthy subjects participated in the EEG data

TABLE I
SUMMARY OF THE SEED-IV EMOTIONAL EEG DATA SET

Item	Properties
# subject	15
# session	3
feature	differential entropy (DE)
# electrode	62
frequency bands	*Delta*, *Theta*, *Alpha*, *Beta*, *Gamma*
emotional states	*sad*, *fear*, *happy*, *neutral*

collection experiment. Four types of emotional states, *i.e.,* *sad*, *fear*, *happy* and *neutral*, were elicited by 72 well-chosen video clips. Each subject participated in three data collection experiments, corresponding to the three sessions. Therefore, there are total 45 sessions in SEED-IV because each of the 15 subjects has three sessions. In each session, 24 video clips were played among which six clips correspond to one emotional state. During watching the videos, EEG data was simultaneously recorded by a 62-channel electrode cap according to the international 10-20 placement. The sampling rate is 1000 Hz.

Each session is sliced into four-second non-overlapping segments, each of which is regarded as one sample for model training. For EEG signal preprocessing, raw EEG data were first down-sampled to the sampling rate of 200 Hz and then band-pass filtered to 1-75 Hz for artifact removal. Differential entropy (DE) features were extracted within each segment at five frequency bands, *i.e., Delta* (1-3 Hz), *Theta* (4-7 Hz), *Alpha* (8-13 Hz), *Beta* (14-30 Hz) and *Gamma* (31-50 Hz). The DE feature is defined as

$$h(\mathbf{X}) = -\int_x f(x)\ln f(x)dx, \tag{18}$$

where $\mathbf{X}$ is a random variable with probability density function $f(x)$. We assume that EEG data follows the Gaussian distribution, *i.e.*, $f(x) = \mathcal{N}(x;\mu,\sigma^2)$, and then calculate its differential entropy by

$$\begin{aligned} h(X) &= -\int_{-\infty}^{\infty} f(x)\ln\frac{1}{\sqrt{2\pi\sigma^2}}\exp\frac{(x-\mu)^2}{2\sigma^2}dx \\ &= \frac{1}{2}\ln(2\pi\sigma^2) + \frac{Var(X)}{2\sigma^2} = \frac{1}{2}\ln(2\pi e\sigma^2). \end{aligned} \tag{19}$$

where μ is the expectation and σ is the standard deviation of the probability density function, respectively. By concatenating the 62 values (corresponding to the 62 EEG channels) of each of the five frequency bands together, we obtained the sample vectors whose dimensionality is 310. Because the time durations of different video clips are slightly different, the three sessions generate different numbers of EEG samples, *i.e.*, 851, 832 and 822, respectively. The general properties of the SEED-IV data set are shown in Table I.

B. Experimental Settings

On the control experiments, we compare the emotion recognition performance of OGSSL with some related semi-supervised classification models. They are the semi-supervised PCAN (semiPCAN) [20], the semi-supervised support vector machine (semiSVM), the semi-supervised Linear Square Regression (semiLSR), the Rescaled Linear Square Regression (RLSR) [27], [28], and the Robust Discriminative Sparse regression (RDSR) [29]. In semiPCAN, we first learn an optimal graph based on PCAN, and then perform the label propagation to predict the emotional states of unlabeled data. In semiSVM, linear kernel is used. It is worth mentioning that semiLSR is a degenerated version without explicitly considering the feature auto-weighting variable. The label indicator matrix $\mathbf{Y}_u$ is initialized as an all-one matrix multiplied by $1/c$. The maximum number of iterations for model training is set as 50.

On the experimental paradigm, we perform the subject-dependent cross-session EEG emotion recognition in chronological order. Because each subject has three different sessions, we have three different cross-session emotion recognition tasks, *i.e.*, session1→session2, session1→session3, and session2→session3. Taking the 'session1→session2' task as an example, EEG samples from session 1 are fully labeled and those from session 2 as unlabeled. Then, the inputs to OGSSL are ($\mathbf{X}_l \in \mathbb{R}^{310\times 851}, \mathbf{Y}_l \in \mathbb{R}^{851\times 4}$) and $\mathbf{X}_u \in \mathbb{R}^{310\times 832}$, where 310 is the sample dimensionality, 851 and 832 are the numbers of samples in session 1 and session 2, respectively. Then, our goal is to estimate the emotional states of these unlabeled EEG samples as accurate as possible. In this sense, OGSSL essentially is a transductive semi-supervised learning model to estimate the emotional state of unlabeled samples which are involved in model training. Given the differential entropy-based EEG samples, considering that there might be distribution divergences for EEG data collected from different sessions, we perform a two-step data preprocessing before model training. One is mapping the feature values into [0, 1] to avoid the scale mismatch. The other is centering the data by removing the sample mean.

C. Results and Analysis

In Table II, we present the experimental results of these compared models on SEED-IV, where the bold numbers are used to highlight the best accuracies. These results provide us with the following inspirations.

