Objective Depression Detection Using EEG and Eye Movement Signals Induced by Oil Paintings

Luyu Liu, Dan Peng, Wei-Long Zheng and Bao-Liang Lu* *Fellow, IEEE*

***Abstract*— Depression is a mental disorder characterized by persistent sadness and loss of interest, which has become one of the leading causes of disability worldwide. There are currently no objective diagnostic standards for depression in clinical practice. Previous studies have shown that depression causes both brain abnormalities and behavioral disorders. In this study, both electroencephalography (EEG) and eye movement signals were used to objectively detect depression. By presenting 40 carefully selected oil paintings—20 positive and 20 negative—as stimuli, we were able to successfully evoke emotions in 48 depressed patients (DPs) and 40 healthy controls (HCs) from three centers. We then used Transformer, a deep learning model, to conduct emotion recognition and depression detection. The experimental results demonstrate that: a) Transformer achieves the best accuracies of 89.21% and 92.19% in emotion recognition and depression detection, respectively; b) The HC group has higher accuracies than the DP group in emotion recognition for both subject-dependent and subject-independent experiments; c) The neural pattern differences do exist between DPs and HCs, and we find the consistent asymmetry of the neural patterns in DPs; d) For depression detection, using single oil painting achieves the best accuracies, and using negative oil paintings has higher accuracies than using positive oil paintings. These findings suggest that EEG and eye movement signals induced by oil paintings can be used to objectively identify depression.**

I. INTRODUCTION

Depression is one of the most common and recurrent psychological disorders and a leading contributor to the global burden of diseases, which is characterized by persistent sadness, loss of interest, hopelessness, and inability or reduced ability to experience pleasure in normally enjoyable activities [1]. It significantly affects a person's family, personal relationships and other general health aspects. However, current diagnostic methods of depression are human-intensive and rely almost exclusively on the evaluation of symptoms using questionnaires and interviews, and the results are dependent on doctors' experience, may leading to subjective biases. There are currently no objective diagnostic standards for depression in clinical practice.

Depression can not only cause brain abnormalities, but also leads to behavioral disorders. Many imaging technologies, including electroencephalography (EEG) [2], functional magnetic resonance imaging (fMRI) [3], and positron emission tomography (PET) [4] etc., have been widely used to explore the abnormal brain activity of depression. EEG has the characteristics of higher time resolution and lower cost compared with fMRI and PET, which is more possible to be promoted for routine use. Detecting depression from behavioral cues, including vocal utterance [5], facial expression [6], head movements [7], and gaze [8], is of increasingly interest. Eye movement signals are powerful for depression detection. For example, depression is characterized by reduced maintenance of gaze on positive stimuli [9]. Because of the complexity of depression, analysis from only one aspect is not comprehensive. In this study, we use EEG and eye movement signals to perform depression detection since they can represent emotions more comprehensively.

In the literature, existing studies have used the abnormalities of resting-state EEG to characterize individuals with depression [2]. However, the ability of resting-state EEG to distinguish individuals with depression is limited [10]. To address this limitation, employing a depression-sensitive task to elicit EEG responses that differ strongly between depressed and healthy individuals proves helpful. Luo *et al.* [11] successfully used oil paintings as stimulate materials to induce three types of emotions in healthy individuals. In this study, we collected EEG and eye movement signals as depressed patients (DPs) and healthy controls (HCs) engaged with the selected positive and negative oil paintings, and applied Transformer to conduct emotion recognition and depression detection tasks. To our knowledge, no prior study has used EEG and eye movement signals induced by oil paintings to detect depression.

II. EXPERIMENT DESIGN

A. Stimuli

The stimuli used in this experiment were 40 oil paintings, selected from the 114 oil paintings [11]. We consider the following two main reasons for adopting these oil paintings as stimulus materials: 1) oil paintings are rich and diverse; 2) they make experimental paradigm more closer to real life and more interesting. These oil paintings were created from the mid-16th century to the 19th century and were collected from the publicly available artwork dataset WikiArt, covering

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. 2021SHZDZX), Shanghai Pujiang Program (Grant No. 22PJ1408600), SJTU Global Strategic Partnership Fund (2021 SJTU-HKUST) and GuangCi Professorship Program of RuiJin Hospital Shanghai Jiao Tong University School of Medicine.

L Liu, W.-L Zheng, and B.-L. Lu are with Center for Brain-Like Computing and Machine Intelligence, Department of Computer Science and Engineering, the Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, and Brain Science and Technology Research Center, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai 200240, People's Republic of China.

