IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 27, 2021, accepted March 16, 2021, date of publication March 24, 2021, date of current version April 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3068316

Effects of Data Augmentation Method Borderline-SMOTE on Emotion Recognition of EEG Signals Based on Convolutional Neural Network

YU CHEN, RUI CHANG, AND JIFENG GUO

College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China

Corresponding author: Jifeng Guo (guojifeng@nefu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61300098, in part by the Natural Science Foundation of Heilongjiang Province under Grant F201347, and in part by the Fundamental Research Funds for the Central Universities under Grant 2572015DY07.

ABSTRACT In recent years, with the continuous development of artificial intelligence and brain-computer interface technology, emotion recognition based on physiological signals, especially electroencephalogram signals, has become a popular research topic and attracted wide attention. However, the imbalance of the data sets themselves, affective features' extraction from electroencephalogram signals, and the design of classifiers with excellent performance, pose a great challenge to the subject. Motivated by the outstanding performance of deep learning approaches in pattern recognition tasks, we propose a method based on convolutional neural network with data augmentation method Borderline-synthetic minority oversampling technique. First, we obtain 32-channel electroencephalogram signals from DEAP data set, which is the standard data set of emotion recognition. Then, after data pre-processing, we extract features in frequency domain and data augmentation based on the data augmentation algorithm above for getting more balanced data. Finally, we train a one dimensional convolutional neural network for three classification on two emotional dimensions valence and arousal. Meanwhile, the proposed method is compared with some traditional machine learning methods and some existing methods by other researchers, which is proved to be effective in emotion recognition, and the average accuracy rate of 32 subjects on valence and arousal are 97.47% and 97.76% respectively. Compared with other existing methods, the performance of the proposed method with data augmentation algorithm Borderline-SMOTE shows its advantage in affective emotional recognition than that without Borderline-SMOTE.

INDEX TERMS Electroencephalogram, emotion recognition, Borderline-Synthetic minority oversampling technique, convolutional neural network.

I. INTRODUCTION

Emotion is a group of continuous or discontinuous reaction of people to their own and things of outside world, which is indispensable to humans [1]–[3]. It is the carrier of non-verbal communication between people, and it plays an important role in people's social activities. However, due to the acceleration of social development and the pace of people's daily life, many people tend to feel stressed and anxious. If this condition continues this way, it may lead to various health issues or depression, even influence people's daily life and self-development. Therefore, emotion recognition gradually becomes a hot and realistic research topic which has received extensive attention from researchers [4], [5].

At present, emotion recognition is applied to many fields, meanwhile, many effective methods are pushing this subject forward. For instance, we can determine the emotion state of the information host when writing an article or recording a speech through text or speech [6]–[8], or recognize facial expressions and gestures of people through video equipment to judge their current emotions [9]–[11]. However, these fields, to some extent, are not be reliable enough, mainly

The associate editor coordinating the review of this manuscript and approving it for publication was Francesco Mercaldo.

because emotion can be behaved affectedly by others, then the authenticity of emotion cannot be well-guaranteed. Just as Picard *et al.* [12] said, if someone has the ability to disguise his or her emotion, the result of emotion recognition may have a high error rate.

As opposed to this, some researchers focus on the physiological signals,especially electroencephalogram(EEG) signals for emotion recognition, considering EEG signals have strong objectivity: they cannot be forged easily and are expressed from the inside out, which makes the result more real and more reliable [13]. Meanwhile, with the continuous development of artificial intelligence and brain-computer interface technology [14], [15], the application prospect of emotion recognition based on EEG signals is broad, and it has gradually become the mainstream of emotion recognition [16], [17].

Inspired by this, we use EEG signals as the data sources to carry out research on emotion recognition. The overall idea of this research can be summarized as follows: data pre-processing, feature extraction, classification and evaluation of the performance of model [17], in order to design an emotion recognition model with excellent generalization performance. Nowadays, more and more researchers at home and abroad have begun to pay attention to the field of emotion recognition of EEG signals, and propose many effective methods for this research.

Some researchers focus on traditional machine learning methods. Many kinds of well-designed features are extracted, then recognized by effective classifiers, which promotes the progress of emotion recognition of EEG signals. Islam and Ahmad [18] proposed a system of emotion recognition from EEG signals based on discrete wavelet transform. The extracted features wavelet energy and wavelet entropy were trained in the knearest neighbor(KNN) algorithm for classification. On DEAP dataset, the average accuracy was about 63%. Wichakam and Vateekul [19] extracted bandpower and power spectral density(PSD) from EEG signals, and used support vector machine(SVM) classifier for binary classification. The best accuracy of valence, arousal and liking was 64.90%, 64.90%, and 66.80% respectively. Yoon and Chung [20] proposed a supervised learning algorithm based on the Bayesian weighted logarithmic posterior function and perceptron convergence algorithm. For binary classification task, it achieved accuracy of 70.90% on valance and 70.10% on arousal, while for three classification task, it achieved accuracy of 53.40% on valance and 53.10% on arousal. Parui *et al.* [21] extracted multiple features from brain electric signals, and then the correlation matrix information gain was used to calculate and recursive feature elimination methods optimized the feature set, which was used to classify the XGBoost classifier. With testing in the DEAP data set, the four emotional dimensions with the best accuracy were 75.97%, 74.20%, 75.23%, and 76.42%.