- Obviously, OGSSL obtained the best average performance. In most cases, it achieves the best accuracies compared to the other five models. In the three cross-session emotion recognition tasks, the average accuracies of OGSSL are 76.51%, 77.08%, 81.29%, respectively. Compared with the second-best model RSDR, OGSSL made performance improvements by 5.6%, 4.5%, 5.35% in respective tasks. Therefore, we conclude that collaboratively optimizing the graph similarity matrix and the label indicator matrix of the unlabeled EEG data is beneficial for obtaining higher recognition accuracy.
- The performance of semiPCAN is worse than the other models. Specifically, semiPCAN obtained the average accuracies of 55.85%, 57.66%, 58.86% corresponding to the three tasks. The reasons are two folds. One is that the unsupervised graph learning process cannot make effective use of the existing label information. The other

TABLE II
THE EEG-BASED EMOTION RECOGNITION RESULTS (%) OF DIFFERENT MODELS ON THE SEED-IV DATA SET

	session1→session2						session1→session3						session2→session3					
	m1	m2	m3	m4	m5	m6	m1	m2	m3	m4	m5	m6	m1	m2	m3	m4	m5	m6
s1	52.76	39.42	57.09	55.65	60.22	**65.14**	54.99	50.12	64.72	70.68	69.95	**80.54**	53.41	56.69	61.44	62.65	65.09	**73.36**
s2	76.32	78.97	91.23	89.18	87.62	**97.36**	75.43	71.05	84.55	89.29	86.37	**91.85**	81.75	79.20	85.40	83.21	82.36	**90.63**
s3	51.68	53.37	60.10	69.71	63.46	**75.48**	45.13	57.66	48.42	48.78	54.87	**66.42**	43.31	67.88	63.75	67.27	72.26	**79.20**
s4	37.50	31.13	63.22	**68.39**	64.30	65.99	56.69	57.66	71.05	71.17	68.73	**79.68**	54.14	67.64	82.48	80.17	75.18	**82.73**
s5	40.63	47.24	59.50	67.67	67.91	**69.95**	43.07	43.07	52.55	58.39	**60.34**	52.43	57.66	64.23	76.64	72.87	75.55	**77.98**
s6	50.12	48.32	69.83	71.03	67.19	**71.39**	74.70	60.46	80.17	83.45	83.45	**85.89**	60.22	65.57	76.64	83.70	79.56	**89.90**
s7	66.95	68.99	82.93	80.77	75.00	**85.94**	66.55	57.91	86.98	88.44	82.60	**91.85**	82.00	85.04	84.31	88.56	86.62	**90.15**
s8	66.59	70.91	68.87	69.95	71.39	**83.05**	65.33	71.41	78.22	80.78	78.10	**80.90**	55.96	64.48	76.52	82.97	72.14	**84.18**
s9	57.81	60.82	67.67	78.73	**81.13**	78.73	47.57	48.30	61.68	62.77	72.02	**77.98**	53.04	59.25	47.45	61.80	75.18	**78.59**
s10	52.64	58.17	46.75	53.85	68.27	**74.88**	60.83	61.31	45.26	49.64	63.99	**73.97**	45.01	53.16	72.87	78.35	78.10	**85.16**
s11	50.36	53.37	50.00	52.04	60.58	**67.79**	54.99	60.71	76.28	71.17	76.76	**78.83**	47.81	52.55	52.68	59.49	59.61	**61.19**
s12	44.23	40.26	**60.34**	53.13	56.73	53.25	40.63	33.09	**65.82**	65.45	66.18	55.84	37.47	42.09	**67.88**	63.63	64.36	66.79
s13	42.79	54.81	58.05	68.63	67.67	**75.60**	55.84	61.80	54.87	62.41	61.31	**64.48**	57.18	41.85	56.45	64.48	74.94	**76.40**
s14	67.79	68.75	79.33	76.92	75.00	**85.70**	61.56	55.84	77.49	82.85	74.94	**84.31**	68.61	71.53	88.44	87.10	87.83	**92.70**
s15	79.57	81.37	88.58	87.14	97.12	**97.36**	61.68	80.78	85.52	85.40	89.05	**91.24**	85.28	85.40	88.08	89.90	90.27	**90.39**
Avg.	55.85	57.06	66.90	69.52	70.91	**76.51**	57.66	58.08	68.90	71.38	72.58	**77.08**	58.86	63.77	72.07	75.08	75.94	**81.29**

In this table, s1 to s15 denote the 15 subjects in SEED-IV. m1 to m6 respectively denote the semiPCAN, semiSVM, semiLSR, RLSR, RDSR and OGSSL.

is that it is a two-stage method in which the graph construction and the label propagation cannot effectively adapt to each other. Though semiSVM obtained slightly better accuracies than semiPCAN, its overall performance is still not satisfactory. The possible reason is that directly characterizing the data correlations by linear kernel in the original data representations is not powerful enough to remove the external influences and capture the semantic connections of EEG samples.

- By comparing the results respectively obtained by semiLSR and RLSR, the latter achieved performance improvements by 2.62%, 2.48%, 3.01% in comparison with the former in the three recognition tasks. The only difference between them is that RLSR additionally incorporates a feature self-weighting variable to adaptively learn the contributions of different feature dimensions in emotion recognition. Therefore, we have reason to believe that such feature weighting technique is beneficial for improving the model discriminative ability.
- By comparing the results of respectively obtained by RLSR and RDSR, we find that the results of RDSR are generally better than those of RLSR. Specifically, RLSR obtains the average accuracies of 69.52%, 71.38%, 75.08% and RDSR obtains the average accuracies of 70.91%, 72.58%, 75.94%. That is, RDSR improves the performance of RLSR by 1.39%, 1.20%, 0.86% in the three semi-supervised EEG emotion recognition tasks. RLSR directly builds mapping between data matrix and the label indicator matrix while in RDSR, the local label consistency is jointly considered according to the local invariance idea.

