Feature-level Fusion for Depression Recognition Based on fNIRS Data

Shuzhen Zheng, Chang Lei, Tao Wang, Chunyun Wu, Jieqiong Sun, Hong Peng*
Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering
Lanzhou University, Lanzhou, China
{zhengshzh19, leich19, wangtao2018, wuchy18, sunjq18, pengh }@lzu.edu.cn

2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 978-1-7281-6215-7/20/$31.00 ©2020 IEEE DOI: 10.1109/BIBM49941.2020.9313549

***Abstract*—Tens of millions of people suffer from depression worldwide. It is urgent to explore an effective method for diagnosing depression. This study developed a novel of multi-modal feature fusion depression recognition method based on functional near-infrared spectroscopy (fNIRS). Sixty volunteers, including thirty patients with depression and thirty healthy controls, participated in the study. The 22-channel fNIRS device recorded the participants' brain oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) concentration changes in the positive, neutral and negative affective words' stimulation. K-nearest neighbors (KNN) and support vector machine (SVM) classifiers were used to recognize depressed patients from normal people, and 10-fold cross-validation was used to verify the classification result. Under the three single-mode features, the accuracy rates were 85.69%, 88.32% and 86.77%, corresponding to the positive condition, neutral condition and negative condition. Then, we used concatenation and linear combination for feature fusion. For the concatenation fusion method, the principal component analysis (PCA) was used to reduce the dimension. The result showed that feature fusion can relatively improve the recognition rate of people with depression, compared with single-model features. The optimal feature fusion method is to concatenate the neutral features and negative features, and the best accuracy reaches 94.45%. The study may provide a more accurate and convenient method for depression detection.**

***Index Terms*—depression, functional near-infrared spectroscopy (fNIRS), feature fusion, support vector machine (SVM), k-nearest neighbor (KNN)**

I. Introduction

Depression is a common mental illness characterized by persistent low mood and slow thinking states [1] [2]. According to incomplete statistics, 340 million human beings suffer from depression pain [3] [4]. Depression has a serious impact on patients' mental and physical health [5]. It is reported that approximately one million people commit suicide every year due to depression [6]. Depression can be measured and evaluated based on perceptual, behavioral and physiological responses [7]. Traditionally, depression was diagnosed by a doctor based on a questionnaire. However, the method of evaluating depression based on questionnaires is more subjective and more dependant on doctors clinical experience [8] [9]; thus finding an objective and effective depression detection method is essential. Assessing depression by monitoring changes in physiological signals is an objective and convincing method [10]. Functional near-infrared spectroscopy (fNIRS) is one of the most commonly used neuroimaging techniques for researching brain activities and conditions [11] [12] and it is commonly used to detect depression.

The fNIRS device records the changes in optical intensity before and after infrared light passes through the cerebral cortex. Then the modified Beer-Lambert law is utilized to convert the optical intensity to concentrations of oxygen-hemoglobin (HbO) and deoxygen-hemoglobin (HbR) [13]. Studies have shown that when stimulated, the concentration of HbO in the brain increases and HbR decreases [14]. For the same stimulus task, people with depression and healthy people have different levels of prefrontal activation [29]. Several experts have performed research on identifying depression based on fNIRS signals. In 2014, Ryu Takizawa *et al*. [16] conducted a multi-site study and the results showed that the frontal hemodynamic detected by fNIRS accurately distinguished people with depression from normal people. This study demonstrated that fNIRS may be used as a tool in assisting the diagnosis of major depression. Song *et al*. [17] achieved a classification accuracy of 86.77% using SVM to distinguish between MDD patients and healthy adults from fNIRS data collected while participants performed a cognitive task, which illustrated the feasibility of discriminating people with depression from healthy people through fNIRS signals.

Previous studies on depression recognition by fNIRS extracted features in a single mode. However, the features from a single mode cannot fully reflect the brain activities of the subjects. In recent years, the technology of combining information in multiple modalities through feature fusion has become increasingly popular in various areas [18]. The fusion of EEG signals under positive and negative audio stimuli obtained higher depression recognition accuracy rates compared with individual modalities [19]. Alghowinem *et al*. [20] fused paralinguistic, head pose and eye gaze features for depression detection and found that fusing all modalities shows a remarkable improvement compared to the unimodal system. Their research results show that fused features show better performance than single-mode features in identifying depression.

This paper proposed a method of fusing fNIRS data in different modes to obtain the normalized features in the combined mode and created a depression classification model under multimodal fNIRS data. We first extracted fNIRS features of depressed patients and normal people in three single modes of positive, neutral and negative. Then SVM and KNN were

used to classify patients and the healthy in single modes. On the basis of single-mode features, we used linear combination and concatenation to perform feature fusion of different modes. By comparing the classification accuracy of single modes and fusion modes, we obtained the optimal combination method.

