SIFIAE: an adaptive emotion recognition model with EEG feature-label inconsistency consideration

Yikai Zhang[a], Yong Peng[a,b,*], Junhua Li[c], Wanzeng Kong[a,b]

[a]*School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, 310018, Zhejiang Province, China*
[b]*Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province, Hangzhou, 310018, Zhejiang Province, China*
[c]*School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, U.K.*

Abstract

Background: A common but easily overlooked affective overlap problem has not been received enough attention in electroencephalogram (EEG)-based emotion recognition research. In real life, affective overlap refers to the current emotional state of human being is sometimes influenced easily by his/her historical mood. In stimulus-evoked EEG collection experiment, due to the short rest interval in consecutive trials, the inner mechanisms of neural responses make subjects cannot switch their emotion state easily and quickly, which might lead to the affective overlap. For example, we might be still in sad state to some extent even if we are watching a comedy because we just saw a tragedy before. In pattern recognition, affective overlap usually means that there exists the feature-label inconsistency in EEG data.
New method: To alleviate the impact of inconsistent EEG data, we introduce a variable to adaptively explore the sample inconsistency in emotion recognition model development. Then, we propose a semi-supervised emotion recognition model for joint sample inconsistency and feature importance exploration (SIFIAE). Accordingly, an efficient optimization method to SIFIAE model is proposed.
Results: Extensive experiments on the SEED-V dataset demonstrate the effectiveness of SIFIAE. Specifically, SIFIAE achieves 69.10%, 67.01%, 71.5%,

*Corresponding author
Email address: yongpeng@hdu.edu.cn (Yong Peng)

73.26%, 72.07% and 71.35% average accuracies in six cross-session emotion recognition tasks.
Conclusions: The results illustrated that the sample weights have a rising trend in the beginning of most trials, which coincides with the affective overlap hypothesis. The feature importance factor indicated the critical bands and channels are more obvious compared with some models without considering EEG feature-label inconsistency.

Keywords:
Emotion recognition, affective overlap, feature-label inconsistency, automatic weighting, EEG, pure semi-supervised learning.

---

1. Introduction

Emotion Recognition is the key to achieving human-computer intelligent interaction. In the past decades, many studies have used different data for emotion recognition, such as facial expression [1], text [2], speech [3] and physiological signals [4]. Compared with other data sources, physiological signals has received increasing attention since its unique properties of high time resolution and non-camouflage [5, 6].

Electroencephalogram (EEG), as the most widely used physiological signals in brain-computer interface, reflects the activity of cerebral cortex to a certain extent [7, 8, 9]. Many affective neural mechanisms studies show that brain region activation is closely related to emotion and cognition [10, 11]. Based on this affective knowledge, lots of EEG-based methods are proposed for high effective emotion recognition [12, 13, 14, 15, 16]. For example, Liu *et al.* proposed an attention mechanism to give the spatial information to input signals to enrich the information of EEG data [17]. In [18], Wang et al. applied the self-supervised method with convolutional neural network to improve the efficiency of resource usage in the task of EEG-based emotion recognition. Wang *et al.* proposed a multi-modal domain adaptive variational autoencoder to reduce the amount of calibration samples. Olamat *et al.* designed a multi-variate empirical mode decomposition to decompose EEG signals and used deep learning methods to recognize emotions. Zheng *et al.* used deep belief network find the channels on temporal lobe and prefrontal lobe are critical to emotion recognition [19]. Song *et al.* used dynamical graph neural network for EEG feature learning and emotion recognition, in which the graph aims to learn intrinsic relationship among different

channels [20]. Li et al. combined the bidirectional long short term memory network and R2G-STNN to learn discriminative spatial-temporal EEG features [21]. Multimodal feature fusion is also an effective way to improve the performance of emotion recognition [22]. Zheng *et al.* designed a multimodal emotion recognition framework termed EmotionMeter to jointly utilize the information of EEG and eye movements [23]. Zhang *et al.* used kernel matrices to learning the comprehensive information from multimodal physiological signals such as EEG, galvanic skin response and electromyography, and then a deep network was utilized to learn the task-specific representation for each signals [24]. In [25], Wu *et al.* introduced a new experimental paradigm that incorporates odors at various stages of video-evoked emotions to explore the effectiveness of olfactory-enhanced videos in inducing subjects' emotions.

However, there is a common but easily overlooked problem during the emotion recognition processing, *i.e.*, the emotion state of human being is sometimes influence easily by his/her historical mood. In other words, the emotions are persistent and cannot quickly switch. In order to understand and quantify this problem more easily, we remove the scenario from real life to the existing experimental paradigm of EEG-based affective BCI (aBCI) [26]. Stimulus-evoked EEG is a classical experimental paradigm of aBCI, whose main process is to induce the emotion of subjects by displaying film clips or images as the emotional stimulus [19]. In this paper, we will consider this affective overlap problem in stimulus-evoked emotion EEG paradigm.

Because emotions are persistent and cannot quickly switch, it will easily lead to affective overlap phenomenon when the rest interval of two adjacent trials is not enough and especially two consecutive trials correspond to opposite emotional states, *i.e.*, the current trial may contain emotions induced by previous trials. As shown in Fig. 1, we consider the situation that the subject watched the film clip with sad theme and then watched the film clip with happy theme after a short break (usually several seconds). Because the subject was asked to immersed in the film clip with sad theme, he could not clear his sad mood in a few seconds. We can naturally imagine that in the subsequent trial, sad mood act as the background emotion of happy mood. Thus, the label of the K-th trial is annotated as happy but the obtained EEG features are the mixture of both sad and happy states. From the perspective of pattern recognition, this affective overlap phenomenon means that there exists feature-label inconsistency in collected EEG data [27]. Therefore, how to explore the inconsistency of stimulus-evoked EEG data and then suppress its side effect are significant to improve the performance of emotion

recognition.

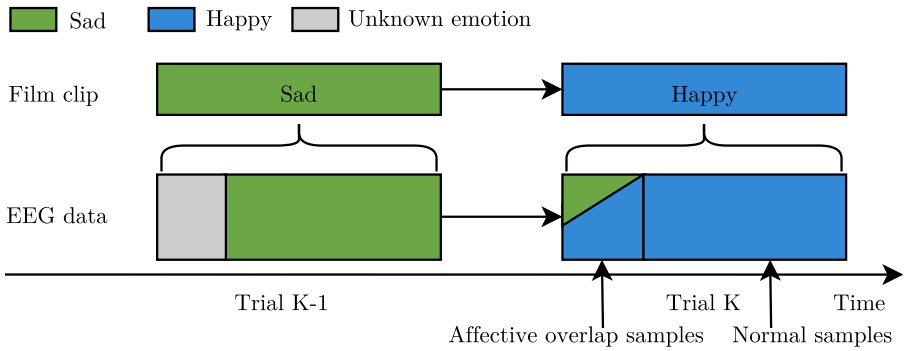


Figure 1: A schematic diagram of affective overlap phenomenon.

The other property we should consider is that EEG signals are multi-rhythm and multi-channel inherently. The current studies showed that different rhythms and channels have different contributions in emotion recognition, which indicates that some EEG features are redundant or have low contributions [19]. To our best knowledge, there is no existing aBCI literature simultaneously considering the sample inconsistency and feature importance.

In this paper, we proposed a model termed SIFIAE to jointly and adaptively explore the feature importance and sample inconsistency. Specifically, we introduced a sample inconsistency factor and a feature importance factor to measure the degree of feature-label inconsistency of each sample and the contribution of each feature, respectively. The sample inconsistency is calculated based on the model approximation error and the feature importance is calculated by the normalized $\ell_{2,1}$-norm of each row of the projection matrix. Then the sample inconsistency and feature importance can help model to self-weighting samples and features to identify the low-inconsistency samples and high-contribution features in emotion recognition. Besides, we implement our SIFIAE model within the pure semi-supervised learning framework that is more suitable for practical emotion recognition applications [28].