In addition, we use confusion matrices to reorganize the recognition accuracies of these compared models in Fig. 2. From this figure, we obtain the recognition rate of each model on each emotional state and also the misclassification rates of each emotional state into the others. Taking the confusion matrix of OGSSL as an example, the recognition rates of the four emotional states, *sad*, *fear*, *happy* and *neutral*, are

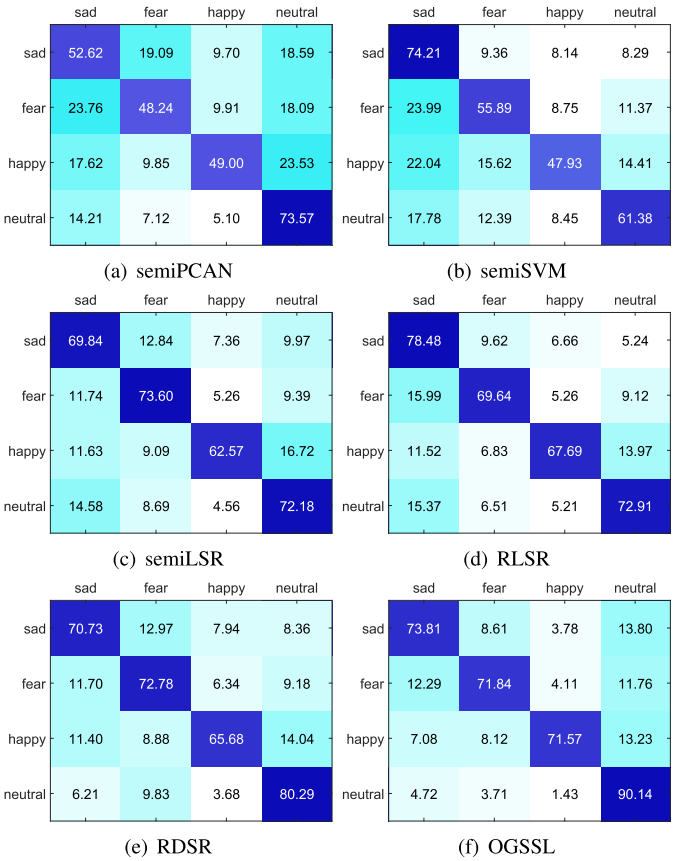


Fig. 2. Emotion recognition accuracies (%) of different models represented by confusion matrices.

73.81%, 71.84%, 71.57% and 90.14%, respectively. Obviously, our model has the best accuracy 90.14% in recognizing the EEG samples in *neutral* state, which is higher than the other three emotional states. In addition, we know that 4.72%, 3.71%, and 1.43% of the *neutral* samples were respectively misclassified as the *sad*, *fear*, and *happy* states.

In OGSSL, the graph similarity matrix is adaptively updated according to the subspace representations of EEG samples to describe their relationship. To intuitively show the iterative optimization process, we visualize the learned graphs in terms of different iterations, *i.e.*, when the numbers of iterations are 0, 5, 20 and 30, respectively. According to the structured graph learning theory [20], [21], the number of connected graph components should be equal to the number of classes ideally. As shown in Fig. 3, we see that the contours of the four diagonal blocks are gradually clear, corresponding to the four emotional states in SEED-IV. This demonstrate that as the OGSSL model iterates, the learned graph similarity matrix can better characterize the relationship among EEG samples. That is, the connections of samples with the same emotional state become gradually enhanced. Besides, this figure shows us the nice convergence property of the proposed optimization method to OGSSL objective function, which monotonically decreases the objective values as the number of iterations increases.

Moreover, we demonstrate the similarity learning process of some sample pairs from the 'subject2: session1→session2' case in Fig. 4. As shown in Fig. 4(a), we select five sample pairs and the two EEG samples in each pair are belonging to the same emotional state. As the number of iterations increases, we observe that their similarity values gradually increase and eventually stay on a high level. On the contrary, in Fig. 4(b), the two samples in each of the five pairs are corresponding to different emotional states. Accordingly, their similarity values decrease as the number of iterations increases.

Below we perform the sensitivity analysis of the OGSSL emotion recognition performance in terms of the three regularization parameters α, β and γ by taking the case of 'subject2: session1→session2' as an example. Concretely, we show the recognition accuracies of OGSSL in terms of two parameters by fixing the third one. The resultant bar plots are shown in Fig. 5, where γ, β and α are respectively fixed as 2^4, 2^9 and 2^9, corresponding to the three subfigures. From these results, we can observe that there are many choices of these parameters to make OGSSL achieve high recognition rates. Specifically, we can set α, β, γ respectively from $\{2^8, 2^9, 2^{10}\}$, $\{2^6, 2^7, \cdots, 2^{11}\}$ and $\{2^5, 2^6, \cdots, 2^{10}\}$ in this case. Similar results can be found in the other emotion recognition cases. Therefore, we generally conclude that the proposed OGSSL model is not very sensitive to these regularization parameters.

