DAformer: Transformer with Domain Adversarial Adaptation for EEG-Based Emotion Recognition with Live-Oil Paintings

Zhong-Wei Jin[1], Jia-Wen Liu[1], Wei-Long Zheng[1,2], and Bao-Liang Lu[1,2,3,4(✉)]

[1] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
{izzie_kim,ljw_venn,weilong,bllu}@sjtu.edu.cn
[2] Key Laboratory of Shanghai Commission for Intelligent Interaction and Cognitive Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[3] Center for Brain-Machine Interface and Neuromodulation, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200020, China
[4] RuiJin-Mihoyo Laboratory, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200020, China

Abstract. The emergence of domain adaptation has brought remarkable advancement to EEG-based emotion recognition by reducing subject variability thus increasing the accuracy of cross-subject tasks. A wide variety of materials have been employed to elicit emotions in experiments, however, artistic works that aim to evoke emotional resonance of observers are relatively less frequently utilized. Previous research has shown promising results in electroencephalogram(EEG)-based emotion recognition on static oil paintings. As video clips are widely recognized as the most commonly used and effective stimuli, we adopted animated live oil paintings, a novel set of emotional stimuli in the live form which are essentially a type of video clip while possessing fewer potential influencing factors for EEG signals compared to traditional video clips, such as abrupt switches on background sound, contrast, and color tones. Moreover, previous studies on static oil paintings focused primarily on the subject-dependent task, and further research involving cross-subject analysis remains to be investigated. In this paper, we proposed a novel DAformer model which combines the advantages of Transformer and adversarial learning. In order to enhance the evocative performance of oil paintings, we introduced a type of innovative emotional stimuli by transforming static oil paintings into animated live forms. We developed a new emotion dataset SEED-LOP (SJTU EEG Emotion Dataset-Live Oil Painting) and constructed DAformer to verify the effectiveness of SEED-LOP. The results demonstrated higher accuracies in three-class emotion recognition tasks when watching live oil paintings, with a subject-dependent accuracy achieving 61.73% and a cross-subject accuracy reaching 54.12%.

B. Luo et al. (Eds.): ICONIP 2023, CCIS 1963, pp. 402–414, 2024.
https://doi.org/10.1007/978-981-99-8138-0_32

Keywords: Emotion recognition · Transfer learning · Electroencephalogram · Oil paintings

1 Introduction

Affective computing aims at identifying, analyzing, and interpreting human emotional states by employing various techniques such as natural language processing, data mining, and machine learning to recognize and understand emotional information effectively, among which electroencephalogram(EEG)-based emotion recognition draws the most attention for its neural pattern stability in emotion [1,2]. In recent years, a diverse range of emotion elicitation materials has been employed in studies focusing on emotion recognition based on EEG signals: images [3], video clips [4], musical segments [5], etc. However, artistic works which aim to evoke the emotional resonance of observers are relatively less frequently utilized. Luo *et al.* innovatively utilized oil paintings as emotional stimuli and demonstrated their effectiveness in a subject-dependent EEG-based emotion recognition task [6]. Subsequently, Lan *et al.* employed Transformer to conduct subject-dependent emotion recognition on the oil painting dataset and achieved promising performance [7].

In recent years, cross-subject emotion recognition has emerged as a research focus due to its compatibility with the practical requirements of real-world applications. With the development of transfer learning, remarkable advancement has been brought to cross-subject emotion recognition tasks by reducing subject variability. Domain adaptation (DA) is a highly important branch of transfer learning [8] whose primary objective is to effectively map features from various source domains into a unified feature space, with the goal of eliminating the domain discrepancy between the source and target domains, thus increasing the accuracy of the target domain. Methods based on domain adversarial neural networks (DANNs) were proposed to identify common representations of EEG signals among all subjects, thereby improving the cross-subject performance [9]. Apart from adversarial methods, subdomain adaption (DSAN) has also yielded promising results. DSAN partitions similar samples within a domain into subdomains by certain criteria and aligns these subdomains rather than the global domain alignment [10].