The rest of paper is organized as follows. In section 2, the model formulation and optimization of SIFIAE are introduced. Afterwards, a pure semi-supervised cross-session emotion recognition experiment is reported in section 3. In section 4, some discussion about our model can be obtained. Finally, we conclude our paper in section 5.

Notations. In this paper, we use $\mathbf{m}_i$ and $\mathbf{m}^i$ to denote the i-th column

and i-th row of matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$. The $\ell_{2,1}$-norm of $\mathbf{M}$ is defined as $\|\mathbf{M}\|_{2,1} = \sum_{i=1}^{m}\sqrt{\sum_{j=1}^{n} m_{ij}^2} = \sum_{i=1}^{m}\|\mathbf{m}^i\|_2$.

2. Model Formulation

Considering the feature importance and sample inconsistency simultaneously, we show the main structure of our proposed SIFIAE model in Fig. 2.

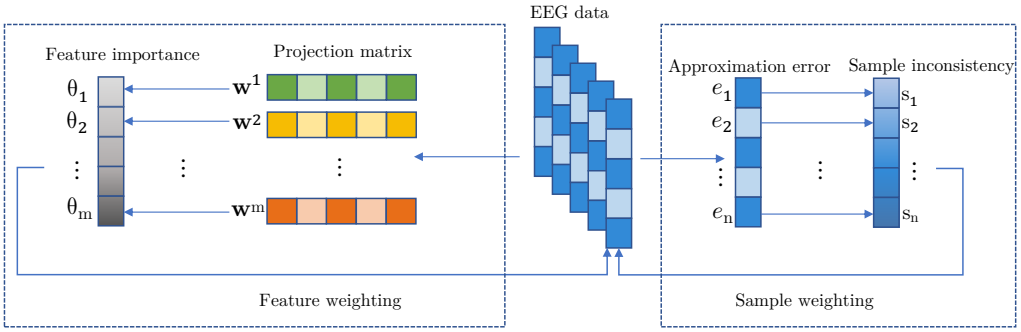


Figure 2: The overall structure of our proposed SIFIAE model. Feature weighting (left rectangle) and sample weighting (right rectangle) are two main blocks in SIFIAE. In the feature weighting block, the normalized $\ell_{2,1}$-norm of the i-th row of projection matrix is calculated to measure the importance of the i-th feature θ_i. In the sample weighting block, the inconsistency feactor s_i of the i-th sample is calculated by the model approximation error e_i of the i-th sample. Finally, feature importance factor $\boldsymbol{\theta}$ and sample inconsistency factor $\mathbf{s}$ are used to weighting EEG data.

In the pure semi-supervised learning framework, we are given an EEG training data $[\mathbf{X}_l, \mathbf{X}_u] \in \mathbb{R}^{d\times n}$ consisting of l labeled and u unlabeled d-dimensional samples and an EEG testing data $\mathbf{X}_t \in \mathbb{R}^{d\times t}$ consisting of t unlabeled samples. $[\mathbf{Y}_l, \mathbf{Y}_u] \in \mathbb{R}^{n\times c}$ is a label matrix where $\mathbf{Y}_l \in \mathbb{R}^{l\times c}$ is the label indicator matrix of labeled samples $\mathbf{X}_l$ and $\mathbf{Y}_u \in \mathbb{R}^{u\times c}$ is an unknown label matrix corresponding to unlabeled samples $\mathbf{X}_u$. Similarly, $\mathbf{Y}_t \in \mathbb{R}^{t\times c}$ is an unknown label matrix of testing data $\mathbf{X}_t$. Here, $\mathbf{Y}_l$, $\mathbf{Y}_u$ and $\mathbf{Y}_t$ are formed by one-hot encoding to represent the membership degree of each sample to c emotion states. To evaluate the feature-label inconsistency of EEG samples, we introduce a weighting vector $\mathbf{s} = (s_1; s_2; \cdots; s_n) \in \mathbb{R}^n$ where $\mathbf{s}$ satisfies the normalization constraint (*i.e.*, $\mathbf{s}^T\mathbf{1} = 1$) and non-negative constraint (*i.e.*, $s_i|_{i=1}^n > 0$). If s_i is larger, it means that the i-th sample is less inconsistent;

otherwise, it means that the i-th sample is more likely to be an affective overlap sample. Based on the semi-supervised linear regression framework, we formulate the objective function of SIFIAE as

$$\begin{aligned} \min_{\mathbf{s},\mathbf{W},\mathbf{b},\mathbf{Y}_u} \sum_{i=1}^{n} s_i^r \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 + \lambda\|\mathbf{s}^{\frac{r}{2}}\|_2^2 + \gamma\|\mathbf{W}\|_{2,1}^2, \\ s.t.\ \mathbf{s}^T\mathbf{1} = 1, s_i > 0, \mathbf{Y}_u\mathbf{1}_c = \mathbf{1}_c, y_{ij} \geq 0, \end{aligned} \tag{1}$$

where $r > 1$ is a weighting factor to control the sensitivity of affective overlap samples, $\mathbf{b} \in \mathbb{R}^m$ represents the bias vector, $\mathbf{y}^i \in \mathbb{R}^c$ is the i-th label vector, $\mathbf{W} \in \mathbb{R}^{m\times c}$ is a projection matrix to bridge the EEG feature space and its label space and $\mathbf{1}_c$ is an all-one column vector with length c. From the first term in objective function (1), we understand the connection between s_i and the model approximation error of sample $\mathbf{x}_i$. Specifically, the smaller the approximation error of the i-th sample $e_i = \|\mathbf{x}_i^T\mathbf{W}+\mathbf{b}^T-\mathbf{y}^i\|_2^2$, the larger the weight of the i-th sample. Otherwise, when e_i is larger, we think the feature-label inconsistency of the i-th sample is more significant and SIFIAE will accordingly assign a lower weight to reduce its side impact in model training process. The second term in objective function (1) is a regularization term of $\mathbf{s}$ to prevent the appearance of trivial solution in $\mathbf{s}$ (*i.e.*, only one element is one and all the others are zeros in $\mathbf{s}$). The third term is an $\ell_{2,1}$-norm regularizer defined on $\mathbf{W}$, which aims to enforce the row sparsity of $\mathbf{W}$ and then feature ranking can be naturally achieved. Based on the $\ell_{2,1}$-norm based feature selection theory [29], the quantitative importance of the i-th feature $\theta_i|_{i=1}^m$ can be obtained by

$$\theta_i = \frac{\|\mathbf{w}^i\|_2}{\sum_{i=1}^m \|\mathbf{w}^i\|_2}. \tag{2}$$

Based on the above analysis, when SIFIAE is fitted by given EEG data, not only the sample inconsistency factor $\mathbf{s}$ is obtained, but also the feature importance factor $\boldsymbol{\theta} \triangleq (\theta_1; \theta_2; \cdots; \theta_m) \in \mathbb{R}^m$ is obtained.

Since there have four variables in objective function (1), whose optimization process can be divided into the below four blocks.