D. Feature Selection Experiments

In OGSSL objective function, the row-sparsity property of the projection matrix is achieved by the $\ell_{2,1}$-norm, which assigns OGSSL the ability of feature selection. Since EEG features originate from different frequency bands and brain regions, they should have different discriminative abilities in emotion recognition. Therefore, we evaluate the feature selection ability of OGSSL by comparing it with several feature selection models including two supervised ones (maximum relevance minimum redundancy (mRMR) [30], the $\ell_{2,1}$-norm [31]) and three semi-supervised ones (RLSR [27], max-relevance and min-redundancy criterion based on Pearson's correlation coefficient (RRPC) [32], adaptive discriminant analysis for semi-supervised feature selection (SADA) [33]). Linear SVM serves as the classifier on the new EEG data formed by the selected features, whose regularization parameter C is searched from $\{2^{-20}, 2^{-19}, \cdots, 2^{20}\}$. We set the parameters in respective models as those in the original papers. For each model, we selected 10, 20, 30, 50, 100, and 200 features respectively and recorded the number of features when the best recognition accuracy was achieved. We report the feature selection results in Tables III, IV, and V, where we use boldface to highlight the best accuracy in each case. The number of selected features corresponding to the best classification accuracy of each model is appended in brackets.

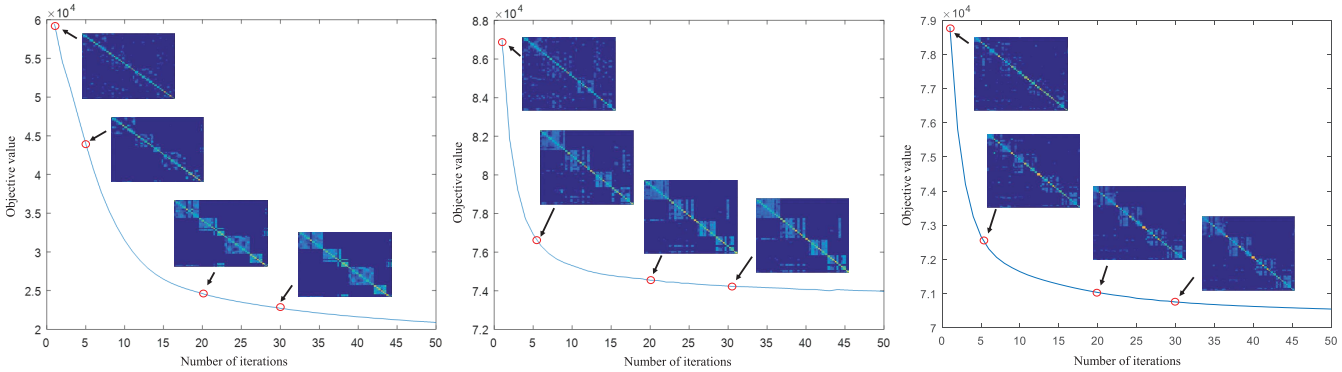
TABLE III
FEATURE SELECTION RESULTS (%) ON THE 'SESSION1→SESSION2' TASK

	mRMR	L21	PRPC	RLSR	SADA	OGSSL
s1	43.03(200)	41.83(200)	67.07(20)	69.95(30)	50.00(10)	**71.03**(100)
s2	84.13(200)	81.85(100)	75.72(200)	77.76(200)	85.34(200)	**85.58**(30)
s3	55.77(50)	57.93(20)	49.76(100)	48.44(10)	**60.94**(100)	56.85(20)
s4	66.23(30)	53.25(200)	50.36(30)	59.25(100)	56.85(20)	**69.47**(10)
s5	58.17(30)	54.33(200)	67.79(20)	53.97(200)	64.66(100)	**69.95**(50)
s6	42.67(30)	39.90(20)	44.23(200)	51.56(20)	54.57(200)	**60.82**(10)
s7	60.94(100)	60.10(10)	64.42(200)	68.39(100)	66.11(50)	**69.11**(100)
s8	68.03(200)	68.03(100)	67.19(200)	71.03(20)	65.50(20)	**74.40**(100)
s9	71.27(50)	70.91(50)	58.77(10)	54.93(10)	70.07(30)	**72.60**(10)
s10	31.49(50)	33.77(200)	34.62(10)	31.61(10)	**49.28**(100)	31.73(10)
s11	43.87(200)	45.91(200)	44.11(50)	44.23(100)	**49.04**(200)	**49.04**(10)
s12	39.78(20)	37.14(10)	42.07(30)	43.15(200)	41.71(50)	**43.63**(30)
s13	57.09(200)	63.94(30)	62.38(100)	62.50(10)	58.17(200)	**65.99**(20)
s14	67.79(200)	72.24(200)	78.25(200)	77.88(50)	70.19(200)	**79.93**(50)
s15	94.23(200)	91.23(100)	**98.56**(200)	93.27(200)	93.03(200)	**98.56**(200)

We have the following insights into these results. 1) OGSSL exhibits the best performance in most cases. Specifically, it achieved the best feature selection effect in 39 out of the total 45 cases, indicating its excellent ability in selecting discriminative EEG features. From the machine learning perspective, OGSSL is an embedded semi-supervised feature selection model to implement the feature selection jointly with the target label estimation. That is, the feature selection process can be well guided by the adaptive graph-based semi-supervised label propagation process. 2) We find that in many cases, OGSSL achieved the best accuracies when only 10 features are selected, which is significantly less than both the original feature dimensionality (*i.e.*, 310) and the maximum number of selected features (*i.e.*, 200). This shows that EEG data in the original feature representation has much redundancy and it is meaningful to perform feature selection in EEG-based emotion recognition.