In addition to advanced classification methods, more neural network architectures for EEG signals to extract efficient representations in various domains have also garnered significant attention in the field of affective computing. Wang *et al.* employed an attention mechanism to fuse EEG and eye signals [11]. Spectral-spatial-temporal adaptive graph convolutional neural network (SST-AGCN) was designed to extract EEG features from spectral, temporal, and spatial domains based on a graph convolutional neural network and achieved outstanding performance in a subject-dependent confidence estimation task [12].

As mentioned before, previous studies of oil painting datasets focused mainly on the subject-dependent tasks [6,7]. However, the effectiveness of this paradigm has not yet been validated across subjects which lays the foundation for practical cross-subject applications. This paper aims to address this issue by introducing a novel DAformer model. By combining the advantage of the attention mechanism in feature extraction of EEG signals with the effectiveness of domain adversarial methods for the cross-subject task, we employed DAformer to recognize three different classes of emotions (negative, neutral, positive) from EEG signals. Since the emotional film clip is one of the most popular stimuli which has been proven to be effective [13], we fabricated a novel set of emotional stimuli of animated live oil paintings in order to enhance the emotion-inducing effects. We conducted experiments to collect EEG signals and eye movements of participants under the new stimuli and developed a new dataset: SEED-LOP (SJTU Emotion EEG Dataset-Live Oil Painting)[1]. Finally, we demonstrated the superior performance of the DAformer model as well as the feasibility of the SEED-LOP dataset. The main contribution of this new dataset lies in its revolutionary implementation of 2D live art pieces as stimuli, which has established a more streamlined and lightweight experimental paradigm compared to traditional emotional video clips and provided innovative methods in the field of affective computing.

2 Methods

We constructed the DAformer model to verify the effectiveness of SEED-LOP (Fig. 1). DAformer is composed of a feature extractor G_f, an emotion predictor G_y, and a domain discriminator G_d. We employed a multi-layer encoder based on the multi-head self-attention mechanism as the feature extractor, a fully connected feed-forward network for emotion prediction, together with a domain adversarial module to eliminate the domain distribution discrepancy across different subjects. The gradient reversal layer propagates the gradients from the domain discriminator reversely to the feature extractor during backward propagation, thereby enabling the feature extractor to extract invariant EEG features in different domains. The training dataset included EEG signals from both labeled data in the source domain and unlabeled data in the target domain. All data are labeled by their corresponding domains to train the domain discriminator, while only source data and its emotion labels are utilized to train the emotion predictor to predict three types of emotions (0: negative, 1: neutral, 2: positive). The test dataset consists of target data and their emotion labels to evaluate the performance of the feature extractor and emotion predictor.

2.1 Data Preprocessing

The raw EEG signals were first downsampled to 200 Hz and processed with a 1–70 Hz bandpass filter and a 50 Hz notch filter. We extracted the differential entropy (DE) features of EEG signals as it is proven to have excellent performance in previous studies [11,12,14]. The preprocessed EEG data underwent the short-time Fourier transform (STFT) using a one-second Hanning window. This

[1] https://bcmi.sjtu.edu.cn/home/seed/.