- *Update* $\mathbf{s}$. By fixing $\mathbf{W}$, $\mathbf{b}$, and $\mathbf{Y}_u$, we transform problem (1) as

$$\min_{\mathbf{s}^T\mathbf{1}=1, s_i>0} \sum_{i=1}^{n} s_i^r \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 + \lambda \sum_{i=1}^{n} s_i^r. \tag{3}$$

We denote $d_i = \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 + \lambda = e_i + \lambda$. Then the Lagrangian function of objective function (3) can be written as

$$\mathcal{L}(\mathbf{s}) = \sum_{i=1}^{n} s_i^r d_i - \alpha(\mathbf{s}^T\mathbf{1} - 1), \tag{4}$$

where α is an Lagrangian multiplier. Taking the derivative of $\mathcal{L}(\mathbf{s})$ with respect to s_i and setting it to zero, we have $s_i = \left(\frac{\alpha}{rd_i}\right)^{\frac{1}{r-1}}$. Combined with the constraint $\mathbf{s}^T\mathbf{1} = 1$, the closed-form solution of $\mathbf{s}$ can be obtained by

$$s_i = \frac{1}{d_i^{\frac{1}{r-1}} \sum_{j=1}^{n} (\frac{1}{d_j})^{\frac{1}{r-1}}}. \tag{5}$$

- *Update* $\mathbf{b}$. When we fix $\mathbf{s}$, $\mathbf{W}$ and $\mathbf{Y}_u$, the objecitve function associated with $\mathbf{b}$ is

$$\begin{aligned} \mathcal{O}(\mathbf{b}) &= \sum_{i=1}^{n} s_i^r \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 \\ &= \mathrm{Tr}((\mathbf{X}^T\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y})^T\mathbf{\Lambda}(\mathbf{X}^T\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y})), \end{aligned} \tag{6}$$

where $\mathbf{\Lambda} \in \mathbb{R}^{n\times n}$ is a diagonal matrix with its i-th diagonal element $\Lambda_{ii} = s_i^r$. Taking the derivative of objective function (6) with respect to variable $\mathbf{b}$ and setting it to zero, we have

$$\mathbf{b} = (\mathbf{W}^T\mathbf{X} - \mathbf{Y}^T)\mathbf{\Lambda}\mathbf{1}. \tag{7}$$

- *Update* $\mathbf{W}$. When we update $\mathbf{W}$, we first rewrite problem (1) as

$$\min_{\mathbf{W}} \sum_{i=1}^{n} s_i^r \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 + \gamma\|\mathbf{W}\|_{2,1}^2. \tag{8}$$

Based on the usual way to deal with $\ell_{2,1}$-norm regularization [29], we introduce a diagonal matrix $\mathbf{Q} \in \mathbb{R}^{m\times m}$ with

$$q_{ii} = \frac{\sum_{j=1}^{m} \sqrt{\|\mathbf{w}^j\|_2^2 + \epsilon}}{\sqrt{\|\mathbf{w}^i\|_2^2 + \epsilon}}, \tag{9}$$

where ϵ is a small constant to avoid the non-differentiable problem. Then we rewrite $\|\mathbf{W}\|_{2,1}^2 = \mathrm{Tr}(\mathbf{W}^T\mathbf{Q}\mathbf{W})$. The closed-form solution of $\mathbf{W}$ can be obtained by setting the derivative of objective function (8) w.r.t. $\mathbf{W}$ to zero; that is

$$\mathbf{W} = (\mathbf{X}\mathbf{H}\mathbf{X}^T + \lambda\mathbf{Q})^{-1}\mathbf{X}\mathbf{H}\mathbf{Y}, \tag{10}$$

where $\mathbf{H} = \mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{1}\mathbf{1}^T\mathbf{\Lambda}$.

- *Update* $\mathbf{Y}_u$. Since $\mathbf{s}$, $\mathbf{W}$ and $\mathbf{b}$ are fixed, the pseudo-label indicator vector $\mathbf{y}^i$ of the $i|_{l+1}^n$-th unlabeled EEG sample can be obtained by solving the following objective function

$$\min_{\mathbf{y}^i\mathbf{1}_c=1,\mathbf{y}^i\geq\mathbf{0}} \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2, \tag{11}$$

  which defines an Euclidean projection on a simplex constraint. The standard optimization can be found in [30].

After the SIFIAE model was learned, we obtain the optimally fitted variables $\mathbf{W}^*$, $\mathbf{b}^*$ and $\mathbf{s}^*$, based on which we can calculate the emotion label indicator vector $\mathbf{y}^j$ of the $j|_{j=1}^t$-th test EEG sample in $\mathbf{X} \in \mathbb{R}^{m\times t}$ by

$$\mathbf{y}^j = \mathbf{x}_j^T\mathbf{W}^* + (\mathbf{b}^*)^T. \tag{12}$$

Naturally, the emotion state corresponding to the maximal value in $\mathbf{y}_t^j$ is the predicted emotion label of the j-th test EEG sample.

To easier capture the workflow of our proposed SIFIAE model, we summarize its learning and testing process in Algorithm 1.

3. Experiments and analysis

In this section, we try to answer the following three questions through experiments. 1) Whether the joint exploration of feature importance and sample inconsistency can improve the accuracy of cross-session emotion recognition? 2) Whether the sample inconsistency factor $\mathbf{s}$ can alleviate the impact of affective overlap samples? 3) How SIFIAE explores the affective activation pattern by considering the EEG feature-label inconsistency? Below we first introduce the data set and experimental setup, and then analyze the experimental results point by point in response to the above mentioned three questions.

Algorithm 1 The proposed SIFIAE model

Input: Training EEG feature matrix $\mathbf{X} = [\mathbf{X}_l, \mathbf{X}_u]$, known emotion label matrix $\mathbf{Y}_l$, testing EEG feature matrix $\hat{\mathbf{X}}$, regularization parameters λ and γ, weighting factor r;

Output: Pseudo-label matrix $\mathbf{Y}_u$ and $\mathbf{Y}_t$, sample inconsistency factor $\mathbf{s}$, projection matrix $\mathbf{W}$;

Initialize $\mathbf{Y}_u = \frac{\mathbf{1}_c \mathbf{1}_c^T}{c}$ and $\mathbf{s} = \mathbf{1}_n$;
while not converged **do**
  Update $\mathbf{W}$ by equation (10);
  Update the diagonal matrix $\mathbf{Q}$ by equation (9);
  Update $\mathbf{b}$ by equation (7);
  Update $\mathbf{s}$ by solving problem (5);
  Update $\mathbf{Y}_u$ by solving problem (11);
end while
Calculate $\mathbf{y}^j|_{j=1}^t$ for each $\hat{\mathbf{x}}_j \in \hat{\mathbf{X}}$ by equation (12).

3.1. Data set

In this paper, we use the SEED-V [31] emotional EEG data set to evaluate the performance of our proposed SIFIAE model. SEED-V includes 16 subjects, which participated in the EEG data collection experiment three times. EEG data of each subject can be divided into three sessions, and each session contains 15 trials. As shown in Fig. 3, in each trial, subjects first receive a 15-second hint of start including the background of the stimulus materials and the emotion the film clips to be displayed. Then, they watch a film clip which lasts about 2-4 minutes for emotion-inducing. Finally, subjects have 15 or 30 seconds for self-evaluation and rest, depending on the types of film clips. If the theme of stimulus film clips are happiness, neutrality or sadness, the rest interval between two adjacent trials is 15 seconds and the rest time for disgust and fear is 30 seconds. The detailed description for the SEED-V data set can be found in [31]. Similar to the data organization in [32], we arrange the EEG data of each subject in each session as a two-dimensional matrix composed of samples and differential entropy features.

3.2. Performance analysis

In this subsection, we conducted a series of comparative experiments to answer the first question. For specific, we selected some popular semi-supervised classification models based on linear square regression, including

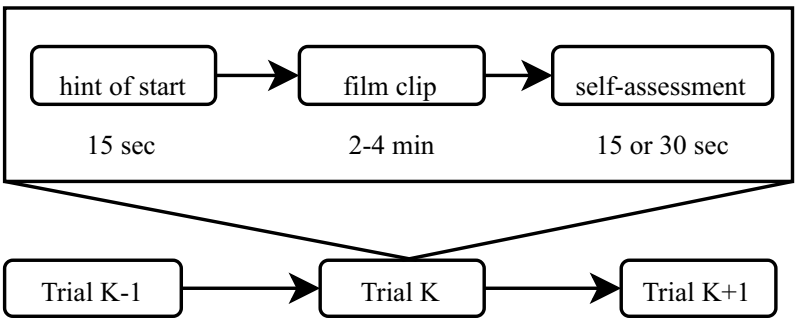


Figure 3: The general stimulus-evoked paradigm in SEED-V [31].

the semi-supervised least square regression (sLSR), the rescaled linear square regression (RLSR) [29], the robust semi-supervised least square regression (RSSLSR) [33] and the semi-supervised joint sample and feature importance evaluation (sJSFE) [34].