II. Materials and methods

A. Participant

Sixty subjects, including 30 patients with depression, and 30 normal controls, participated in the study. All depressed subjects were selected from the diagnosed patients in the hospital and their PHQ-9 scores were greater than or equal to 5, diagnostic criteria of MINI met the criteria of depression. The depression group ranged in age 32±9.91, the control group ranged in age 39±9.5. All participants signed informed consent forms before the experiment.

B. Experimental Procedure

The experiment was conducted in a quiet and empty room. During the experiment, the subjects were asked to stare at the words that appeared on the screen and then press button based on the word's color. The experiment consisted of three parts. Pre-experiment: Before the formal experiment, a 12-word preliminary experiment was used to familiarize the subjects with the experimental process. Resting time: To ensure that the patient's attention was focused during the task, there was a rest period before the start of each stimulation task. Stimulus task: The formal experiment involved three kinds of stimulation: positive word stimuli, neutral word stimuli and negative word stimuli. Thirty-two trials for each stimulus, and each word was displayed for five seconds£after each word ended, a "+" appeared for 2 seconds on the screen. Fig. 1 shows experimental flowchart. As displayed in Fig. 2, the 22-channel fNIRS data of the prefrontal lobe were collected synchronously in this study.

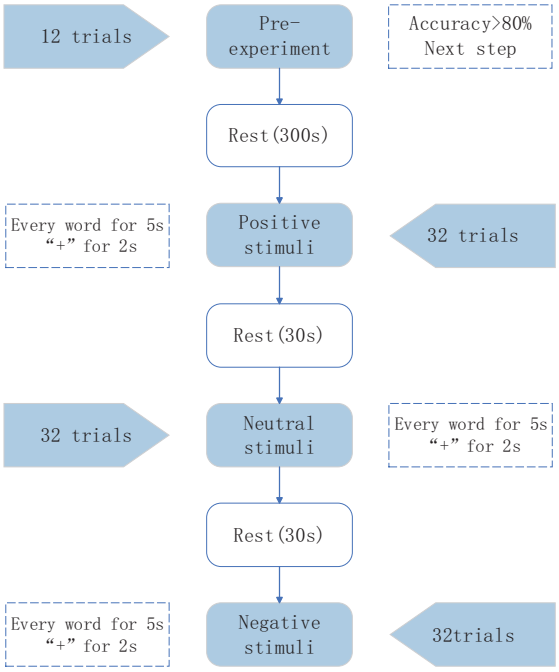


Fig. 1: Experimental flowchart.

C. FNIRS Data Preprocessing

The raw signals recorded by fNRIS equipment are two-wavelength (760nm and 850nm) optical intensity of 22 channels of prefrontal lobe. The sampling rate is 7.81Hz in the experiment. In the process of collecting experimental data, it is inevitable that some noise signals will be introduced. There are generally three types of noise in functional near-infrared signals (fNIRS): (a). physiological signals, including heartbeat (1Hz-2Hz), respiration (0.4Hz) and blood pressure (0.1Hz) signals; (b). high-frequency noise caused by hair interference and the contact of optodes to the head; (c). motion artifacts are mainly composed of head movement, eye movement and blinking. There are three manifestations of motion artifacts: a high-frequency spike, a shift from baseline intensity and low-frequency variations. Among them, high amplitude and high frequency spikes are easily detectable and removed, yet lower frequency content is harder to distinguish from normal hemodynamic signals. Commonly used methods to remove motion artifacts are spline interpolation [21], principal component analysis (PCA) [22], wavelet filtering [23], correlation-based signal improvement(CBSI) [24] and Kalman filtering [25]. In this article, we used spline interpolation to remove artifacts. The preprocessing was completed in the Homer2 toolbox of MATLAB. The detialed flow is as follows.

1) Bad trials/channels rejection: Remove bad trials/channels through visual inspection and calculation of coefficient of variation ($CV = \frac{\sigma}{\mu} \times 100\%$, where μ represents the mean value of trial or channel, and σ represents the standard deviation. Rejection trials with $CV_{trial} > 5\%$ and channels with $CV_{channel} > 15\%$).
2) Optical intensity to optical density (OD): Convert the optical intensity signals to optical density according to the modified Lambert-Beer Law ($OD = log(I_O/I_I)$, I_O refers to the intensity of outgoing light, I_I refers to the intensity of the input light).
3) Motion correction: Spline interpolation (First, detect artifacts segment by the moving standard deviation. The period of motion artifact is then modeled via a cubic spline interpolation. Finally, subtract the interpolation from the original signal and adjust the mean value).
4) Bandpass filter (0.01Hz-0.2Hz).
5) OD to hemoglobin (HbO, HbR).
6) Block average (-1s 6s): Calculate the block average given the stimulation conditions in seconds over the time range. The baseline of the average is set to zero by subtracting the mean of the average for t<0. If a stimulus occurs too close to the start or the end of the data such that time range extends outside of the data range, then the trial is excluded from the average.