D. Peng and B.-L. Lu are with the RuiJin-Mihoyo Laboratory, Clinical Neuroscience Center, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin 2nd Rd., Shanghai 200020, People's Republic of China.

*Corresponding author (bllu@sjtu.edu.cn).

most of the major art-styles and genres. Luo *et al.* [11] adjusted these oil paintings uniformly to reduce the impact of distracting factors and distributed questionnaires to get their emotion labels among students of China Academy of Art and Shanghai Jiao Tong University. Based on these questionnaires, we selected 20 positive and 20 negative oil paintings with the highest consistency and intensity.

B. Experimental Procedure

The experimental protocol is shown in Fig. 1, which has been approved by the local ethics committee. At first, we introduced the experiment and showed two examples of positive painting and negative painting, which would help the subjects get prepared for the whole experiment. There were 40 trials in total, and the display sequence of the oil paintings was randomized. For each trial, there was a 3-s hint of process. Each oil painting was displayed for 20 s and the subjects were told to freely view the painting. After the presentation of each oil painting, the subjects were required to rate their feelings for the painting with valence (1 to 9) and arousal (1 to 9), and the rating process had no time limit.

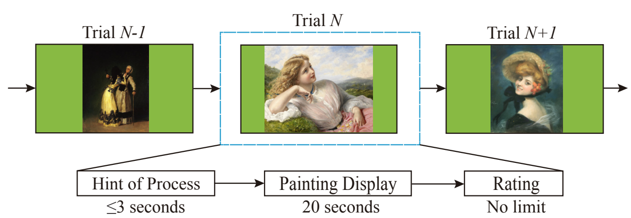


Fig. 1. The procedure of the whole experiment, including 40 trials of process hint, painting display and rating.

During the experiment, EEG and eye movement signals were collected simultaneously. The EEG data were recorded using the DSI-24[1] dry electrode EEG system at a 1000 Hz sampling rate. The DSI-24 conforms to the international 10-20 system and has 19 electrodes on the head, 2 earclip sensors, and 3 built-in auxiliary inputs which were not used in this experiment. Among the 21 channels, three were references, leaving a total of 18 EEG channels used for further analysis. The eye movements data were recorded using the screen-based eye tracker Tobii Pro Fusion[2] at a sampling rate of 250 Hz.

C. Subjects

We conducted the experiment in three medical centers, including Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Taizhou Second People's Hospital, and Shanghai Mental Health Center. The HC subjects were recruited from the public community. Data were acquired after obtaining informed consent from the subjects. Each subject was required to complete a depression assessment conducted by psychiatrists. In this paper, we used the data of 88 subjects with high data quality, which consist of 48 DP subjects (15 males and 33 females; age 21.75 ± 6.69 years) and 40 HC subjects (9 males and 31 females; age 25.35 ± 3.71 years). Table I lists the basic information.

TABLE I

BASIC INFORMATION OF THE SELECTED SUBJECTS FROM THREE MEDICAL CENTERS.

Center	Variable	DP	HC	p-value
Ruijin Hospital	Gender	4 M, 3 F	4 M, 28 F	0.007
	Age (Years)	25.43 ± 5.44	26.13 ± 3.74	0.684
Taizhou Second People's Hospital	Gender	10 M, 18 F	5 M, 3 F	0.185
	Age (Years)	19.36 ± 6.15	22.25 ± 1.04	0.198
Shanghai Mental Health Center	Gender	1 M, 12 F	-	-
	Age (Years)	24.92 ± 6.59	-	-
All	Gender	15 M, 33 F	9 M, 31 F	0.365
	Age (Years)	21.75 ± 6.69	25.35 ± 3.71	0.003

III. METHODOLOGY

A. Data Preprocessing and Feature Extraction

For raw EEG signals, we used a bandpass filter between 1 Hz and 50 Hz to eliminate the noise, and downsampled the signals from 1000 Hz to 200 Hz to reduce computation complexity. We extracted differential entropy (DE) features [12] using short time fourier transform (STFT) with a 4-s Hanning window with 3-s overlapping. The DE features were divided into five frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), beta (14–31 Hz), and gamma (31–50 Hz). There were $18 \times 5 = 90$ dimensions for 18 EEG channels in total. The DE features were smoothed by the linear dynamic system (LDS) method [13].

For eye movement signals, we extracted 33 eye movement features [14], including pupil diameter, dispersion, fixation, blink and saccade. The LDS method was also used to smooth the extracted features.