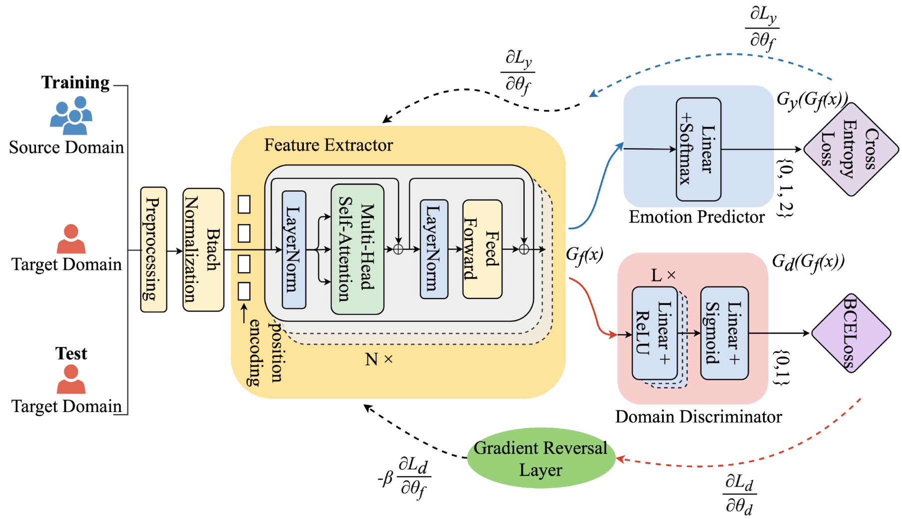

Fig. 1. The overall architecture of the DAformer. Solid arrows represent the forward propagation and dashed arrows represent the backward propagation. The gradient reversal layer propagates the gradient reversely to the feature extractor during backward propagation, thereby enabling the feature extractor to learn invariant representations in the source and target domains.

transformation was performed on each epoch to extract the differential entropy (DE) features from five distinct frequency bands: delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–50 Hz). The extracted DE features of EEG signals are defined as $X \in \mathbb{R}^{N \times F \times C}$ where N stands for the total number of samples, F represents five frequency bands and C denotes the EEG channels.

2.2 DAformer

Feature Extractor. As EEG signals are temporal signals, we employed an encoder based on the multi-head self-attention mechanism as the feature extractor. A batch normalization layer is applied at the beginning to avoid over-fitting and accelerate the training process. After batch normalization and position encoding of a sequence length L_{seq}, the representations of EEG signals are fed into the encoder which consists of N identical layers, each of the layers contains two sub-layers: a multi-head attention layer and a fully connected feed-forward network. Both sub-layer is preceded by a layer normalization and applied with a residual connection.

A multi-head attention layer consists of a parallel concatenation of N scaled dot product attention layers. First of all, the original input $X \in \mathbb{R}^{B \times T \times D}$ where B denotes the batch size, T the overlapping window, and D the feature dimension of EEG signals after preprocessing. After a linear transformation, a query of dimension d_k, a key of dimension d_k, and a value of dimension d_v are generated to compose the input of the scaled dot-product attention. The output of a single scaled dot-product attention is calculated as below, where Q, K, and V stand for the packed matrices of queries, keys, values, and d_k denotes the scaling factor in order to avoid extremely small gradients:

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V. \tag{1}$$

The output of all multi-head attention layers is shown as below:

$$\begin{aligned} \text{MultiHead}(Q, K, V) &= \text{Concat}(\text{head}_1, ..., \text{head}_\text{h})W^O, \\ \text{head}_i &= \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), \end{aligned} \tag{2}$$

where the weight matrices $W_i^Q \in \mathbb{R}^{d_{en} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{en} \times d_k}$, $W_i^V \in \mathbb{R}^{d_{en} \times d_v}$, $W_i^O \in \mathbb{R}^{hd_v \times d_{en}}$ and d_{en} denotes the output dimension of the encoder.

The output is then fed into a residual connection before the fully connected feed-forward network which comprises two linear transformations connected by a ReLU activation function:

$$\begin{aligned} x &= \text{LayerNorm}(x + \text{Sublayer}(x)), \\ \text{FFN}(x) &= \max(0, xW_1 + b_1)W_2 + b_2. \end{aligned} \tag{3}$$

Emotion Predictor. We utilized a single layer fully connected feed-forward network as the emotion predictor.

Domain Discriminator. The domain discriminator is composed of L layers where each is a single linear transformation with a ReLU activation function. The output of the last linear transformation is activated by a Sigomoid activation function. Between feature extractor G_f and domain discriminator G_d, a gradient reversal layer is employed to propagate the gradient reversely, thereby fostering an adversarial training process between G_f and G_d. As G_y and G_d gain increasing precision in classifying emotion types and domain types, the reversed gradient is intended to guide G_f in extracting features that are indistinguishable for G_d.