In recent years, with the improvement of computer automation and the advancement of pattern recognition task, some deep learning methods have been gradually developed, and used in the field of EEG-based emotion recognition [22], [23]. Even up to a point, this kind of method outperforms traditional machine learning methods. Xing *et al.* [24] established a stacked autoencoder(SAE) to decompose EEG signals and classified them by the long-short-term memory(LSTM) model. The observed accuracy of valence was 81.10% and that of arousal was 74.38%. Alhagry *et al.* [25] proposed a deep learning approach to identify emotions from raw EEG signals, using LSTM neural networks to learn features from EEG signals, and then categorized these features into low/high arousal, valence and liking. The method was tested on the DEAP data set. The average accuracy of the method was 85.45%, 85.65%, and 87.99% respectively, on arousal, valence and liking, respectively. Zhan *et al.* [26] extracted PSD of four frequency bands from EEG signals and designed a shallow depthwise parallel convolutional neural network(CNN). This approach achieved the competitive accuracy of 82.95% and 84.07% on valence and arousal respectively in DEAP data set. Meanwhile, the method shows extensive application prospects for EEG-based emotion recognition on resource-limited devices. Yang *et al.* [27] implemented a 2DCNN-LSTM module to extract spatial and temporal features respectively and combined the features for binary classification, which achieved high accuracy of 91.03% and 90.80% on arousal and valence in the emotion recognition task.

However, to some extent, these methods above ignore an important issue more or less: the handling of data imbalance. Especially in the multi-classification task, the number of samples in different classes of the original data set may vary greatly, which will directly affect the final model performance [28], [29]. Therefore, considering the problems above, and the application of deep learning approaches in various pattern recognition tasks, we propose a method based on CNN with data augmentation method Borderline-Synthetic minority oversampling technique(Borderline-SMOTE). The specific contributions of this paper can be summarized as follows:

(1) We extract features in frequency domain based on the Short-Time Fourier Transform(STFT) method and sliding window, to obtain enough feature set related to emotion.

(2) We introduce data augmentation method Borderline-SMOTE, for getting more balanced feature set of EEG signals.

(3) We train a 1DCNN model for three classification on two emotional dimensions valence and arousal, and make comparison with some existing traditional machine learning methods(Decision Tree, KNN, SVM, XGBoost), to verify that the proposed model has a higher performance on emotion recognition of EEG signals.

The remainder of this paper is organized as follows: In section II, we elaborate the methodology of the proposed method, including the overview of DEAP data set(used for evaluating the performance of the model), and the details of pre-processing, feature extraction, data augmentation and classification; In Section III, we present results under the three classification task of every subject without and with data augmentation on emotion dimensions valence and arousal with using DEAP data set, evaluate the proposed model and make some discussion about the proposed method with some existing traditional machine learning methods(Decision Tree, KNN, SVM, XGBoost). In Section IV, we conclude our work and plan the following research objectives.

II. METHODOLOGY

The technical route of the proposed method in this paper is shown in Fig. 1. More details of each step are as follows.

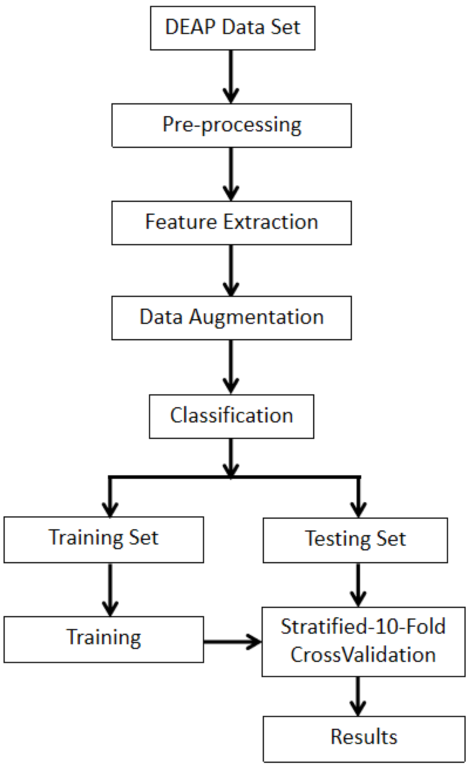

FIGURE 1. The technical route of the proposed method in this paper.