Data settings. Then we evaluate the performance of different models in a pure semi-supervised learning manner. Specifically, the unlabeled test EEG samples are unseen during the semi-supervised model learning process. In the following experiments, we informally used the notation '{labeled training data, unlabeled training data}→test data' to explain how the training and test data were set. Taking the '{s1, s2}→s3' as an example, for each subject, we used the EEG samples and the associated emotion labels from the first session and EEG samples without labels from the second session as training data. The EEG samples from the third session are set as test data, meaning that we aim at predicting their emotion states as accurately as possible. It is noteworthy that the EEG data from session 3 cannot be used in the model learning process. From the perspective of emotion recognition in practical scenarios, both the labeled training data and the unlabeled training data are obtained in advance, which coincides with the motivation of semi-supervised learning that unlabeled data is used to facilitate the ability of learning model to capture the underlying data properties. However, the test data cannot be collected in advance, and therefore cannot be involved in model training, which is used for evaluating the out-of-sample extension ability of learning models. Compared with the semi-supervised learning in transductive paradigm employed by our previous works [32, 34], pure semi-supervised learning is more closer to real applications in emotion recognition.

Parameter settings. To be fair, the regularization parameters in each

model were searched from $\{10^{-3}, 10^{-2}, \cdots, 10^{3}\}$. In sJSFE, the self-paced parameter was initialize to 0.001 and the step parameter was tuned from $\{1.1, 1.2, \cdots, 3\}$. In SIFIAE, the weighting factor r was tuned with a grid search from $\{1.1, 1.2, \cdots, 2\}$.

Performance result. Based on the above settings, the accuracies of cross-session emotion recognition on SEED-V are respectively reported in Tables 1, 2, and 3, where the best accuracy in each case is highlighted in bold. It is observed that our proposed SIFIAE model ranks the first in '{s2, s3}→s1', '{s3, s2}→s1', '{s1, s3}→s2', '{s3, s1}→s2', '{s1, s2}→s3' and '{s2, s1}→s3' for 10 times, 9 times, 11 times, 10 times, 12 times, and 7 times, respectively. To be specific, the average accuracies of SIFIAE respectively exceed the second-best model by 3.76%, 2.72%, 2.85%, 4.85%, 4.25%, and 1.83%. The above results depict that the exploration of both sample inconsistency and feature importance effectively improve the performance of cross-session emotion recognition.

Table 1: Cross-session emotion recognition results (%) of different models on SEED-V. Take the EEG data from session 1 as test samples.

ID	{s2, s3}→s1					{s3, s2}→s1				
	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE
sub1	73.72	68.43	48.90	68.87	**87.67**	68.87	68.87	58.44	72.25	**72.69**
sub2	45.81	59.62	47.14	60.21	**65.79**	64.76	60.65	62.85	**68.43**	65.93
sub3	56.24	56.24	56.39	52.57	**60.79**	48.75	48.75	45.96	58.15	**63.14**
sub4	68.87	74.74	66.52	76.06	**80.91**	68.28	69.31	59.91	**80.91**	79.44
sub5	53.89	55.36	59.03	67.40	**69.31**	39.35	52.57	52.86	65.93	**75.62**
sub6	60.50	**63.14**	56.68	57.86	**63.14**	33.77	35.83	46.70	52.13	**58.88**
sub7	55.95	64.02	46.70	**67.69**	66.81	37.15	45.52	39.21	46.99	**75.04**
sub8	61.23	60.65	58.15	73.27	74.74	52.72	55.36	50.07	**69.60**	68.14
sub9	53.01	35.54	59.62	**72.69**	69.31	70.48	56.39	61.67	70.48	**73.72**
sub10	49.05	47.72	52.57	51.25	**58.74**	41.56	43.47	43.17	**60.65**	50.51
sub11	51.40	54.63	51.54	**67.40**	60.35	66.67	65.64	64.61	**79.59**	71.37
sub12	53.89	58.30	56.09	**77.68**	76.36	51.84	58.00	45.96	64.17	**70.78**
sub13	56.83	65.79	52.13	65.35	**73.27**	44.49	54.92	46.84	61.67	**67.11**
sub14	48.02	41.26	47.87	57.86	**65.93**	43.02	45.67	43.91	**49.93**	48.16
sub15	51.25	46.70	54.19	**63.14**	62.85	48.02	58.59	52.72	66.96	**71.95**
sub16	50.81	47.14	60.35	66.08	**69.60**	38.03	49.78	45.96	**60.79**	59.62
AVG.	55.65	56.20	54.62	65.34	**69.10**	51.11	54.33	51.30	64.29	**67.01**

3.3. Affective overlap analysis

Before answering the second question, we try to explain it in more detail. First, our goal is to demonstrate how SIFIAE alleviates the affective overlap phenomenon in emotion recognition through introducing the sample inconsistency factor **s**. Our underlying hypothesis is that the affective

Table 2: Cross-session emotion recognition results (%) of different models on SEED-V. Take the EEG data from session 2 as test samples.

ID	{s1, s3}→s2					{s3, s1}→s2				
	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE
sub1	42.33	65.99	48.43	74.12	**79.67**	65.62	67.65	63.03	**82.62**	80.59
sub2	57.49	65.80	57.49	65.80	**68.39**	85.77	76.89	79.48	**87.80**	83.73
sub3	49.17	61.37	49.54	**65.25**	63.77	55.82	64.70	51.02	**79.67**	77.08
sub4	79.11	83.36	70.98	86.32	**91.13**	70.24	61.74	66.36	59.33	**88.72**
sub5	56.75	69.32	51.02	**83.92**	73.57	37.89	46.58	62.48	63.96	**72.64**
sub6	58.04	61.00	58.04	62.85	**76.16**	30.68	26.43	37.71	55.45	**71.53**
sub7	73.20	**75.79**	59.15	70.61	71.35	55.45	57.67	43.81	**79.11**	71.90
sub8	58.23	63.40	49.35	**70.06**	69.69	77.45	77.45	48.98	**90.20**	80.41
sub9	63.77	58.04	47.32	**68.21**	**68.21**	50.65	60.81	56.01	54.90	**65.80**
sub10	34.38	38.45	44.55	51.20	**56.19**	39.93	41.40	43.25	52.13	**56.19**
sub11	32.90	47.13	42.14	56.75	**73.38**	43.81	49.72	46.40	55.82	**66.73**
sub12	65.80	74.49	54.71	77.63	**83.92**	70.06	71.16	75.60	77.63	**81.70**
sub13	77.63	73.57	62.66	77.63	**78.74**	75.97	67.84	54.16	70.06	**80.22**
sub14	48.43	52.87	46.40	62.48	**64.14**	60.44	55.45	41.96	50.09	**60.81**
sub15	68.76	68.76	65.62	**71.16**	65.62	51.39	55.08	52.13	**73.38**	67.10
sub16	44.73	50.65	60.07	54.34	**60.07**	43.81	61.55	43.81	62.48	**66.91**
AVG.	56.92	63.12	54.22	68.65	**71.50**	57.19	58.88	54.14	68.41	**73.26**

Table 3: Cross-session emotion recognition results(%) of different models on SEED-V. Take the EEG data from session 3 as test samples.