We denoted $x_{d,i}$ the ith input of the model, $d \in \{s, t\}$ which represents the domain of $x_{d,i}$ (s for source domain and t for target domain), the feature extractor produces its output $G_f(x_{d,i})$. The extracted features of source domain data are then fed into the emotion predictor and domain discriminator while those of the target domain are only applied to the domain discriminator. The emotion predictor predicts the emotion label $G_y(G_f(x_{s,i}))$ only for source data, while the domain discriminator makes domain prediction $G_d(G_f(x_{d,i}))$ for both source data and target data.

We utilized the cross entropy loss for the emotion predictor. The loss function can be presented as:

$$L_y = -\frac{1}{|s|} \sum_{i=1}^{|s|} y_{s,i} \log \hat{y}_{s,i}, \tag{4}$$

where $\hat{y}_{s,i}$ denotes the output of the emotion predictor: $\hat{y}_{s,i} = G_y(G_f(x_{s,i}))$ and $y_{s,i}$ denotes the true emotion label of $x_{s,i}$.

We applied the binary cross entropy loss for the domain discriminator. The loss functions for the source domain and target domain are expressed as:

$$L_s = -\frac{1}{|s|}\sum_{i=1}^{|s|}(d_{s,i}\log\hat{d}_{s,i} + (1-d_{s,i}) * \log\hat{d}_{s,i}), \tag{5}$$

$$L_t = -\frac{1}{|t|}\sum_{i=1}^{|t|}(d_{t,i}\log\hat{d}_{t,i} + (1-d_{t,i}) * \log\hat{d}_{t,i}), \tag{6}$$

where $\hat{d}_{s,i}$ and $\hat{d}_{t,i}$ denote the predictions of the domain discriminator: $\hat{d}_{s,i} = G_d(G_f(x_{s,i}))$, $\hat{d}_{t,i} = G_d(G_f(x_{t,i}))$, respectively, and $d_{s,i}$, $d_{t,i}$ denote the true domain label of $x_{s,i}$, $x_{t,i}$, respectively.

The total loss of DAformer is composed as below:

$$L = L_y + \beta(L_s + L_t), \tag{7}$$

where β is a hyperparameter that balances the weight between the emotion class loss and the domain loss.

We optimized the parameters of the feature extractor and the emotion predictor by minimizing the total loss function, and we updated the parameters of the domain discriminator by maximizing the total loss function:

$$\begin{aligned} \hat{\theta}_f, \hat{\theta}_y &= argmin_{\theta_f,\theta_y}\mathcal{L}(\theta_f, \theta_y, \hat{\theta}_d), \\ \hat{\theta}_d &= argmax_{\theta_d}\mathcal{L}(\hat{\theta}_f, \hat{\theta}_y, \theta_d). \end{aligned} \tag{8}$$

3 Experiments

We designed a novel emotional experiment paradigm using animated live oil paintings as stimuli. The value of this innovative dataset lies in its groundbreaking use of 2D live art pieces as stimuli, which has achieved a more lightweight experimental paradigm compared to video clips and yielded effective classification results. A total of 40 healthy, right-handed participants (20 males, 20 females) aged between 17 and 29 years (mean age: 21.6 years, standard deviation: 2.95) were recruited from Shanghai Jiao Tong University for the experiment. Each participant viewed 60 oil paintings (20 positive, 20 neutral, 20 negative) during the experiments, and both their EEG signals and eye movements were recorded during the observation of paintings. As previous studies proved that the power of the theta band and the alpha band power of EEG signals differ between artists and non-artists in response to abstract and representational paintings [15], our participants were selected intentionally as half artistically experienced and half artistically naive under the results of the art experience questionnaire [16]. We successfully collected EEG signals and eye movements of 40 participants with a statistic distribution of 28% negative emotion samples, 34% neutral emotion samples, and 38% positive emotion samples.