B. Classification Models

In this study, we used Transformer for both emotion recognition and depression detection. Our Transformer model used the naive Transformer encoder to extract features and used a fully connected layer as a classifier to detect emotions or depression instead of a decoder. We used an overlapping time window on the preprocessed signals for they were too long to be fed into the network directly. There were four hyper-parameters in our Transformer architecture: size of overlapping time window, number of layers, number of heads of attention, and the learning rate, which were chosen from [2, 4, 5], [2, 4], [2, 4, 8] and $10^{[-2,-3,-4,-5]}$, respectively. We used the support vector machine (SVM) with a linear kernel as the baseline model and used grid search to find the best C from 10^{-5} to 10^{4}.

C. Emotion Recognition and Depression Detection

In this study, we performed two classification tasks, namely emotion recognition and depression detection, to see the differences of EEG and eye movement signals induced by oil paintings between DPs and HCs.

[1]https://wearablesensing.com/dsi-24/

[2]https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-fusion

1) Emotion Recognition: For the emotion recognition experiment, we used the subjects' own valence ratings as the corresponding emotion labels with a threshold of 5. We performed the subject-dependent and subject-independent emotion recognition experiments in the DP group and HC group, respectively. For subject-dependent experiment, we used a five-fold cross-validation to evaluate the performance of each subject, and averaged the results of the same group for comparison. For subject-independent experiment, we used a leave-one-out cross-validation to evaluate the performance of trained models.

2) Depression Detection: We tried to use the selected oil paintings to classify DPs and HCs based on the assessment results given by psychiatrists. We first used single oil painting, the data of different subjects watching the same individual oil painting (20 s), to identify DPs and averaged the 40 results for comparison. Then we used the same valence oil paintings (positive or negative, $20 \times 20 = 400$ s) to conduct classification to evaluate the differences when watching different kinds of paintings. Finally, we used all 40 oil paintings (800 s) to conduct depression detection, to evaluate the feasibility of identifying DPs using oil paintings. We used a 20-fold cross-validation for all these experiments and took the average accuracies for comparison.

IV. Results and Discussion

A. Emotion Recognition

Table II presents the results of SVM and Transformer in the subject-dependent task. Transformer achieves the highest accuracies in the HC group for all three kinds of modalities, 86.53 ± 10.27% for EEG, 86.10 ± 8.64% for eye movement signals, and 89.21 ± 7.93% for the combination of EEG and eye movement signals. Regardless of the various modalities and models tested, the HC group outperforms the DP group, and the accuracy of the HC group's Transformer utilizing the combination of EEG and eye movements is significantly ($p<0.01$) higher than that of the DP group.

TABLE II

The average accuracies (mean ± std)% of different groups and models in subject-dependent emotion recognition.

Group	Model	EEG	Eye Movements	EEG + Eye
DP	SVM	69.95 ± **7.12**	75.76 ± **7.29**	74.56 ± **7.04**
	Transformer	**85.77** ± 10.26	**84.49** ± 9.07	**86.66** ± 9.01
HC	SVM	71.27 ± **6.51**	77.54 ± **6.65**	76.73 ± **5.99**
	Transformer	**86.53** ± 10.27	**86.10** ± 8.64	**89.21** ± 7.93

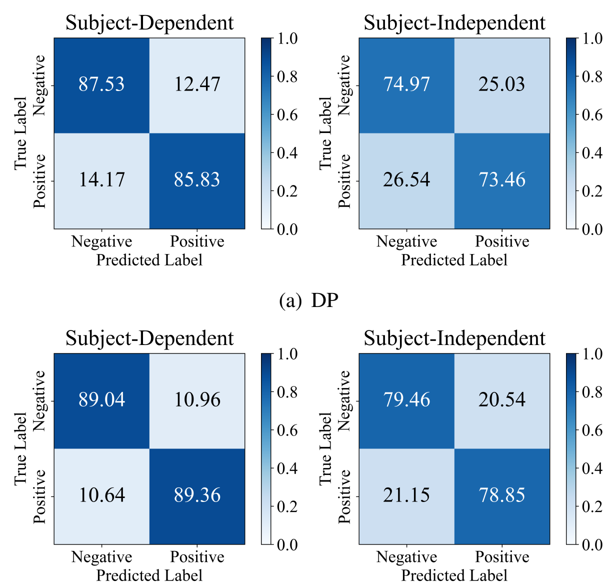
For the subject-independent task in Table III for both SVM and Transformer, using the combination of EEG and eye movement signals achieves better performance than eye movements and significantly better performance than EEG ($p<0.001$), which indicates that the two modalities have complementary characteristics for emotion recognition. Transformer obtains the best performance of 69.56 ± 5.59%, 78.03 ± 7.04% and 79.15 ± 6.35% for EEG, eye movements and the combination of EEG and eye movements, respectively. Similarly, the HC group has higher accuracies than the DP group for all three modalities on both models, and the Transformer results of HC group using eye movement and using the combination of EEG and eye movements are significantly higher than those of the DP group ($p<0.01$ and $p<0.001$, respectively).