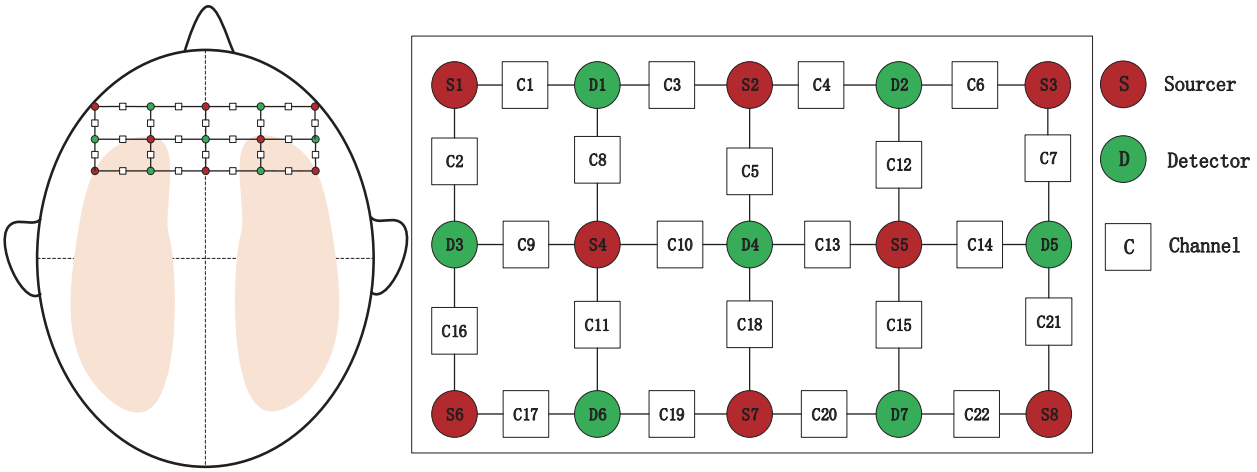


Fig. 2: Electrode location and channel map.

As the preprocessing was completed, we obtained the pure block average signal of 60 subjects. Next, we will extract features for each channel of 60 subjects.

D. Feature Extraction

We extracted five metrics of HbO and HbR respectively. They were the mean value, standard deviation, beta value, area under curve and the left slope. Thus we obtained a 600×22 dimensions feature matrix, where 600 ($60 subjects \times 5 metrics \times 2 Hb$) refers to the number of samples, and 22 represents the number of features for each sample. The feature matrix was normalized by the mapminmax function in MATLAB. The function formula is as equation (1)

$$y = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

E. Feature Fusion

Information fusion can be divided into three levels: feature-level fusion, score-level fusion and decision-level fusion. Feature-level fusion refers to combining features from different modes to obtain fusion features. Score fusion refers to the fusion of matching scores corresponding to each feature in a single mode. Decision-level fusion refers to making decisions on different feature subsets and then synthesizing their results to obtain the final classification result.

Feature-level fusion was utilized in the present study. There are two feature fusion methods: linear combination and concatenation. In the previous section, we extracted the HbO and HbR features of the subjects under the positive, neutral and negative stimuli respectively, and obtained the feature matrix under the three single modes. Denoted as $U_{pos} = (u_1, u_2, ..., u_m)$, $V_{neu} = (v_1, v_2, ..., v_m)$, $W_{neg} = (w_1, w_2, ..., w_m)$. Where, U_{pos} represents the feature matrix in the positive emotion word stimulation mode, V_{neu} represents the feature matrix in the neutral emotion word stimulation mode, and W_{neg} represents the feature matrix in the negative emotion word stimulation mode.

The fusion matrix obtained by linear combination can be expressed as follows:

$$L_{pos-neu} = k_1 \times U_{pos} + k_2 \times V_{neu} \tag{2}$$

$$L_{pos-neg} = k_1 \times U_{pos} + k_3 \times W_{neg} \tag{3}$$

$$L_{neu-neg} = k_2 \times V_{neu} + k_3 \times W_{neg} \tag{4}$$

$L_{pos-neu}$ represents the linear combination of the feature matrix under positive stimulation and neutral stimulation, $L_{pos-neg}$ represents the linear combination of the feature matrix under positive stimulation and negative stimulation, $L_{neu-neg}$ represents the linear combination of feature matrix under neutral stimulation and neutral stimulation. Where, k_1, k_2, and k_3 are linear combination coefficients, which represent the weight coefficients of the feature matrix under three modes of positive, neutral and negative, respectively. In this article, weight coefficient used two value methods:

1) Assume that the three single-mode features have equal weights, that is, $k_1 = k_2 = k_3$.
2) The weights of the three single-mode features are different. The classification accuracy of the three single-mode features served as the combination coefficient. k_1, k_2, and k_3 respectively represented the classification accuracy of the positive stimuli, neutral stimuli and negative stimuli.