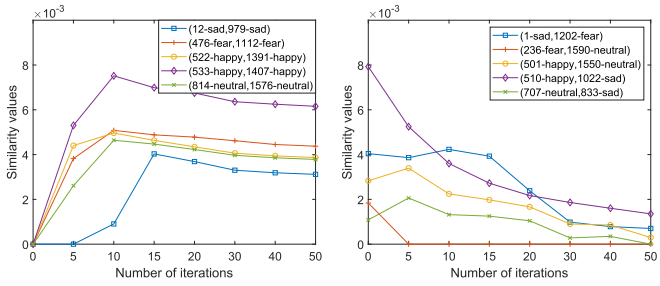
E. Affective Activation Patterns Exploration

From the above two subsections, we have realized that OGSSL not only improves the emotion recognition accuracy, but also identifies the discriminative EEG features. In this section, we perform analysis on where these discriminative EEG features are mainly from, *i.e.*, frequency bands and channels, from the pattern recognition perspective.



(a) subject2: session1→session2 (b) subject2: session1→session3 (c) subject2: session2→session3

Fig. 3. The iterative processes of optimal graph learning in OGSSL. The four diagonal blocks respectively correspond to the emotional states of *sad*, *fear*, *happy* and *neutral* in the SEED-IV data set.



(a) Intra-class sample pairs (b) Inter-class sample pairs

Fig. 4. The similarity learning process of some EEG sample pairs by OGSSL.

TABLE IV
FEATURE SELECTION RESULTS (%) ON THE 'SESSION1→SESSION3' TASK

	mRMR	L21	PRPC	RLSR	SADA	OGSSL
s1	61.44(30)	64.60(100)	71.05(200)	83.94(100)	70.32(100)	**84.31**(100)
s2	74.94(50)	75.67(100)	78.95(200)	88.44(100)	76.89(20)	**90.15**(200)
s3	56.57(50)	45.26(30)	47.32(200)	47.57(200)	51.70(20)	**59.85**(200)
s4	70.56(30)	**72.26**(200)	52.07(200)	70.19(100)	70.92(100)	68.98(50)
s5	53.28(100)	55.11(50)	57.54(100)	56.69(50)	54.74(100)	**58.64**(50)
s6	72.26(50)	73.11(30)	68.37(200)	72.87(30)	73.72(200)	**85.04**(50)
s7	68.25(200)	69.59(10)	68.73(200)	79.20(10)	64.48(200)	**93.92**(10)
s8	77.49(200)	77.62(200)	71.65(200)	77.74(50)	76.28(200)	**81.87**(200)
s9	61.68(100)	67.76(20)	50.00(100)	49.88(200)	61.19(200)	**77.86**(10)
s10	41.00(100)	35.16(200)	**45.74**(200)	**45.74**(20)	40.02(200)	**45.74**(10)
s11	60.46(200)	62.90(200)	64.72(50)	60.34(200)	66.79(200)	**67.15**(10)
s12	43.80(10)	45.50(30)	51.46(100)	49.88(20)	48.91(50)	**65.94**(10)
s13	50.85(200)	47.32(100)	52.19(20)	51.58(100)	49.03(100)	**52.31**(10)
s14	58.76(10)	75.06(30)	72.38(10)	79.08(20)	68.86(200)	**83.58**(30)
s15	85.64(10)	87.59(100)	85.28(200)	77.13(200)	85.52(200)	**89.54**(10)

To this end, on one hand we should obtain the quantitative importance measures of all feature dimensions; on the other hand, we should establish the correspondence between EEG frequency bands (channels) and feature dimensions.

We propose to use the normalized ℓ_2-norm of each row of the projection matrix $\mathbf{W}$ to quantitatively characterize the importance of each feature dimension. Specifically, the importance of the i-th EEG feature, $\theta_i (i = 1, 2, \cdots, 310)$, can be calculated by

$$\theta_i = \frac{\|\mathbf{w}^i\|_2}{\sum_{j=1}^{d} \|\mathbf{w}^j\|_2}, \tag{20}$$

TABLE V
FEATURE SELECTION RESULTS (%) ON THE 'SESSION2→SESSION3' TASK

	mRMR	L21	PRPC	RLSR	SADA	OGSSL
s1	54.87(200)	55.84(200)	72.26(100)	68.37(200)	63.63(100)	**73.36**(10)
s2	77.25(10)	76.89(10)	86.37(200)	74.82(100)	79.20(50)	**86.86**(200)
s3	59.98(200)	63.02(10)	64.72(30)	55.72(30)	66.42(100)	**74.09**(10)
s4	68.86(200)	70.07(50)	63.14(100)	65.45(50)	66.42(30)	**70.44**(50)
s5	53.16(200)	60.83(10)	55.35(200)	56.45(100)	62.29(50)	**63.63**(10)
s6	81.27(100)	83.21(10)	58.03(200)	79.08(20)	78.95(200)	**85.04**(30)
s7	84.18(100)	83.70(50)	80.41(200)	85.77(10)	82.36(200)	**86.50**(50)
s8	61.44(200)	69.34(50)	67.88(20)	69.22(50)	65.94(200)	**84.18**(10)
s9	71.17(10)	61.56(200)	50.49(50)	60.34(10)	66.30(200)	**72.99**(10)
s10	79.68(200)	80.05(20)	59.25(200)	69.10(50)	71.29(50)	**80.78**(10)
s11	51.09(100)	54.87(200)	**57.06**(20)	55.72(20)	53.77(200)	55.35(200)
s12	60.58(50)	60.22(50)	58.39(200)	56.57(10)	**62.29**(100)	58.52(10)
s13	63.63(30)	59.00(200)	53.04(20)	54.74(50)	53.65(50)	**69.95**(10)
s14	68.00(200)	89.90(100)	**91.00**(200)	89.66(50)	85.04(200)	89.29(200)
s15	88.20(20)	88.08(50)	87.59(200)	**90.39**(100)	88.08(200)	**90.39**(50)