TABLE III

The average accuracies (mean ± std)% of different groups and models in subject-independent emotion recognition.

Group	Model	EEG	Eye Movements	EEG + Eye
DP	SVM	59.93 ± 7.94	67.58 ± 7.70	69.15 ± 7.44
	Transformer	**69.33 ± 5.22**	**73.57 ± 6.38**	**74.20 ± 5.72**
HC	SVM	61.23 ± 6.92	73.10 ± 8.21	74.00 ± 6.44
	Transformer	**69.56 ± 5.59**	**78.03 ± 7.04**	**79.15 ± 6.35**

Fig. 2 presents the confusion matrices of Transformer using the combination of EEG and eye movements for emotion recognition for the two groups. As illustrated in Fig. 2(a), EEG and eye movement signals can better distinguish negative emotion compared with positive emotion for DPs. By comparing Fig. 2(a) and Fig. 2(b), we can see that the accuracies of HCs are all higher than those of DPs for both negative and positive emotions.



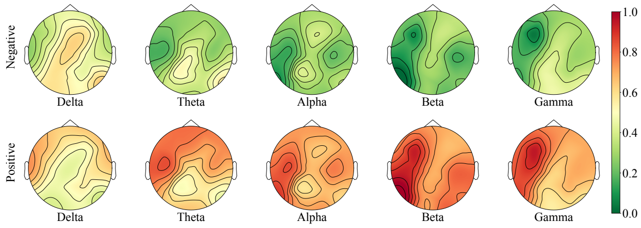
(a) DP

(b) HC

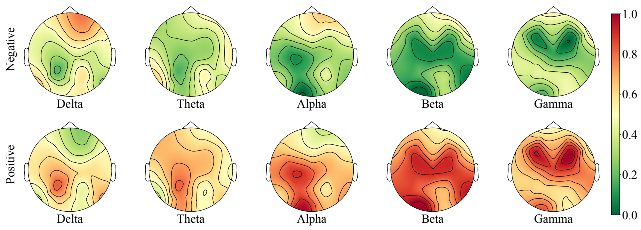
Fig. 2. The confusion matrices of DPs and HCs for subject-dependent and subject-independent emotion recognition on Transformer using the combination of EEG and eye movements.

B. Neural Patterns

To further study the differences between the DP and HC groups in recognizing emotions, we visualize the brain topography in Fig. 3. We find that for both DP and HC subjects, positive emotion has higher activation than negative emotion in all frequency bands. For DPs, we find high activation in the alpha, beta, and gamma bands at the left hemisphere for positive emotion, which is related to the alpha asymmetry [10], a possible EEG marker of depression. For the positive emotion of HCs, the prefrontal lobe is less activated in the beta and gamma bands, and the temporal lobes on both hemispheres have higher activation, which is consistent with the previous research [13].



(a) DP



(b) HC

Fig. 3. The average neural patterns of the DP and HC groups. The EEG features are normalized and averaged within each group, respectively.

C. Depression Detection

For depression detection, four subtasks using SVM and Transformer are performed and the experimental results are shown in Table IV. Transformer using single oil painting obtains the best performance, with accuracies of 91.67 ± 10.44%, 91.71 ± 9.75% and 92.19 ± 10.22% for EEG, eye movements and the combination of EEG and eye movements, respectively. The performance is much better than those of using positive, negative and all oil paintings ($p<0.001$), for they are less affected by the differences between different oil paintings. When using the same valence oil paintings to identify depression using Transformer, positive oil paintings get the accuracies of 78.71 ± 8.03%, 72.82 ± 7.11%, 79.69 ± 7.67%, and negative oil paintings get 79.16 ± 8.69%, 74.33 ± 5.85% and 81.09 ± 6.87% for EEG, eye movements and the combination of EEG and eye movements, respectively. The results indicate that negative oil paintings have better ability to distinguish DPs from HCs.

TABLE IV

THE AVERAGE ACCURACIES (MEAN ± STD)% IN DEPRESSION DETECTION USING FOUR TYPES OF OIL PAINTINGS.