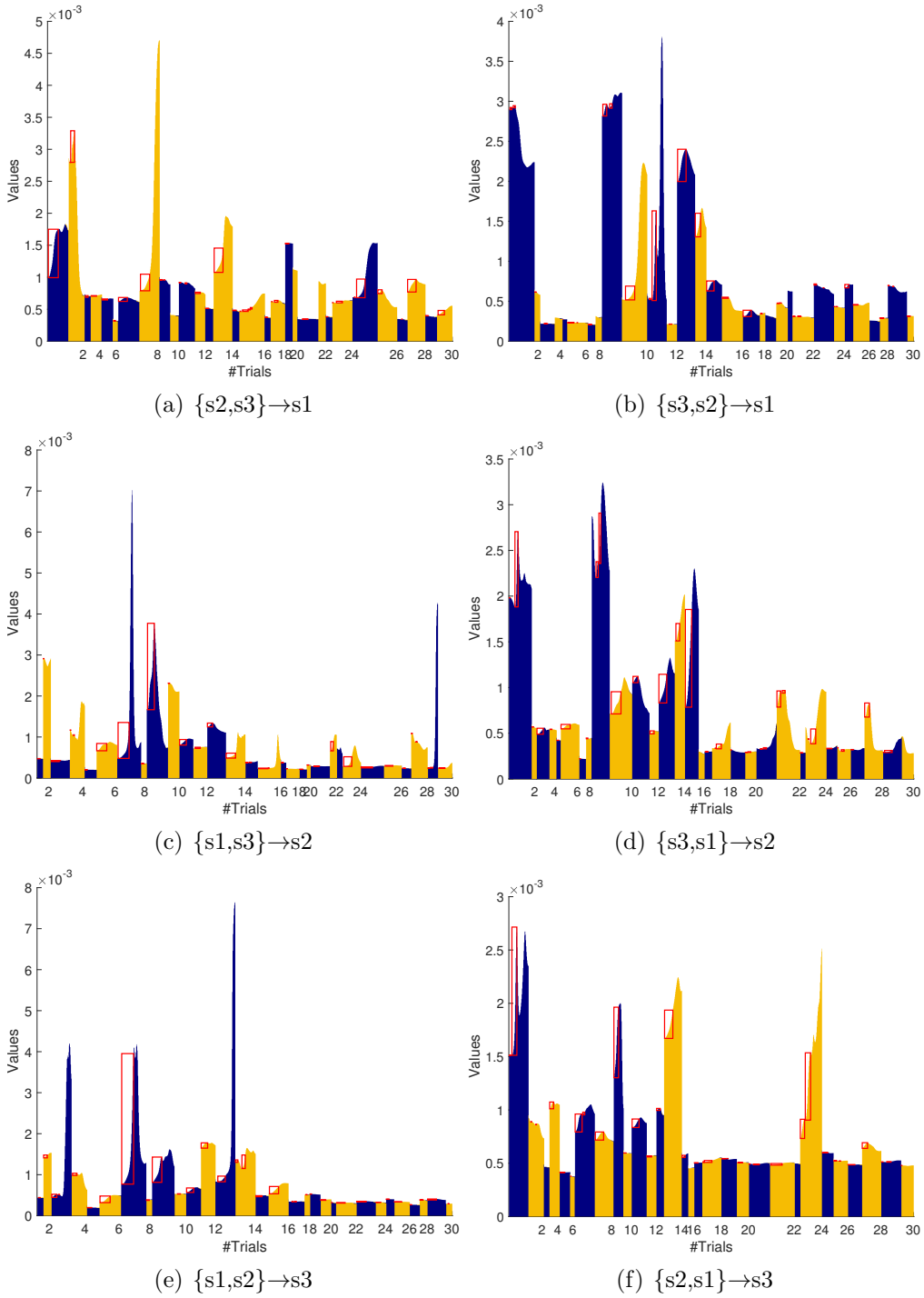
ID	{s1, s2}→s3					{s2, s1}→s3				
	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE	sLSR	RLSR	RSSLSR	sJSFE	SIFIAE
sub1	62.40	64.56	68.39	70.88	**73.54**	92.85	92.85	65.22	92.85	**92.85**
sub2	51.41	67.89	55.41	**74.88**	73.54	63.06	69.38	71.55	**86.19**	82.70
sub3	57.90	64.73	61.06	63.06	**70.88**	50.58	50.75	43.59	**71.55**	71.38
sub4	83.69	81.36	70.88	77.37	**86.69**	77.87	73.88	67.55	71.71	**80.20**
sub5	58.57	68.72	50.92	**80.87**	76.04	44.76	60.90	46.92	60.90	**61.40**
sub6	42.93	36.94	45.76	**56.57**	**56.57**	38.27	41.26	49.75	55.91	**65.72**
sub7	42.76	45.59	49.75	**71.21**	69.72	65.56	66.72	48.92	71.05	**85.86**
sub8	40.43	55.24	53.24	58.40	**66.56**	37.44	53.41	49.75	**78.37**	66.39
sub9	73.88	78.87	53.24	**84.53**	79.87	**75.21**	73.71	69.38	69.22	73.71
sub10	36.11	34.44	47.09	54.74	**56.07**	34.11	40.43	48.25	54.58	59.23
sub11	64.06	61.40	66.56	69.38	**78.37**	73.38	73.04	68.22	**76.71**	74.71
sub12	48.42	51.75	49.75	73.54	**85.52**	57.24	62.23	50.25	67.05	**67.72**
sub13	55.74	56.07	58.07	72.38	**76.04**	**78.87**	77.20	45.76	75.87	73.38
sub14	43.93	54.91	51.75	62.23	**67.72**	40.77	37.94	44.43	49.08	**64.06**
sub15	46.26	30.28	51.75	54.91	**71.05**	54.24	54.24	52.25	**71.38**	64.06
sub16	61.73	63.73	49.75	60.23	**64.89**	28.45	39.43	50.25	**59.90**	58.24
AVG.	54.39	57.28	55.21	67.82	**72.07**	57.04	60.46	54.50	69.52	**71.35**

overlap phenomenon will gradually disappear as the subjects immerse themselves in the new stimulus material (*e.g.*, film clips); therefore, the sample inconsistency in a trial will decrease and $\mathbf{s}$ will gradually increase. Besides, we should constrain our receptive field within a limited time period because we believe that the affective overlap phenomenon may not always exist in a whole trial. According to our understanding, we think that the samples in the beginning of a trial are more likely to have mixed emotions induced by both the previous and the current trials.

To intuitively observe the sample inconsistency, we show the learned sample inconsistency factor $\mathbf{s}$ in Fig. 4, where the six subfigures respectively correspond to the six different data setting cases in Tables 1, 2, and 3. In each case, we calculate the mean value of sample inconsistency factor $\mathbf{s}$s across the 16 subjects. Taking the emotion recognition task '{s2, s3}→s1' in Fig. 4(a) as an example, there are 30 trials in the two sessions (*i.e.*, session 2 and session 3) and we use the blue and yellow colors to differentiate the adjacent trials. Based on the consensus that a more significant affective overlap sample will lead to a more obvious feature-label inconsistency value (*i.e.*, a smaller s_i), we explicitly mark the increasing trend of s_is with a red rectangular box, which mostly correspond to the EEG samples at the beginning of a trial. In this case, we find that the learned sample inconsistency factors (*i.e.*, $\mathbf{s}$s) in 9 out of the 30 trials (Nos. 1, 2, 7, 8, 14, 25, 26, 28, 30) have significant increasing trends in the beginning of these trials. Besides, another 18 of the 30 trials have slight increasing trends in their learned samples inconsistency factors. For the other cross-session emotion recognition tasks, such increasing trends of learned $\mathbf{s}$s are also observed in the beginning of most trials, indicating that the affective overlap phenomenon does exist in adjacent trials and our proposed SIFIAE model finds an effective way to suppress its adverse effect in emotion recognition by adaptively quantifying the contributions of samples in model learning.