In the concatenation method, the features from the different single modes were concatenated along the horizontal direction to yield a multimodal feature matrix. In our study, the fusion matrix obtained by concatenation can be expressed as equation(5)(6)(7)

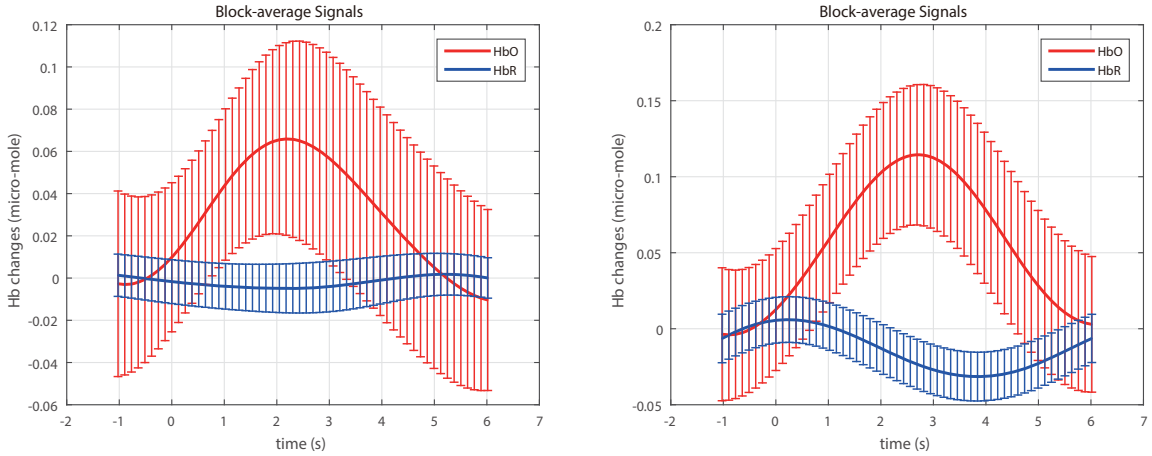


Fig. 3: (a)The hemoglobin of subject 9 channel 1 in depression group (b)The hemoglobin of subject 45 channel 1 in normal group

$$P_{pos-neu} = [U_{pos} \quad V_{neu}] \tag{5}$$

$$P_{pos-neg} = [U_{pos} \quad W_{neg}] \tag{6}$$

$$P_{neu-neg} = [V_{neu} \quad W_{neg}] \tag{7}$$

where $P_{pos-neu}$ represents the fusion matrix obtained by concatenating the feature matrix in the positive stimulation mode and the neutral stimulation mode, $P_{pos-neg}$ represents the fusion matrix obtained by concatenating the feature matrix in the positive stimulation mode and the negative stimulation mode, $P_{neu-neg}$ represents the fusion matrix obtained by concatenating the feature matrix in the neutral stimulation mode and the negative stimulation mode. Since the dimension of the fusion matrix obtained by concatenating will increase to twice the original, principal component analysis (PCA) was used to reduce the dimension of the fusion matrix obtained by concatenating.

Principal component analysis (PCA): The principal component analysis method is one of the most widely used data dimensionality reduction algorithms. The main idea of PCA is to map n-dimensional features to k-dimensions. This k-dimension is a brand-new orthogonal features, also called principal component, which is reconstructed on the basis of the original n-dimensional features [26].

F. Classifier

1) KNN: KNN is a commonly used supervised learning method, and its working mechanism is very simple: find the K points in the training set that are closest to the test sample, and according to the principle of voting, select the category that appears more in these K samples as the prediction result [27].

2) SVM: The basic idea of SVM for binary classification is to find a hyperplane that can divide the training samples into two categories with the largest interval [28], and then predict the category of the test sample based on the position of the test sample in the feature space. The kernel function used in this study was "RBF".

In this study, accuracy, precision and recall as metrics were utilized to assess the classification performance of the model.

III. Results

A. Single mode

Some research results have shown that when stimulated by the outside world, the concentration of HbO in the human brain will increase, while the concentration of HbR will decrease, and the brain activation levels of healthy people and patients with depression are different. As shown in Fig. 3 (a) displays oxyhemoglobin and deoxyhemoglobin concentration of subject 9 channel 1 in the depression group, (b) displays the oxyhemoglobin and dexhemoglobin concentration of subject 45 channel 1 in the normal group. Thus, we extracted the features of HbO and HbR under the positive, neutral and negative stimuli. Then the extracted features were fed into SVM and KNN to recognize the depressed patients from healthy people under the three single modes. In Table I, the classification accuracy, precision and recall under single modes are displayed. Among them, U_{pos} refers to the feature matrix under the positive stimulus, V_{neu} refers to the feature matrix under the neutral stimulus, W_{neg} refers to the feature matrix under the negative stimulus.