Type	Model	EEG	Eye Movements	EEG + Eye
Single	SVM	71.42 ± 16.02	74.34 ± 14.15	71.64 ± 15.85
	Transformer	**91.67 ± 10.44**	**91.71 ± 9.75**	**92.19 ± 10.22**
Positive	SVM	62.66 ± 8.59	64.35 ± 9.49	62.59 ± 8.27
	Transformer	78.71 ± 8.03	72.82 ± 7.11	79.69 ± 7.67
Negative	SVM	65.28 ± 9.92	63.27 ± 7.66	64.88 ± 10.06
	Transformer	79.16 ± 8.69	74.33 ± 5.85	81.09 ± 6.87
All	SVM	63.47 ± 8.79	63.43 ± 8.32	62.93 ± 8.76
	Transformer	76.54 ± 8.58	72.11 ± 5.53	77.14 ± 7.66

V. CONCLUSIONS

In this study, we aim at developing an objective depression detection approach that can assist clinical diagnosis. We have collected the EEG and eye movement signals of DPs and HCs induced by oil paintings, and applied Transformer to conduct the emotion recognition and depression detection tasks. From the experimental results, we find that for both subject-dependent and subject-independent emotion recognition, the HC group has higher accuracies than the DP group. From the brain topographic mapping, we find that positive emotion has higher activation than negative emotion for both DPs and HCs. The asymmetry of the neural patterns in DPs is consistent to previous studies to some extent, which needs to be further studied. When using oil paintings to identify depression, single oil painting has better performance, and the results of negative oil paintings are better than those of positive oil paintings, which pave the way for further study. Our study provides an effective method for depression detection, which in the future could be used for the auxiliary diagnosis of depression.

REFERENCES

[1] A. P. Association., *Diagnostic and statistical manual of mental disorders : DSM-5*, 5th ed. American Psychiatric Association, 2013.

[2] F. S. de Aguiar Neto and J. L. G. Rosa, "Depression biomarkers using non-invasive EEG: A review," *Neuroscience & Biobehavioral Reviews*, vol. 105, pp. 83–93, 2019.

[3] H. Zheng, F. Li, Q. Bo, X. Li, L. Yao, Z. Yao, C. Wang, and X. Wu, "The dynamic characteristics of the anterior cingulate cortex in resting-state fMRI of patients with depression," *Journal of Affective Disorders*, vol. 227, pp. 391–397, 2018.

[4] M. Dubol, C. Trichard, C. Leroy, B. Granger, E. T. Tzavara, J.-L. Martinot, and E. Artiges, "Lower midbrain dopamine transporter availability in depressed patients: Report from high-resolution PET imaging," *Journal of Affective Disorders*, vol. 262, pp. 273–277, 2020.

[5] T. Taguchi, H. Tachikawa, K. Nemoto, M. Suzuki, T. Nagano, R. Tachibana, M. Nishimura, and T. Arai, "Major depressive disorder discrimination using vocal acoustic features," *Journal of Affective Disorders*, vol. 225, pp. 214–220, 2018.

[6] X. Zhou, K. Jin, Y. Shang, and G. Guo, "Visually interpretable representation learning for depression recognition from facial images," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 542–552, 2020.

[7] S. Alghowinem, R. Goecke, M. Wagner, G. Parkerx, and M. Breakspear, "Head pose and movement analysis as an indicator of depression," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 283–288.

[8] S. Alghowinem, R. Goecke, M. Wagner, G. Parker, and M. Breakspear, "Eye movement analysis for depression detection," in *2013 IEEE International Conference on Image Processing*, 2013, pp. 4220–4224.

[9] T. Armstrong and B. O. Olatunji, "Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis," *Clinical Psychology Review*, vol. 32, no. 8, pp. 704–723, 2012.

[10] N. Van Der Vinne, M. A. Vollebregt, M. J. Van Putten, and M. Arns, "Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? a meta-analysis," *Neuroimage: Clinical*, vol. 16, pp. 79–87, 2017.

[11] S. Luo, Y.-T. Lan, D. Peng, Z. Li, W.-L. Zheng, and B.-L. Lu, "Multimodal emotion recognition in response to oil paintings," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 4167–4170.

[12] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 81–84.

[13] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 417–429, 2017.

[14] Y. Lu, W.-L. Zheng, B. Li, and B.-L. Lu, "Combining eye movements and EEG to enhance emotion recognition," in *IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press, 2015, pp. 1170–1176.