3.4. Affective activation pattern analysis

In this section, we investigate the spatial-frequency activation patterns derived by SIFIAE and two other models, RLSR and sJSFE, which did not take the feature-label inconsistency of EEG data into consideration. Such analysis aims to answer the third question raised in the beginning of Section 3. There are two prerequisites to obtain the affective activation patterns, *i.e.* the quantitative importance of each feature and the correspondence between the feature dimensions and frequency bands (channels). As expressed by



(a) {s2,s3}→s1

(b) {s3,s2}→s1

(c) {s1,s3}→s2

(d) {s3,s1}→s2

(e) {s1,s2}→s3

(f) {s2,s1}→s3

Figure 4: Visualization of the average sample inconsistency in different tasks. In each task, we calculated the mean value of the sample inconsistency factor **s** across the 16 subjects.

equation (2), the first prerequisite is satisfied and the feature importance factor $\boldsymbol{\theta}$ is achieved once the SIFIAE model learning process is completed. As explained by [35], the second prerequisite has also been satisfied. Considering that there are five frequency bands and 62 EEG channels in the SEED-V data set, we define the quantitative importance of each frequency band and channel by the following rules

$$\omega(i) = \theta_{(i-1)*62+1} + \theta_{(i-1)*62+2} + \cdots + \theta_{(i-1)*62+62},\ i = 1, 2, 3, 4, 5 \tag{13}$$

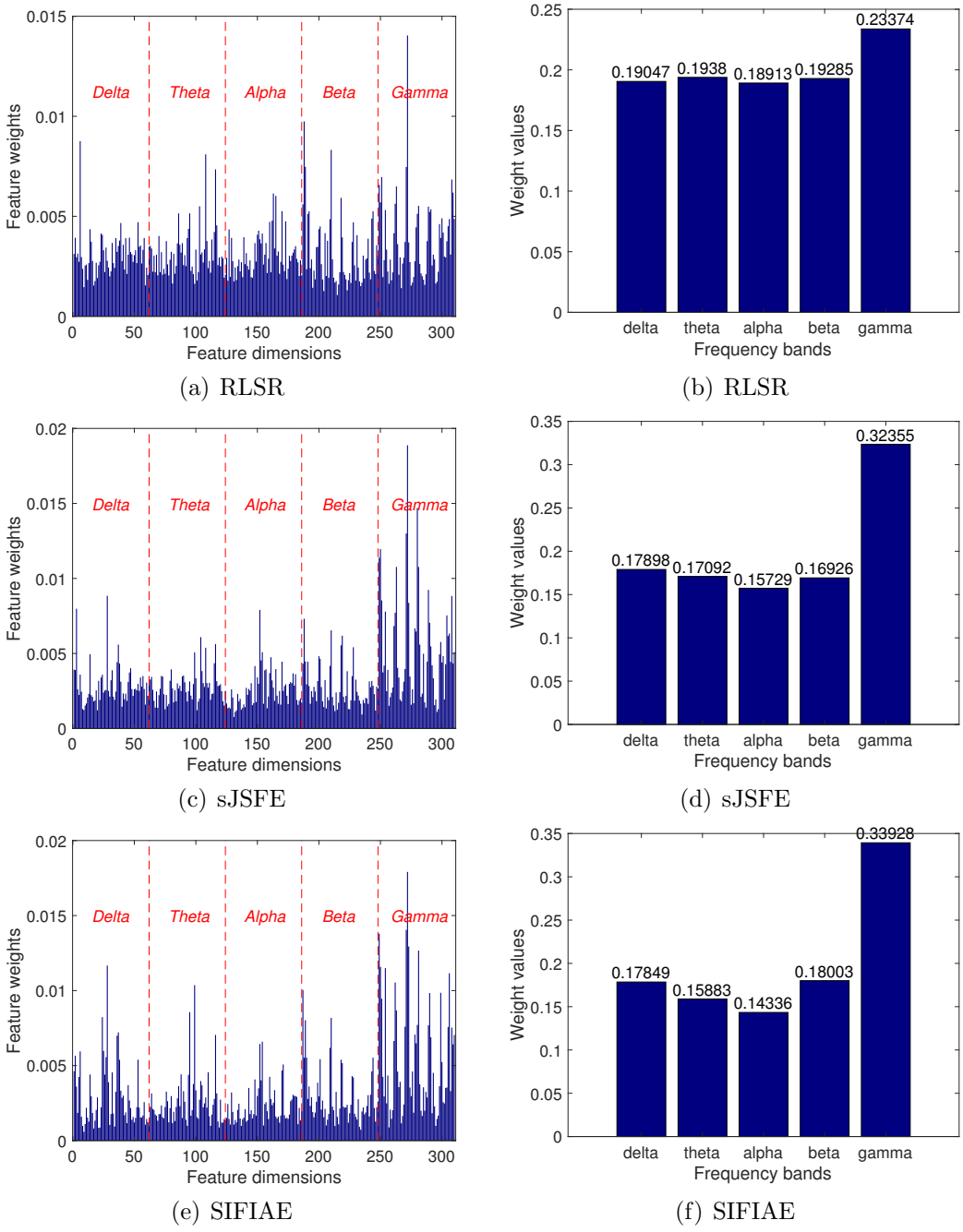
and

$$\psi(j) = \theta_j + \theta_{j+62} + \theta_{j+124} + \theta_{j+186} + \theta_{j+248},\ j = 1, 2, \cdots, 62 \tag{14}$$

where $\omega(i)$ and $\psi(j)$ represent the importance of the i-th frequency band and the j-th channel, respectively. As shown in Figs. 5 and 6, we visualized the average importance across all these 96 cross-session emotion recognition cases (*i.e.*, there are 16 subjects and each subject has six cross-session emotion recognition tasks) of feature dimensions, frequency bands and channels. As provided by the results in both figures, we have the following two findings.

- *The three models identified similar results on critical frequency bands and channels in cross-session emotion recognition.* Specifically, the *Gamma* band is considered as the most important one among the five frequency bands, as shown in Fig. 5. The channels in prefrontal, parietal and lateral temporal lobes are more activated than those in other brain areas, as demonstrated in Fig. 6. These identified critical frequency bands and channels in cross-session emotion recognition are also consistent with some existing researches [19] and illustrate that there might be specific frequency-spatial patterns of neural processes in response to visual emotional stimulus.

- *In SIFIAE, the critical frequency bands and channels are more obvious than other two models.* Specifically, the weight value of *Gamma* band achieves 0.33928 in SIFIAE while in RLSR and sJSFE, the weights of *Gamma* band only achieve 0.23374 and 0.32355 respectively. Similarly, this phenomenon also happens in the critical channels we mentioned above. It indicates that these critical frequency bands and channels can be used more effectively on a higher quality EEG data obtained by SIFIAE to improve the performance of emotion recognition.

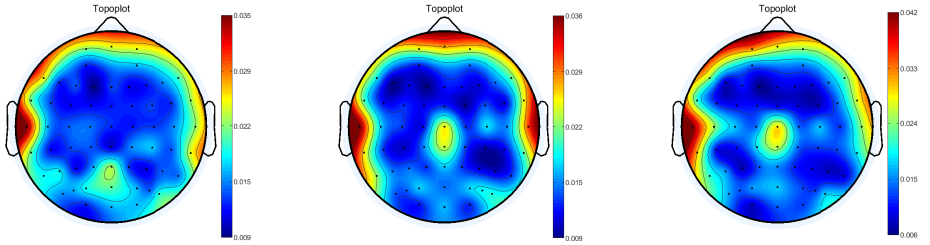


(a) RLSR (b) RLSR

(c) sJSFE (d) sJSFE

(e) SIFIAE (f) SIFIAE

Figure 5: The average importance of feature dimensions (left column) and frequency bands (right column) learned by three models.



(a) RLSR (b) sJSFE (c) SIFIAE

Figure 6: The average importance of channels learned by three models.

4. Discussion

In this section, we discuss the similarities and differences between SIFIAE and two related models, *i.e.*, sJSFE [34] and RSSLSR [33], from the perspectives of motivation, model formulation, optimization and experimental results. Especially, we analyze the underlying connection between the sample weighting technique and the $\ell_{2,p}$-norm based loss function in achieving the model robustness.

- *Motivation.* Obviously, both SIFIAE and sJSFE are task-driven methods, which aim at improving the performance of EEG-based emotion recognition by referring appropriate machine learning models to solve certain limitations within this field. In more detail, SIFIAE is more specific to the affecitve overlap problem usually happened in adjacent trials of stimulus-evoked EEG data acquisition experiment, while sJSFE is more general for purifying EEG data from both horizontal and vertical axes when vectorized EEG samples are organized as columns of the data matrix. RSSLSR is a model-driven method, which was motivated by enhancing the robustness of LSR in processing noisy data. Therefore, such motivation is more straightforward and the model itself might be more general for multiple applications.