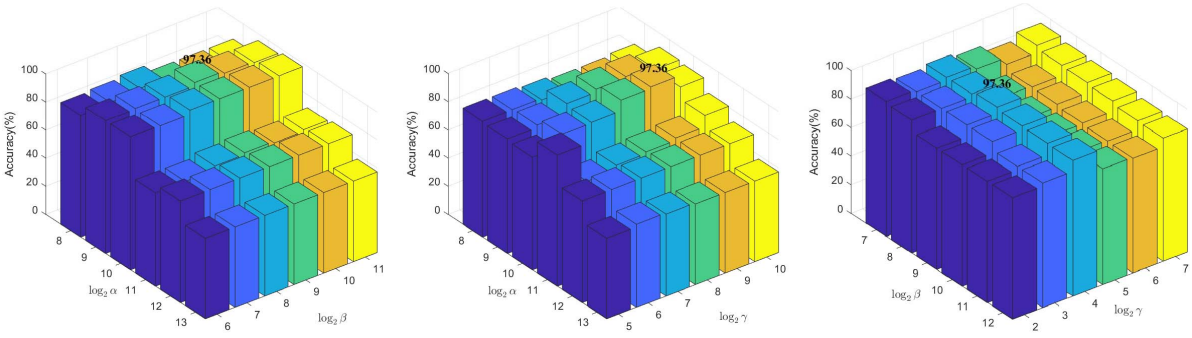
where $\mathbf{w}^i$ is the i-th row of the projection matrix $\mathbf{W}$. Obviously, the greater the value of θ_i, the more discriminative the i-th EEG feature dimension in differentiating emotional states. Considering that we have p EEG frequency bands and q channels, according to the established correspondence between EEG frequency bands and feature dimensions [4], the importance of the $i|_{i=1}^{p}$-th frequency band can be calculate by

$$\omega(i) = \theta_{(i-1)\times q+1} + \theta_{(i-1)\times q+2} + \cdots + \theta_{i\times q}. \tag{21}$$

For the $j|_{j=1}^{q}$-th channel, we measure its importance by

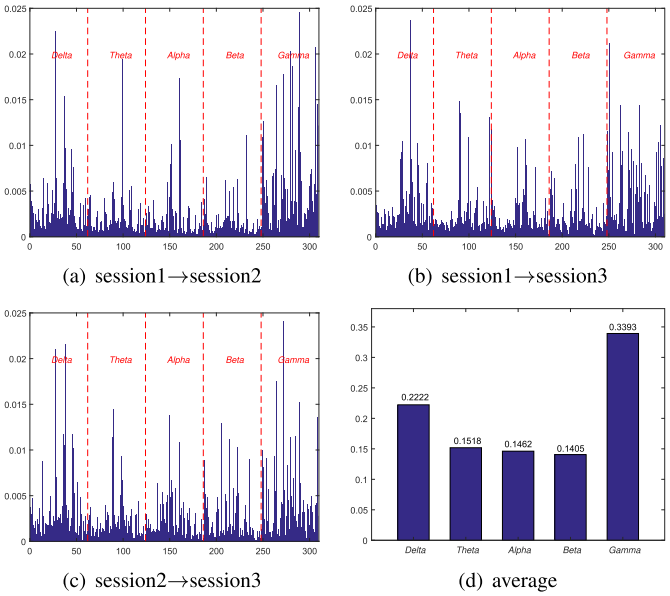
$$\psi(i) = \theta_j + \theta_{j+q} + \cdots + \theta_{j+(p-1)\times q}. \tag{22}$$

In SEED-IV, there are 62 channels and 5 frequency bands. Therefore, we respectively set p to 5 and q to 62 in equations (21) and (22). As shown by Fig. 6(a)-(c), the lengths of the vertical lines represent the different importance values of EEG features, which respectively correspond to the three cross-session emotion recognition tasks. In Fig. 6(d), we use the bar plots to represent the importance values of different frequency bands, which are explicitly provided on the top of bars. Obviously, features extracted from the *Gamma* band contribute the most in differentiating emotional states; therefore, we consider the *Gamma* frequency band as the most important one in emotion recognition. Based on the formula (22), the



(a) Accuracy vs. (α, β). (b) Accuracy vs. (α, γ). (c) Accuracy vs. (β, γ).

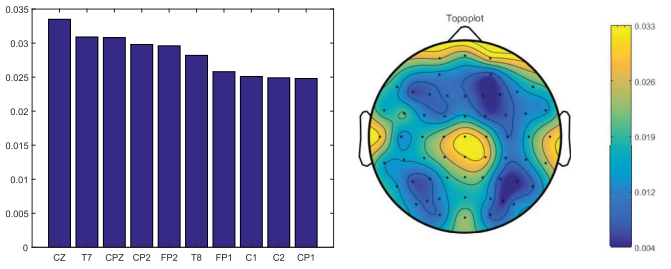
Fig. 5. Parameter sensitivity analysis on the case of 'subject 2: session 1→ session 2'.