3.1 Dataset

Stimuli. Luo *et al.* employed a total of 114 world-renowned paintings spanning from the mid-16th century to the 19th century [6]. The paintings encompassed a diverse range of genres, including portraiture, animal depiction, still life, landscape, and cityscape, which were representative of most major art styles. Questionnaires were distributed among students of Shanghai Jiao Tong University and China Academy of Art on the perception of emotion types (negative, neutral, positive) of the 60 paintings as well as their level of intensity. We selected 60 paintings (20 of each type of emotion) that have the highest intensity ratings and put them into our stimuli set. Before the animation of the oil paintings, we conducted an experiment to collect the eye movement heatmaps during the observations of these paintings in order to locate the parts that draw the most attention. 40 healthy participants were recruited (9 males, 31 females, age: 25.35 ± 3.71) to observe the 60 oil paintings selected. Each painting was presented for 20 s with a one-second interval of focus time between each display. The eye movements of the participants during the observation were recorded by a Tobii Pro X3-120 screen-based eye tracker at a sampling rate of 120 Hz. We produced an attention heatmap for each painting by calculating the average gaze point of all participants (Fig. 2). Based on the attention heatmaps, we successfully transformed 60 static oil paintings into animated live GIFs using Photoshop software and Cartoon Animator 4 software.

Fig. 2. Examples of attention heatmaps of the average eye movements of 40 participants during observations of the oil paintings from our stimuli set. The transition from static oil paintings to animation form was based on the attention heatmaps.

Procedure. The experimental protocol is presented in Fig. 3 to illustrate the details of our main experiment. In the experiment for each participant, 60 animated oil paintings were organized into 5 groups. Each group comprised 3 randomly decided different emotion type batches, with each batch consisting of 4 paintings randomly selected with the same emotion class in order to minimize the occurrence of frequent emotional transformation within a short timeframe. Before the experiment, the participants were briefed on the experimental protocol and shown a tutorial of three example paintings representing negative, neutral, and positive emotions as a reference level for rating. Each painting was preceded by a one-second focus time of a black screen with a white + symbol in the middle. The duration of display of each painting was fixed at 20 s, during which the participants were asked to intently observe and perceive the painting. After a one-second focus session that followed, participants were instructed to report both their emotional perception of the emotion type (negative, neutral, and positive) and the intensity (1–9) at the rating session with no time limit.

During the experiments, the EEG signals of subjects were collected by an ESI NeuroScan System with a 62-channel module arranged according to the international 10–20 system at a sampling rate of 1000 Hz. Eye movements were recorded by a Tobii Pro X3-120 screen-based eye tracker at a sampling rate of 120 Hz.

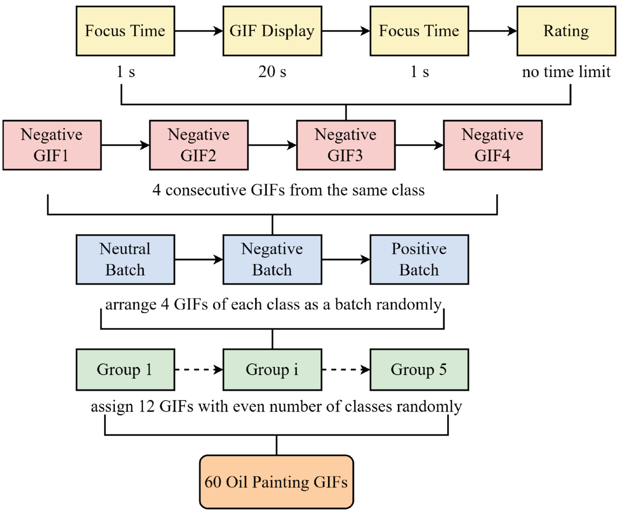

Fig. 3. Illustration of the experimental protocol.