- *Model formulation and optimization.* From the model formulation, SIFIAE and sJSFE share similar objective functions. However, different ways are used in them to achieve the quantitative sample inconsistency (SIFIAE) and importance (sJSFE) measurements. To be specific, SIFIAE adaptively learns the factor $\mathbf{s}$ while in sJSFE, it is

obtained by looking up the self-paced function with the sample approximation error as index. In other words, the former is a completely automatic process while the latter is a semi-automatic one. Therefore, the optimization procedures to variable $\mathbf{s}$ in both models are totally different, which is expected to have continuous values. In RSSLSR, a sample is either identified as a normal data or an outlier and accordingly a binary vector is introduced to characterize the possible types of each sample, which plays similar roles with the factor $\mathbf{s}$ in SIFIAE and sJSFE. Informally, we can understand that both SIFIAE and sJSFE used a soft weighting technique while RSSLSR used a hard selection strategy in determining the impact of samples in model learning. There is also some minor differences between SIFIAE and sJSFE such as the experimental paradigms (*i.e.*, pure semi-supervised learning in SIFIAE while transductive mode in sJSFE) and the constraints defined on factor $\mathbf{s}$ (*i.e.*, non-negative and normaization constraints in SIFIAE and only the non-negative constraint in sJSFE).

- *Experimental Results.* Below we explain the underlying reasons accounting for performance differences between RSSLSR, sJSFE and SIFIAE, followed by a one-way ANOVA (analysis of variance) to check the statistical difference among their experimental results. As analyzed above, in RSSLSR, a sample is identified as either a normal point or an outlier, which is more suitable for those data sets with obvious outliers. However, our EEG data has been preprocessed by filtering and removing artifacts, which may deviate a lot from this type of data set. Therefore, it is challenging to achieve optimal results using the RSSLSR approach on our EEG data. This deduction is supported by the results presented in Tables 1 to 3, which indicate that the mean accuracy of RSSLSR is significantly lower than that of sJSFE and SIFIAE. In addition, we found SIFIAE outperforms sJSFE in most cases. We believe that the primary reason for the performance difference between sJSFE and SIFIAE lies in the self-paced learning function used in sJSFE to determine the importance of samples. This function includes an aging parameter that needs to be initialized manually, which reduces the flexibility of the model compared to SIFIAE. To verify the significant difference between our SIFIAE model and other comparison models, we conducted a one-way ANOVA on the experimental results in Table 1 to Table 3. Each model has a total of 96 recognition accuracies, 16

for each recognition situation. The null hypothesis of ANOVA assumes that the means of the recognition accuracies for different models are equal. Table 4 shows the p-values obtained by ANOVA for each model. A p-value less than 0.05 indicates a substantial difference between the two models, while a p-value less than 0.01 indicates an extremely significant difference. Our results show that the SIFIAE model has significantly better performance compared with sLSR, RLSR, and RSSLSR models. In addition, our SIFIAE model also shows some improvement over sJSFE.

Table 4: The one-way ANOVA results between SIFIAE and other four comparison models (**p-value<0.01, *p-value<0.05).

ANOVA	sLSR	RLSR	RSSLSR	sJSFE
SIFIAE	1.5926e-16**	6.0871e-13**	2.0714e-28**	0.0165*

- *The connection between sample weighting and $\ell_{2,p}$-norm based model robustness.* In RSSLSR, the $\ell_{2,p}$-norm ($0 < p \leq 2$) is used to measure the model approximation loss and accordingly improve its robustness. Below we show how the updating rule of the sample inconsistency factor in SIFIAE connects with the $\ell_{2,p}$-norm. Generally, the $\ell_{2,p}$-norm based least square regression can be written as

$$\min_{\mathbf{W},\mathbf{b}} \|\mathbf{X}^T\mathbf{W} + \mathbf{1}\mathbf{b}^T - \mathbf{Y}\|_2^p + \|\mathbf{W}\|_F^2, \tag{15}$$

where $\|\mathbf{X}\|_2^p = (\sum_{i=1}^m (\sum_{j=1}^n |x_{ij}|^2)^{\frac{p}{2}})^{\frac{1}{p}},\ 0 < p \leq 2,\ \mathbf{X} \in \mathbb{R}^{m\times n}$. Based on the derivation in [33], equation (15) can be rewritten as

$$\min_{\mathbf{W},\mathbf{b}} \sum_{i=1}^n f_i \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2 + \|\mathbf{W}\|_F^2, \tag{16}$$

where $f_i = \frac{p}{2}(\|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2)^{\frac{p}{2}-1}$. Similarly, we denote the approximation error in equation (16) as $e_i = \|\mathbf{x}_i^T\mathbf{W} + \mathbf{b}^T - \mathbf{y}^i\|_2^2$ and the weighting factor $r = \frac{2}{q} \geq 1$. By ignoring the coefficient r in f_i and denoting $g_i = f_i^{\frac{r}{(1-r)^2}}$, we have the normalized g_i as

$$g_i^{'} = \frac{f_i^{\frac{r}{(1-r)^2}}}{\sum_{j=1}^n f_i^{\frac{r}{(1-r)^2}}} = \frac{e_i^{\frac{1}{1-r}}}{\sum_{j=1}^n e_j^{\frac{1}{1-r}}} = \frac{1}{e_i^{\frac{1}{r-1}} \sum_{j=1}^n (\frac{1}{e_j})^{\frac{1}{r-1}}}. \tag{17}$$

It is worthy noting that when $r = 1$, this equivalent transformation is invalid since $(r - 1)$ cannot serve as a denominator in this situation. Obviously, equation (17) shares similar expressions with the updating rule of $\mathbf{s}$ in equation (5), except for the regularization parameter λ we introduced in $\mathbf{s}$ to avoid trivial solution.

5. Conclusion

In this paper, we proposed a new model termed SIFIAE to deal with the affective overlap problem in stimulus-evoked EEG emotion recognition. In SIFIAE, the feature-label inconsistency in EEG data caused by affective overlap problem was quantitatively measured by an introduced weighting factor which was adaptively learned in model training. In addition, the feature importance was jointly optimized with sample inconsistency to improve the effectiveness. Overall, the SIFIAE model was learned within the pure semi-supervised learning framework, which has the out-of-sample extension ability and is more practical for EEG-based emotion recognition in real scenarios. Experimental results demonstrated SIFIAE significantly improved the emotion recognition performance. The visualization of sample inconsistency factor suggested that affective overlap samples were reasonably suppressed. The average feature importance provides us with insights that *Gamma* frequency band and the prefrontal, parietal and lateral temporal lobes are more activated in emotion recognition.

CRediT authorship contribution statement

Yong Peng, Junhua Li and Wanzeng Kong designed this study. Yikai Zhang did the experiments and wrote the first draft of the manuscript. All authors helped with the modification of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was supported by the National Natural Science Foundation of China under Grant 61971173 and the Key Research and Development Project of Zhejiang Province under Grant 2023C03026, Grant 2021C03001, and Grant 2021C03003, and the Planted Talent Plan of Zhejiang Province under Grant 2022R407C066.

References

[1] B. C. Ko, A brief review of facial emotion recognition based on visual information, Sensors 18 (2018) 401. Https://doi.org/10.3390/s18020401.

[2] N. Alswaidan, M. E. B. Menai, A survey of state-of-the-art approaches for emotion recognition in text, Knowl. Inf. Syst. 62 (2020) 2937–2987. Https://doi.org/10.1007/s10115-020-01449-0.

[3] M. B. Akçay, K. Oğuz, Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers, Speech Commun. 116 (2020) 56–76. Https://doi.org/10.1016/j.specom.2019.12.001.

[4] A. Khosla, P. Khandnor, T. Chand, A comparative analysis of signal processing and classification methods for different applications based on EEG signals, Biocybern. Biomed. Eng. 40 (2020) 649–690. Https://doi.org/10.1016/j.bbe.2020.02.002.