(a) session1→session2 (b) session1→session3

(c) session2→session3 (d) average

Fig. 6. The importance analysis of EEG features and frequency bands.



(a) top 10 channels (b) average

Fig. 7. The importance analysis of EEG channels.

importance values of different EEG channels are obtained. In Fig. 7(a), we show the top 10 most important channels in emotion recognition. To get more meaningful insights into the discriminative abilities of different brain regions for emotion recognition, we present the EEG channel importance values by the form of brain topology in Fig. 7(b). Based on our results, the left/right temporal, prefrontal, and (central) parietal lobes are considered to be more correlated to emotion expression. The above critical frequency bands and channels identification results are consistent with some existing studies [4], [34], [35].

IV. Discussion

To highlight the contributions of this work, this section provides discussions on OGSSL and some related models, including the projected clustering with adaptive neighbors (PCAN) [20], the structured optimal graph feature selection (SOGFS) [36] and the semi-supervised projection with graph optimization (SPGO) [37].

Generally, these four models respectively unify the optimal graph learning with different graph-based learning tasks including the semi-supervised label propagation [37], feature selection [36] and spectral clustering [20]. OGSSL couples the optimal graph learning with both semi-supervised label propagation and feature selection. These joint learning models dynamically optimize the graph similarity matrix to accurately characterize the sample connections, which further facilitates the graph-based learning tasks.

These models differ mainly in the following aspects. 1) The learning paradigms are different. PCAN is an unsupervised learning model by unifying the graph construction and spectral clustering together. SOGFS is an unsupervised graph-based feature selection model. Though both SPGO and OGSSL are semi-supervised classification models, they use different methods to achieve the optimal graph learning. In SPGO, a desirable graph is learned to approximate a given graph which is previously constructed by rule-based method. In OGSSL, the optimal graph is directly learned from data. 2) The roles of the projection matrix $\mathbf{W}$ are not exactly the same. In both PCAN and SPGO, $\mathbf{W}$ has only the dimensionality reduction ability. From the model perspective, $\mathbf{W}$ in SOGFS and OGSSL has the dual roles of joint subspace learning and feature selection. However, $\mathbf{W}$ in the present work has special functions, *i.e.*, quantitatively identifying the critical EEG frequency bands and brain regions in emotion expression. 3) The settings of the label indicator matrix $\mathbf{F}$ are different. In PCAN and SOGFS, the indicator matrix is real-valued and is solved by eigen-value decomposition. Accordingly, a post-processing step (e.g., k-means or spectral rotation [38]) is necessary to perform discretization to obtain the final cluster assignments. In SPGO, without the orthogonal constraint, the label indicator matrix can be directly solved; however, the discretization postprocessing is also necessary. In OGSSL, with the help of non-negative and row-normalization constraints, the emotion recognition results can be obtained directly by locating the maximum value in each row of $\mathbf{F}$. Consequently, a different way was proposed to optimization $\mathbf{F}$ row-wisely.

V. Conclusion

In this paper, an Optimal Graph coupled Semi-Supervised Learning (OGSSL) model was proposed for EEG-based emotion recognition. OGSSL combined the optimal graph learning and the emotional label estimation of unlabeled EEG samples together in a unified framework. The graph similarity matrix was dynamically updated based on the sparse subspace representations of EEG samples. Extensive results on benchmark EEG data set demonstrated that 1) the joint learning mode effectively improved the emotion recognition performance; 2) the row-sparsity property of the projection matrix had promising ability in selecting discriminative EEG features for emotion recognition; and 3) the normalized ℓ_2-norm of each row of the projection matrix provided us with a quantitative way to identify the affective activation patterns (*i.e.*, critical EEG frequency bands and brain regions) in emotion recognition. In the present work, the affective activation patterns were explored by considering the discriminative ability of each EEG feature dimension with respect to all emotional states. In the future, we will investigate the label-specific activation patterns; that is, the critical frequency bands and brain regions in expressing a specific emotional state.

References

[1] Z. He *et al.*, "Advances in multimodal emotion recognition based on brain–computer interfaces," *Brain Sci.*, vol. 10, no. 10, p. 687, Sep. 2020.

[2] J. Li, N. Thakor, and A. Bezerianos, "Brain functional connectivity in unconstrained walking with and without an exoskeleton," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 730–739, Mar. 2020.

[3] X. Jin *et al.*, "CTNN: A convolutional tensor-train neural network for multi-task brainprint recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 103–112, 2021.

[4] Y. Peng, F. Qin, W. Kong, Y. Ge, F. Nie, and 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.*, early access, May 21, 2021, doi: 10.1109/TCDS.2021.3082803.

[5] W. L. Zheng and B.-L. Lu, "A multimodal approach to estimating vigilance using EEG and forehead EOG," *J. Neural Eng.*, vol. 14, no. 2, Feb. 2017, Art. no. 026017.

[6] Z. Pei, H. Wang, A. Bezerianos, and J. Li, "EEG-based multiclass workload identification using feature fusion and selection," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–8, 2021.

[7] Y. Jiang *et al.*, "Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2270–2284, Dec. 2017.

[8] Q. She, J. Zou, M. Meng, Y. Fan, and Z. Luo, "Balanced graph-based regularized semi-supervised extreme learning machine for EEG classification," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 4, pp. 903–916, Apr. 2021.