B. Fusing mode

Because the features collected in a single mode cannot fully represent the characteristics of the sample, that is, it cannot fully reflect the separability information of the sample. In addition, the brain activation area are different under different types of emotional words stimuli [29]. Through multimode fusion, multiple features can achieve complementary advantages, thereby improving the classification accuracy of system differentiation.

TABLE I: Accuracy, precision and recall of all single modes and fusing modes.

	Accuracy(%)			Precision(%)			Recall(%)		
	KNN	SVM	mean	KNN	SVM	mean	KNN	SVM	mean
U_{pos}	86.15	85.24	85.70	84.79	83.87	84.33	87.45	86.67	87.06
V_{neu}	87.23	89.41	**88.32**	84.43	89.19	**86.81**	90.56	90.91	**90.74**
W_{neg}	85.23	88.32	86.78	84.11	81.82	82.97	87.04	84.38	85.71
$A_{pos-neu}$	85.39	87.19	86.29	86.87	88.89	87.88	86.06	88.89	87.48
$A_{pos-neg}$	85.69	89.00	87.35	89.00	85.71	87.44	84.89	88.24	86.57
$A_{neu-neg}$	90.77	90.10	**90.43**	94.00	92.31	**93.16**	88.04	90.42	**89.23**
$L_{pos-neu}$	85.54	86.39	85.96	87.05	86.67	86.86	86.05	87.50	86.78
$L_{pos-neg}$	85.85	89.59	87.72	88.75	88.24	88.50	84.49	88.89	86.89
$L_{neu-neg}$	90.46	90.83	**90.65**	93.99	88.24	**91.12**	87.74	87.88	**87.81**
$P_{pos-neu}$	89.54	89.75	89.64	88.14	88.89	88.52	88.19	88.57	88.38
$P_{pos-neg}$	87.85	92.66	90.25	87.66	92.96	90.26	88.19	86.67	87.43
$P_{neu-neg}$	94.77	94.13	**94.45**	95.09	90.32	**92.71**	92.94	96.55	**94.75**

#:U_{pos}, V_{neu}, W_{neg} are the feature matrix under positive, neutral and negative stimuli, respectively.
$A_{pos-neu}$, $A_{pos-neg}$, $A_{neu-neg}$ represent the fusion matrix obtained by linear combination when $k_1 = k_2 = k_3 = 1$.
$L_{pos-neu}$, $L_{pos-neg}$, $L_{neu-neg}$ represent the fusion matrix obtained by linear combination when $k_1 = 0.86, k_2 = 0.88, k_3 = 0.87$.
$P_{pos-neu}$, $P_{pos-neg}$, $P_{neu-neg}$ represent the fusion matrix obtained by concatenating.

	single	k1=k2=k3=1	k1=0.86,k2=0.88, k3=0.87	concatenate
Accuracy	87.66%	88.76%	88.94%	92.18%
Precision	84.96%	88.97%	87.72%	90.69%
Recall	87.32%	89.18%	88.09%	90.60%

Fig. 4: The average classification performances comparison of all fusing modes and single modes when using SVM as the classifier (The accuracy, precision and recall rate in the figure refer to the average value under single mode and each fusion mode).

1) Linear combination:

When the three modes take the same weights for linear combination, that is $k_1 = k_2 = k_3 = 1$, the feature matrixes are represented as follows:

$$A_{pos-neu} = U_{pos} + V_{neu} \tag{8}$$

$$A_{pos-neg} = U_{pos} + W_{neg} \tag{9}$$

$$A_{neu-neg} = V_{neu} + W_{neg} \tag{10}$$

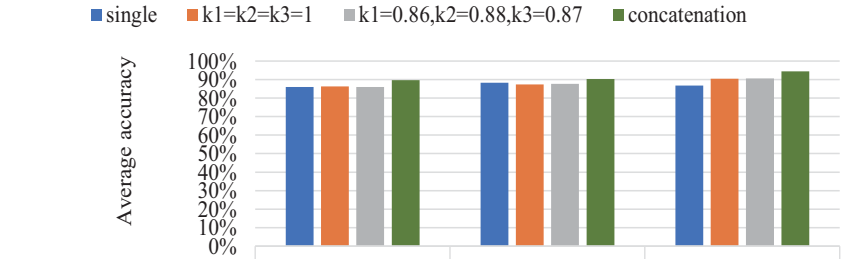
The three modes take different weights. According to the classification results of the single mode, positive: 85.70%, neutral: 88.32%, and negative: 86.78%, we take the weights of the three modes of positive, neutral and negative in order $k_1 = 0.86$, $k_2 = 0.88$, $k_3 = 0.87$. Taking the weights of the three modes as the combination coefficient of the linear combination, the resulting fusion matrix is as follows:

$$L_{pos-neu} = 0.86 \times U_{pos} + 0.88 \times V_{neu} \tag{11}$$

$$L_{pos-neg} = 0.86 \times U_{pos} + 0.87 \times W_{neg} \tag{12}$$

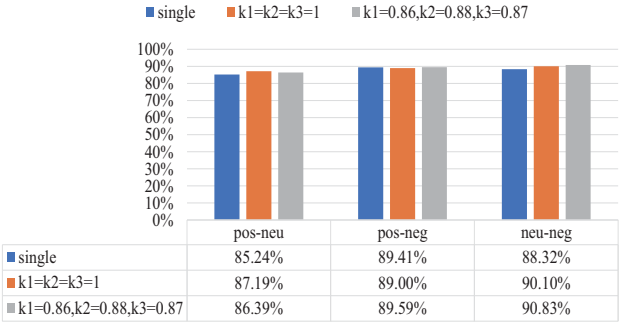
$$L_{neu-neg} = 0.88 \times V_{neu} + 0.87 \times W_{neg} \tag{13}$$

The classification results of linear combination modes are shown in Table I when using accuracy, precision and recall as metrics. We can see that SVM performs better than KNN in



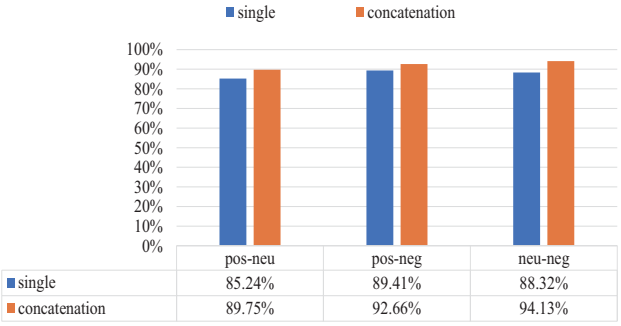
	pos-neu	pos-neg	neu-neg
single	85.96%	88.32%	86.78%
k1=k2=k3=1	86.29%	87.35%	90.43%
k1=0.86,k2=0.88,k3=0.87	85.96%	87.72%	90.65%
concatenation	89.64%	90.25%	94.45%

Fig. 5: Classification accuracy under fusing mode and single mode (single modes order is positive, neutral, negative).



	pos-neu	pos-neg	neu-neg
single	85.24%	89.41%	88.32%
k1=k2=k3=1	87.19%	89.00%	90.10%
k1=0.86,k2=0.88,k3=0.87	86.39%	89.59%	90.83%

Fig. 6: The classification accuracy comparison of linear combination modes and single modes when using SVM as classifier (single modes is positive, neutral, negative).



	pos-neu	pos-neg	neu-neg
single	85.24%	89.41%	88.32%
concatenation	89.75%	92.66%	94.13%

Fig. 7: The classification accuracy comparison of concatenating modes and single modes when using SVM as the classifier (single modes order is positive, neutral, negative).

general in our study. Fig. 6 displays the classification accuracy comparison of two linear combination strategies and single modes when using SVM as the classifier. The results show that compared with the single-mode classification results, the accuracy of depression recognition through feature fusion by linear combination improved slightly, but there is not much difference between the two different weight coefficients.

2) Concatenation:

The concatenating features method in two different modes was used to obtain a new feature matrix and classify depressed patients and healthy people on the obtained fusion features. This method can synthesize the characteristics of different modes, thereby improving the recognition accuracy of the model. However, the dimension of the feature matrix obtained by concatenating is twice the dimension of the feature matrix in a single mode and some redundant information was introduced to our model. As the dimension increased, the computational complexity also increased greatly, so we need to reduce the dimension of the feature matrix. In this article, we used PCA to reduce the dimension of the feature matrix. The classification results of concatenation modes are shown in Table I. The classification accuracy comparison of concatenating modes and single modes when using SVM as the classifier are shown in Fig. 7.

It can be seen from the previous results that SVM performs better than KNN in general. For single modes, the feature matrix extracted under the stimulation of neutral words obtained the best classification results, the average accuracy was 88.32%, the average precision was 86.81%, the average recall was 90.74%. For the fusion modes, whether it was the fusion method of linear combination or the fusion method of concatenation, the best recognition modality in fusion modalities was the fusion of neutral and negative word stimuli. For all modes, the KNN classifier obtained the best depression recognition in the concatenating modality of neutral and negative word stimuli, and the accuracy was 94.77%.