[5] H. Wang, H. Yu, H. Wang, EEG_GENet: A feature-level graph embedding method for motor imagery classification based on EEG signals, Biocybern. Biomed. Eng. 42 (2022) 1023–1040. Https://doi.org/10.1016/j.bbe.2022.08.003.

[6] W. Tao, C. Li, R. Song, J. Cheng, Y. Liu, F. Wan, X. Chen, EEG-based emotion recognition via channel-wise attention and self attention, IEEE Trans. Affect. Comput. 14 (2023) 382–393. Https://doi.org/10.1109/TAFFC.2020.3025777.

[7] J. Shen, X. Zhang, G. Wang, Z. Ding, B. Hu, An improved empirical mode decomposition of electroencephalogram signals for depression detection, IEEE Trans. Affect. Comput. 13 (2022) 262–271. Https://doi.org/10.1109/TAFFC.2019.2934412.

[8] R. Nawaz, K. H. Cheah, H. Nisar, V. V. Yap, Comparison of different feature extraction methods for EEG-based emotion recognition, Biocybern. Biomed. Eng. 40 (2020) 910–926. Https://doi.org/10.1016/j.bbe.2020.04.005.

[9] M. Wu, S. Hu, B. Wei, Z. Lv, A novel deep learning model based on the ICA and riemannian manifold for EEG-based emotion recognition, Neurosci. Methods 378 (2022) 109642. Https://doi.org/10.1016/j.jneumeth.2022.109642.

[10] J. Li, Y. Hao, W. Zhang, X. Li, B. Hu, Effective connectivity based EEG revealing the inhibitory deficits for distracting stimuli in major depression disorders, IEEE Trans. on Affect. Comput. 14 (2021) 694–705. Https://doi.org/10.1109/TAFFC.2021.3054953.

[11] H. Huang, Q. Xie, J. Pan, Y. He, Z. Wen, R. Yu, Y. Li, An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness, IEEE Trans. on Affect. Comput. 12 (2021) 832–842. Https://doi.org/10.1109/TAFFC.2019.2901456.

[12] D. Wu, J. Huang, Affect estimation in 3D space using multi-task active learning for regression, IEEE Trans. on Affect. Comput. 13 (2022) 16–27. Https://doi.org/10.1109/TAFFC.2019.2916040.

[13] S. Gong, K. Xing, A. Cichocki, J. Li, Deep learning in EEG: Advance of the last ten-year critical period, IEEE Trans Cogn. Develop. Syst. 14 (2022) 348–365. Https://doi.org/10.1109/TCDS.2021.3079712.

[14] 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.

[15] A. Olamat, P. Ozel, S. Atasever, Deep learning methods for multi-channel EEG-based emotion recognition, Int. J. Neural Syst. 32 (2022) 2250021. Https://doi.org/10.1142/S0129065722500216.

[16] M. Xing, S. Hu, B. Wei, Z. Lv, Spatial-frequency-temporal convolutional recurrent network for olfactory-enhanced EEG emotion recognition, Neurosci. Methods 376 (2022) 109624. Https://doi.org/10.1016/j.jneumeth.2022.109624.

[17] S. Liu, Z. Wang, Y. An, J. Zhao, Y. Zhao, Y.-D. Zhang, EEG emotion recognition based on the attention mechanism and pre-trained convolution capsule network, Knowl.-Based Syst. 265 (2023) 110372. Https://doi.org/10.1016/j.knosys.2023.110372.

[18] X. Wang, Y. Ma, J. Cammon, F. Fang, Y. Gao, Y. Zhang, Self-supervised EEG emotion recognition models based on CNN, IEEE Trans. Neural Syst. Rehabil. Eng. (2023). Https://doi.org/10.1109/TNSRE.2023.3263570.

[19] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, IEEE Trans. Auton. Ment. Dev. 7 (2015) 162–175. Https://doi.org/10.1109/TAMD.2015.2431497.

[20] T. Song, W. Zheng, P. Song, Z. Cui, EEG emotion recognition using dynamical graph convolutional neural networks, IEEE Trans Affect. Comput. 11 (2020) 532–541. Https://doi.org/10.1109/TAFFC.2018.2817622.

[21] Y. Li, W. Zheng, L. Wang, Y. Zong, Z. Cui, From regional to global brain: A novel hierarchical spatial-temporal neural network model for EEG emotion recognition, IEEE Trans. Affect. Comput. 13 (2022) 568–578. Https://doi.org/10.1109/TAFFC.2019.2922912.

[22] C. Zhang, Z. Yang, X. He, L. Deng, Multimodal intelligence: Representation learning, information fusion, and applications, IEEE J. Sel. Topics Signal Process. 14 (2020) 478–493. Https://doi.org/10.1109/JSTSP.2020.2987728.

[23] W.-L. Zheng, W. Liu, Y. Lu, B.-L. 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.

[24] X. Zhang, J. Liu, J. Shen, S. Li, K. Hou, B. Hu, J. Gao, T. Zhang, Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine, IEEE Trans. Cybern. 51 (2020) 4386–4399. Https://doi.org/TCYB.2020.2987575.

[25] M. Wu, W. Teng, C. Fan, S. Pei, P. Li, Z. Lv, An investigation of olfactory-enhanced video on EEG-based emotion recognition, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1602–1613. Https://doi.org/10.1109/TNSRE.2023.3253866.

[26] X. Li, Y. Zhang, P. Tiwari, D. Song, B. Hu, M. Yang, Z. Zhao, N. Kumar, P. Marttinen, EEG based emotion recognition: A tutorial and review, ACM Comput. Surv. 55 (2022) 57. Https://doi.org/10.1145/3524499.

[27] Y. Xu, Y. Cui, X. Jiang, Y. Yin, J. Ding, L. Li, D. Wu, Inconsistency-based multi-task cooperative learning for emotion recognition, IEEE Trans. on Affect. Comput. 13 (2022) 2017–2027. Https://doi.org/10.1109/TAFFC.2022.3197414.

[28] Y. Yang, Y. Yang, H. T. Shen, Y. Zhang, X. Du, X. Zhou, Discriminative nonnegative spectral clustering with out-of-sample extension, IEEE Trans. Knowl. Data Eng. 25 (2013) 1760–1771. Https://doi.org/10.1109/TKDE.2012.118.

[29] X. Chen, G. Yuan, F. Nie, Z. Ming, Semi-supervised feature selection via sparse rescaled linear square regression, IEEE Trans. Knowl. Data Eng. 32 (2020) 165–176. Https://doi.org/10.1109/TKDE.2018.2879797.

[30] 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.

[31] W. Liu, J.-L. Qiu, W.-L. Zheng, B.-L. Lu, Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition, IEEE Trans. Cognit. Devel. Syst. 14 (2022) 715–729. Https://doi.org/10.1109/TCDS.2021.3071170.

[32] Y. Peng, Y. Zhang, W. Kong, F. Nie, B.-L. Lu, A. Cichocki, S3LRR: A unified model for joint discriminative subspace identification and semisupervised EEG emotion recognition, IEEE Trans. Instrum. Meas. 71 (2022) 1–13. Https://doi.org/10.1109/TIM.2022.3165741.

[33] J. Wang, F. Xie, F. Nie, X. Li, Robust supervised and semisupervised least squares regression using $\ell_{2,p}$ norm minimization, IEEE Trans. Neural Netw. Learn. Syst. (2022). Https://doi.org/10.1109/TNNLS.2022.3150102.

[34] X. Li, F. Shen, Y. Peng, W. Kong, B.-L. Lu, Efficient sample and feature importance mining in semi-supervised EEG emotion recognition, IEEE Trans. Circuits Syst. II 69 (2022) 3349–3353. Https://doi.org/10.1109/TCSII.2022.3163141.

[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. Cognit. Devel. Syst. 14 (2022) 935–947. Https://doi.org/10.1109/TCDS.2021.3082803.