3.2 Implementation Details

During the experiment, the reported emotion categories of the participants (0: negative emotion, 1: negative emotion, 2: positive emotion) were used as classification labels to investigate the effectiveness of this paradigm in emotion classification tasks based on EEG signals. We utilized EEG features in a total of five frequency bands. In the subject-dependent experiment, all models were trained on each participant using three-fold cross-validation, while the data under the observation of the same oil painting did not appear in both the training sets and the testing sets simultaneously. For the cross-subject experiment, we employed Leave-One-Out-Cross-Validation. To evaluate the performance of DAformer, we compared our model with four other classifiers, including support vector machine (SVM), multilayer perceptron (MLP), Transformer [17], and spectral-spatial-temporal adaptive graph convolutional neural network (SST-AGCN) [12], both in the subject-dependent task and the cross-subject task, as well as the comparison of four different domain adaptation methods: correlation alignment (CORAL) [18], domain-adversarial neural networks (DANN) [19], dynamic adversarial adaptation network (DAAN) [20] and deep subdomain adaptation network (DSAN) [10] additionally for the cross-subject task.

In our experiments, we set batch size $B = 128$, window size $= 4$ s thus window number $T = 5$. The number of channels of EEG signals was 62 for our experiment, and together with the 5 frequency bands, the feature dimension D of EEG signals added to 310. For both the subject-dependent task and the cross-subject task, the SVM classifiers were applied with the RBF kernel with the range of parameter C is $2^{[-8:8]}$.

For DAformer, we employed the encoder dimension $d_{en} = 512$, the dimension of feed-forward network $d_{ff} = 2048$, the learning rate in a range of $[1e-3, 5e-4, 1e-4]$, the domain loss weight $\beta = 10$, the layer number of domain discriminator $L = 1$ with 256 dimensions and the sequence length L_{seq} within a range of $[1, 2, 4]$.

3.3 Results Analysis

In this section, we compared the performance of DAformer with the other four classifiers, including SVM, MLP, Transformer [17], and SST-AGCN [12], both in the subject-dependent task and the cross-subject task, with the comparison of 4 different DA methods of CORAL [18], DANN [19], DAAN [20] and DSAN [10] in order to recognize emotions under stimuli of active oil paintings for the cross-subject task. To further analyze the performance of SEED-LOP, the neural patterns under live oil paintings are also investigated.

Table 1. Results of the subject-dependent task on SEED-LOP

Model	SVM	MLP	SST-AGCN	Transformer
mean	53.67	56.47	56.58	**61.73**
std	7.85	8.86	4.46	7.65

Emotion Recognition Performance of DAformer. The mean accuracies and standard deviation of SVM, MLP, SST-AGCN, and Transformer under the subject-dependent task are listed in Table 1 while Table 2 presents the results of SVM, MLP, SST-AGCN, Transformer, DANN, CORAL, DAAN, DSAN and DAformer. The experimental results indicate that for the subject-dependent task, Transformer achieved the best performance with an accuracy of 61.73%/7.65%, while among the nine models for the cross-subject task, DAformer exhibited superior performance with the highest classification accuracy of 54.12%/6.89%.

Table 2. Results of the cross-subject task on SEED-LOP

Model	SVM	MLP	SST-AGCN	Transformer
mean	39.07	47.54	50.59	50.09
std.	8.47	9.86	6.88	6.92

Model	DANN	CORAL	DAAN	DSAN	DAformer
mean	49.23	44.29	46.08	45.54	**54.12**
std.	6.85	8.31	9.31	8.61	6.89