Fig. 4 shows the comparison of classification performance between single modes and fusion modes when SVM as the classifier. Fig. 5 shows the average classification accuracy rate under all fusion modes and single modes. It can be seen from the results of Fig. 4 and Fig. 5 that the classification accuracy, precision and recall obtained by fusing

modes apparently improved compared with the single modes. In addition, the classification result obtained by the fusion method of concatenating was better than the result obtained by linear combination. There was no significant difference in the classification accuracy obtained by the two different linear combination methods.

IV. Conclusion

In this study, we proposed a method of identifying patients with depression among healthy people using feature fusion based on the fNIRS technique. First, we studied the discrimination accuracy of depression in three single modes: positive, neutral and negative. To research depression recognition under fusing features, linear combination and concatenation were used to perform information fusion at the feature level. The three fusion matrixes were pos-neu, pos-neg, and neu-neg. We found that the method of identifying depression by fusing features can improve the recognition accuracy. In addition, the results showed that there was no obvious difference in classification accuracy under two different combination coefficients ($k_1 = k_2 = k_3 = 1$, or $k_1 = 0.86,\ k_2 = 0.88,\ k_3 = 0.87$). We also found that feature fusion by concatenation can provide more accurate diagnostic results for depression detection, compared with the linear combination. On comparing the classification accuracy in different fusion modalities, we found that the best classification result was obtained by concatenating the neutral features and negative features for feature fusion, with the highest accuracy of 94.45%. Different from previous methods for identifying depression with single-mode features, multimode fusion can effectively improve the accuracy of depression recognition. The multimodality fusion method proposed in this paper will provide a more objective and effective method for the diagnosis of depression.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2019YFA0706200), in part by the National Natural Science Foundation of China (Grant No.61632014, No.61627808, No.61210010), in part by the National Basic Research Program of China (973 Program, Grant No.2014CB744600), in part by the Program of Beijing Municipal Science & Technology Commission (Grant No.Z171100000117005), and in part by the Fundamental Research Funds for the Central Universities (lzujbky-2020-kb25).

References

[1] U. Acharya, V. Sudarshan, H. Adeli, J. Santhosh, E. Koh, S. Puthankatti, A. Adeli, "A novel depression diagnosis index using nonlinear features in EEG signals," *European Neurology*, vol. 74, no. 1–2, pp. 79–83, Aug. 2015.

[2] D. Kan, P. Lee, "Decrease alpha waves in depression:an electroencephalogram(EEG) study," *2015 International Conference on Biosignal Analysis, Processing and System(ICBAPS*, pp. 156–161, May. 2015.

[3] C. Liao, Z. Feng, "Mechanism of affective and cognitive-control brain regions in depression," *Advances in Psychological Ence*, vol. 115, no. 3, pp. 1325–1335, 2010.

[4] Y. Ali, N. Hadi, "Quantitative electroencephalography for objective and differential diagnosis of depression: a comprehensive review," *Global Journal of Health Science*, vol. 8, no. 11, pp. 249–256, 2016.

[5] A. O'Neil, K. Sanderson, B. Oldenburg, C. Taylor, "Impact of depression treatment on mental and physical health-related quality of life of cardiac patients: a meta-analysis," *Journal of Cardiopulmonary Rehabilitation and Prevention*, vol. 31, no. 3, pp. 146–156, May. 2011.

[6] J. Shen, X. Zhang, B. Hu, G. Wang, Z. Ding, "An improved empirical mode decomposition of Electroencephalogram signals for depression detection," *IEEE Transactions on Affective Computing*, 2019.

[7] I. Gotlib, J. Hamilton, "Neuroimaging and depression: current status and unresolved issues," *Current Directions in Psychological Science*, vol. 17, no. 2, pp. 159–163, Apr. 2008.

[8] J. Joshi, R. Goecke, S. Alghowinem, A. Dhall, M. Magner, J. Epps, G. Parker, M, Breakspear, "Multimodal assistive technologies for depression diagnosis and monitoring," *Journal on Multimodal User Interfaces*, vol. 7, no. 3, pp. 217–229, Nov. 2013.

[9] H. Cai, J. Han, Y. Chen, X. Sha, Z. Wang, B. Hu, J. Yang, L. Feng, Z. Ding, Y. Chen, J. Gutknecht, "A pervasive approach to EEG-based depression detection," *Complexity*, no. 5, 2018.

[10] J. Mitra, L. Carey, J. Fripp, K. Shen, K. Pannek, K. Bourgeat, O. Salvado B. Campbell, A. Connelyy, S. Palmer, S. Rose, "Predicting poststroke depression from brain connectivity," *Mathematics and Visualization*, vol. 39, pp. 89–99, 2014.

[11] Q. Gong, and Y. He, "Depression, neuroimaging and connectomics: a selective overview," *Biological Psychiatry*, vol. 7, Dec. 2015.