[9] G. Zhang and A. Etemad, "Deep recurrent semi-supervised EEG representation learning for emotion recognition," in *Proc. 9th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Sep. 2021, pp. 1–8.

[10] J. Luo, M. Wu, Z. Wang, Y. Chen, and Y. Yang, "Progressive low-rank subspace alignment based on semi-supervised joint domain adaption for personalized emotion recognition," *Neurocomputing*, vol. 456, pp. 312–326, Oct. 2021.

[11] Y. Dan, J. Tao, J. Fu, and D. Zhou, "Possibilistic clustering-promoting semi-supervised learning for EEG-based emotion recognition," *Frontiers Neurosci.*, vol. 15, pp. 1–13, Jun. 2021.

[12] D. H. Kim, M. K. Lee, D. Y. Choi, and B. C. Song, "Multi-modal emotion recognition using semi-supervised learning and multiple neural networks in the wild," in *Proc. 19th ACM Int. Conf. Multimodal Interact.*, Nov. 2017, pp. 529–535.

[13] Q. She *et al.*, "Safe semi-supervised extreme learning machine for EEG signal classification," *IEEE Access*, vol. 6, pp. 49399–49407, 2018.

[14] J. Jin *et al.*, "A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 20–29, 2022.

[15] C. Dai *et al.*, "Brain EEG time-series clustering using maximum-weight clique," *IEEE Trans. Cybern.*, vol. 52, no. 1, pp. 271–357, Jan. 2022.

[16] G. Kalantar and A. Mohammadi, "Graph-based dimensionality reduction of EEG signals via functional clustering and total variation measure for BCI systems," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 4603–4606.

[17] Y. Peng, B.-L. Lu, and S. Wang, "Enhanced low-rank representation via sparse manifold adaption for semi-supervised learning," *Neural Netw.*, vol. 65, pp. 1–17, May 2015.

[18] H. Gan, Z. Li, W. Wu, Z. Luo, and R. Huang, "Safety-aware graph-based semi-supervised learning," *Expert Syst. Appl.*, vol. 107, no. 1, pp. 243–254, Oct. 2018.

[19] Y. Peng, W. Kong, F. Qin, and F. Nie, "Manifold adaptive kernelized low-rank representation for semisupervised image classification," *Complexity*, vol. 2018, pp. 1–11, Jan. 2018.

[20] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 977–986.

[21] F. Nie, X. Wang, M. Jordan, and H. Huang, "The constrained Laplacian rank algorithm for graph-based clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2016, pp. 1969–1976.

[22] Y. Peng, Q. Li, W. Kong, F. Qin, J. Zhang, and A. Cichocki, "A joint optimization framework to semi-supervised RVFL and ELM networks for efficient data classification," *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 106756.

[23] J. Zheng, P. Yang, S. Chen, G. Shen, and W. Wang, "Iterative re-constrained group sparse face recognition with adaptive weights learning," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2408–2423, May 2017.

[24] M. Yu *et al.*, "A review of EEG features for emotion recognition," *Scientia Sinica Inf.*, vol. 49, no. 9, pp. 1097–1118, Sep. 2019.

[25] Y. Peng, X. Zhu, F. Nie, W. Kong, and Y. Ge, "Fuzzy graph clustering," *Inf. Sci.*, vol. 571, pp. 38–49, Sep. 2021.

[26] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019.

[27] X. Chen, G. Yuan, F. Nie, and J. Z. Huang, "Semi-supervised feature selection via rescaled linear regression," in *Proc. Int. J. Conf. Artif. Intell.*, 2017, pp. 1525–1531.

[28] X. Chen, G. Yuan, F. Nie, and Z. Ming, "Semi-supervised feature selection via sparse rescaled linear square regression," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 1, pp. 165–176, Jan. 2020.

[29] P. Song, W. Zheng, Y. Yu, and S. Ou, "Speech emotion recognition based on robust discriminative sparse regression," *IEEE Trans. Cogn. Develop. Syst.*, vol. 13, no. 2, pp. 343–353, Jun. 2021.

[30] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.

[31] F. Nie, H. Huang, X. Cai, and C. Ding, "Efficient and robust feature selection via joint $\ell_{2,1}$-norms minimization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 23, 2010, pp. 1813–1821.

[32] J. Xu, B. Tang, H. He, and H. Man, "Semisupervised feature selection based on relevance and redundancy criteria," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 9, pp. 1974–1984, Sep. 2017.

[33] W. Zhong, X. Chen, F. Nie, and J. Z. Huang, "Adaptive discriminant analysis for semi-supervised feature selection," *Inf. Sci.*, vol. 566, pp. 178–194, Aug. 2021.

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

[35] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 417–429, Jul./Sep. 2017.

[36] F. Nie, W. Zhu, and X. Li, "Structured graph optimization for unsupervised feature selection," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 3, pp. 1210–1222, Mar. 2021.

[37] F. Nie, X. Dong, and X. Li, "Unsupervised and semisupervised projection with graph optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1547–1559, Apr. 2020.

[38] W. Huang, Y. Peng, Y. Ge, and W. Kong, "A new kmeans clustering model and its generalization achieved by joint spectral embedding and rotation," *PeerJ Comput. Sci.*, vol. 7, p. e450, Mar. 2021.