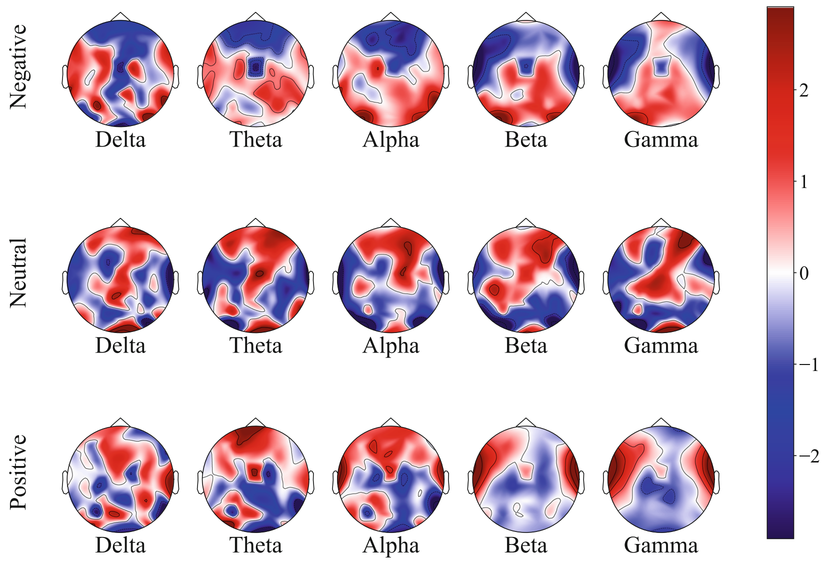

Fig. 4. The average topographic maps of the 40 subjects for three emotion classes with the five frequency bands. The row denotes the different emotion classes and the column denotes the different frequency bands.

Visualization of the Brain Topographic Maps. The neural patterns of the five frequency bands are depicted in Fig. 4. These patterns were derived by averaging the DE features across all 40 participants for each EEG channel. It demonstrates that the lateral temporal areas are more active in the beta and gamma band and the prefrontal sites have higher theta and alpha responses for positive emotions than negative emotions. While the neural patterns of neutral emotions show a greater activation at prefrontal sites for all bands and a stronger gamma response at the occipital sites. We also observed that the negative emotion patterns have significantly higher activation in all bands at the occipital sites and parietal sites. These observations highlight the potential existence of neural patterns associated with the stimuli of animated oil paintings.

Fig. 5. Confusion matrices of experiments. 0: negative, 1: neutral, 2: positive; (a), (b), (c): SVM, SST-AGCN, Transformer in the subject-dependent task, (d), (e), (f): MLP, SST-AGCN, DAformer in the cross-subject task.

Confusion Matrix of the Experiments. To further analyze the experimental outcomes of emotion recognition on SEED-LOP, we present the confusion matrices in Fig. 5. From these confusion matrices, we observe that negative emotion and positive emotion are more readily distinguishable by the model with higher accuracies while there exists a tendency for neutral emotions to be more easily predicted as positive emotion, which aligns with the neural pattern observed in the average topographic map. The topographic map indicates a greater similarity in the neural pattern between neutral and positive emotions as they both show higher activation in prefrontal sites at the theta and alpha bands. This phenomenon also corresponds to the reports of some participants that they believed to have observed a smaller number of neutral emotion paintings in comparison to paintings of the other two emotions. However, we have utilized an equal number of neutral emotion paintings compared to the other two categories in the experiments. This finding suggests that the criteria for rating neutral emotions may be ambiguous for different participants, leading to confusion in the classification task of the models.

4 Conclusion

In this paper, we proposed a novel EEG-based emotion recognition experimental paradigm by utilizing animated oil paintings as stimuli and created the dataset SEED-LOP. The significance of this new dataset lies in its revolutionary utilization of 2D live art pieces as stimuli, which has established a more streamlined and lightweight experimental paradigm in comparison to traditional emotional video clips. We employed Transformer for subject-dependent tasks and introduced DAformer by combining the effectiveness of Transformer with domain adversarial methods for cross-subject tasks to test our dataset SEED-LOP. The experimental results demonstrate that Transformer performs an outstanding accuracy of 61.73% on the subject-dependent task and an improved accuracy of 54.12% of DAformer on the cross-subject task which both serves as an evidence for the effectiveness of the SEED-LOP.

Acknowledgements. This work was supported in part by grants from National Natural Science Foundation of China (Grant No. 61976135), STI 2030-Major Projects+2022ZD0208500, Shanghai Municipal Science and Technology Major Project (Grant No. 2021SHZD ZX), Shanghai Pujiang Program (Grant No. 22PJ1408600), Medical-Engineering Interdisciplinary Research Foundation of Shanghai Jiao Tong University "Jiao Tong Star" Program (YG2023ZD25), and GuangCi Professorship Program of RuiJin Hospital Shanghai Jiao Tong University School of Medicine.