[12] P. Pinti, C. Aichelburg, F. Lind, S. Power, E. Swingler, A. Merla, A. Hamilton, S. Gilbert, P. burgess, L. Tachtsidis, "Using fiberless, wearable fNIRS to monitor brain activity in real-world cognitive tasks," *Journal of Visualized Experiments Jove*, no. 106, 2015.

[13] M. Khan, K. Hong, N. Naseer, M. Bhutta, "A hybrid eeg-fnirs bci: Motor imagery for eeg and mental arithmetic for fnirs," *International Conference on Control Automation and Systems*, pp. 275–278, 2014.

[14] S. Toshikazu, K. Nobutoshi, M. Saori, M. Hanae, O. Yusuke, "Increase of frontal cerebral blood volume during transcranial magnetic stimulation in depression is related to treatment effectiveness: A pilot study with near-infrared spectroscopy," *Psychiatry and Clinical Neurosciences*, 2018.

[15] T. Matsubara, K. Matsuo, M. Nakashima, M. Nakano, K. Harada, T. Watanuki, K. Egashira, Y. Watanabe, "Prefrontal activation in response to emotional words in patients with bipolar disorder and major depressive disorder," *Neuroimage*, vol. 85, pp. 489–497, Jan. 2014.

[16] R. Takizawa, M. Fukuda, S. Kawasaki, K. Kasai, M. Mimura, S. Pu, T. Noda, S. Niwa, Y. Okazaki, "Neuroimaging-aided differential diagnosis of the depressive state," *Neuroimage*, vol. 85, pp. 498–507, Jan. 2014.

[17] H. Song, W. Du, X. Yu, W. Dong, W. Quan, W. Dang, H. Zhang, J. Tian, T. Zhou, "Automatic depression discrimination on fnirs by using general linear model and svm," *International Conference on Biomedical Engineering and Informatics* , 2014.

[18] X. Zhang, B. Hu, J. Shen, Z. Din, J. Liu, G. Wang, "Multimodal depression detection: fusion of electroencephalography and paralinguistic behaviors using a novel strategy for classifier ensemble," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 6, pp. 2265–2275, Nov. 2019.

[19] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, B. Hu, "Feature-level fusion approaches based on multimodal eeg data for depression recognition," *Information Fusion*, vol. 59, pp. 127–138, Jul. 2020.

[20] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Hyett, G. Parker, M. Breakspear, "Multimodal depression detection: Fusion analysis of paralinguistic, head pose and eye gaze behaviors," *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 478–490, Oct. 2018.

[21] F. Scholkmann, S. Spichtig, T. Muehlemann, M. Wolf, "How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation," *Physiological Measurement*, vol. 31, no. 5, pp. 649–662, May. 2010.

[22] Y. Zhang, D. Brooks, M. Franceschini, D. Boas, "Eigenvector-based spatial filtering for reduction of physiological inter- ference in diffuse optical imaging," *J Biomedical Optics*, vol. 10, no. 1, pp. 11014, 2005.

[23] B. Molavi, G. Dumont, "Wavelet-based motion artifact removal for functional near-infrared spectroscopy," *Physiological Measurement*, vol. 33, no. 2, pp. 259-270, Feb. 2012.

[24] X. Cui, S. Bray, A. Reiss, "Functional near infrared spectroscopy (nirs) signal improvement based on negative correlation between oxygenated

and deoxygenated hemoglobin dynamics," *Neuroimage*, vol. 49, no. 4, pp. 3039-3046, 2010.

[25] M. Izzetoglu, P. Chitrapu, S. Bunce, B. Onaral, "Motion artifact cancellation in nir spectroscopy using discrete kalman filtering," *Biomedical Engineering Online*, vol. 9, no. 1, pp. 1–10, 2010.

[26] H. Klumpp, K. Kinney, R. Bhaumik, J. Fitzgerald, "Principal component analysis and brain-based predictors of emotion regulation in anxiety and depression," *Psychological Medicine*, vol. 49, no. 14, pp. 2320–2329, Oct. 2019.

[27] W. Wang, J. Zhen, "An efficient nearest neighbor classifier algorithm based on pre-classify," *Computer Ence*, vol. 34, no. 2, pp. 198–200, 2007.

[28] P. Bhusana, V. Jumlong, S. Nop, T. Christofer, "A novel multiclass support vector machine algorithm using mean reversion and coeeficient of variance," *Mathematics and Statistics*, vol. 9, no. 3, pp. 208–218, 2013.

[29] T. Matsubara, K. Koji Matsuo, M. Nakashima, M. Nakano, "Prefrontal activation in response to emotional words in patient with bipolar disorder and major depressive disorder," *NeuroImage*, vol. 85, pp. 489-497, 2014.