References

1. Alarcao, S.M., Fonseca, M.J.: Emotions recognition using EEG signals: a survey. IEEE Trans. Affect. Comput. **10**(3), 374–393 (2017)
2. Zheng, W.-L., Zhu, J.-Y., Lu, B.-L.: Identifying stable patterns over time for emotion recognition from EEG. IEEE Trans. Affect. Comput. **10**(3), 417–429 (2019)
3. Schaaff, K., Schultz, T.: Towards emotion recognition from electroencephalographic signals. In: 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, pp. 1-6. IEEE (2009)
4. Koelstra, S., Muhl, C., Soleymani, M.: Deap: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2011)
5. Lin, Y.P., et al.: EEG-based emotion recognition in music listening. IEEE Trans. Biomed. Eng. **57**(7), 1798–1806 (2010)

6. Luo, S., Lan, Y.T., Peng, D., Li, Z., Zheng, W.L., Lu, B.L.: Multimodal emotion recognition in response to oil paintings. In: 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 4167-4170. IEEE (2022)
7. Lan, Y.T., Li, Z.C., Peng, D., Zheng, W.L., Lu, B.L.: Identifying artistic expertise difference in emotion recognition in response to oil paintings. In: 11th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 1–4. IEEE (2023)
8. Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., Vaughan, J.W.: A theory of learning from different domains. Mach. Learn. **79**, 151–175 (2010)
9. Li, H., Jin, Y.-M., Zheng, W.-L., Lu, B.-L.: Cross-subject emotion recognition using deep adaptation networks. In: Cheng, L., Leung, A.C.S., Ozawa, S. (eds.) ICONIP 2018. LNCS, vol. 11305, pp. 403–413. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04221-9_36
10. Zhu, Y., et al.: Deep subdomain adaptation network for image classification. IEEE Trans. Neural Networks Learn. Syst. **32**(4), 1713–1722 (2021)
11. Wang, Y., Jiang, W. B., Li, R., Lu, B.L.: Emotion transformer fusion: complementary representation properties of EEG and eye movements on recognizing anger and surprise. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1575–1578. IEEE (2021)
12. Li, R., Wang, Y., Lu, B.L.: Measuring decision confidence levels from EEG using a spectral-spatial-temporal adaptive graph convolutional neural network. In: Tanveer, M., Agarwal, S., Ozawa, S., Ekbal, A., Jatowt, A. (eds.) ICONIP 2022. LNCS, pp. 395–406. Springer, Cham (2022). https://doi.org/10.1007/978-981-99-1642-9_34
13. Schaefer, A., et al.: Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. Cogn. Emot. **24**(7), 1153–1172 (2010)
14. Zheng, W.L., Lu, B.L.: Personalizing EEG-based affective models with transfer learning. In: International Joint Conference on Artificial Intelligence, pp. 2732–2738. AAAI Press, New York (2016)
15. Batt, R., Palmiero, M., Nakatani, C., van Leeuwen, C.: Style and spectral power: processing of abstract and representational art in artists and non-artists. Perception **39**(12), 1659–1671 (2010)
16. Chatterjee, A., Widick, P., Sternschein, R., Smith, W.B., Bromberger, B.: The assessment of art attributes. Empir. Stud. Arts **28**(2), 207–222 (2010)
17. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
18. Sun, B., Saenko, K.: Deep CORAL: correlation alignment for deep domain adaptation. In: Hua, G., Jégou, H. (eds.) ECCV 2016. LNCS, vol. 9915, pp. 443–450. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49409-8_35
19. Ganin, Y., et al.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. **17**(1), 2030–2096 (2016)
20. Wang, J., Feng, W., Chen, Y., Yu, H., Huang, M., Yu, P.S.: Visual domain adaptation with manifold embedded distribution alignment. In: Proceedings of the 26th ACM International Conference on Multimedia, pp. 402-410 (2018)