A. DEAP DATA SET - EMOTION RECOGNITION STANDARD DATA SET

In this paper, we use the standard data set of emotion recognition - DEAP data set, which is an open data set for emotion recognition with using physiological signals [30]. DEAP data set is a multimodal data set that contains EEG and other peripheral physiological signals. It includes not only physiological signals, but also facial expression videos and subjective evaluation of subjects during physiological signals acquisition experiment. 32 subjects were stimulated by Music Video to induce their emotions, and then physiological signals of the subjects were collected to obtain the data set.

After down-sampling, the sampling frequency of EEG signals in the DEAP data set were reduced to 128Hz, and then the EEG signals were filtered to 4 ∼ 45Hz. After averaging to the unified benchmark, the eye electrical artifacts were removed by blind source separation method. The signal

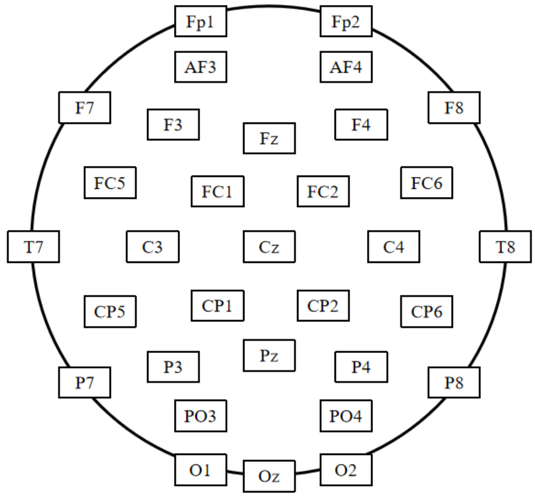

FIGURE 2. Name and corresponding location of the 32-conductor EEG signal electrodes in the brain region.

TABLE 1. DEAP data set signal categories and corresponding signal acquisition channel.

Signal Category	Signal acquisition channel
EEG Signals	Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, O2
EOG Signals	Horizontal Ophthalmic Signal, Vertical Ophthalmic Signal
EMG Signals	Zygomatic Major Signal, Trapezius Signal
Skin Electrical Signals	Galvanic Skin Response
Respiratory Band Signals	Respiratory Band
Body Surface Scanning Signals	Volume Scanning

categories and corresponding signal acquisition channels are shown in Table. 1, and 32-conductor EEG signal electrodes in the brain region are shown in Fig. 2.

Each data file(s01.dat ∼ s32.dat) contains the following two matrices:

(1) Data matrix: 40 * 40 * 8064, the first 40 represents the total number of videos, the second 40 represents the signal collection of the total number of channels, 8064 is 63 seconds of any 1 video 1 channel (63 * 128) on the experimental data, which is the data for 3 seconds before the baseline data obtained before the experiment, and the last 60 seconds of data recorded in the process of the experiment;

(2) Labels matrix: 40 * 4. These four columns represent four affective dimensions: valence, arousal, dominance, and liking scores ranging from 1.0 to 9.0.

B. PRE-PROCESSING

The model of EEG-based emotion recognition is targeted to be highly personalized and is thus trained and tested in units of each subject. First, considering the first 3 seconds of EEG signals is baseline, it is discarded and the last 60s EEG data are selected as the research object of the proposed method.

TABLE 2. The original data distribution of EEG signals on valence.

Subject	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
01	1	2	6	3	4	2	2	1	1	0	0	2	8	1	4	3
02	6	1	1	1	0	0	2	4	5	0	3	0	2	0	2	13
03	0	0	0	0	3	2	5	8	5	1	2	3	7	3	1	0
04	3	7	4	4	2	0	3	1	2	1	4	3	4	0	1	1
05	3	2	0	1	2	1	3	4	2	2	1	6	5	1	5	2
06	0	0	0	0	2	3	2	3	4	2	3	6	6	6	3	0
07	0	1	2	1	5	0	1	2	3	0	1	5	14	1	4	0
08	3	0	0	1	1	3	6	4	7	0	1	4	2	2	4	2
09	0	0	0	3	5	4	5	2	8	2	4	0	6	0	1	0
10	1	2	2	4	0	4	3	4	2	6	0	2	4	3	0	3
11	4	1	2	1	1	2	2	3	5	3	5	2	4	3	2	0
12	2	4	1	4	4	0	2	2	4	0	2	5	6	3	1	0
13	4	2	1	4	1	1	7	2	5	1	1	0	4	4	0	3
14	2	2	4	4	4	3	1	0	0	1	2	4	5	4	4	0
15	0	1	2	2	5	2	6	2	3	0	1	0	4	1	5	6
16	4	1	1	6	1	7	3	2	3	2	4	4	2	0	0	0
17	0	0	0	0	4	5	5	4	1	5	8	2	6	0	0	0
18	0	0	0	1	1	2	6	4	5	4	7	5	4	1	0	0
19	0	1	2	2	5	2	2	3	4	3	4	3	4	5	0	0
20	0	0	1	2	4	0	5	5	6	1	3	2	3	5	2	1
21	0	0	2	1	5	2	5	4	2	2	3	2	6	2	2	2
22	4	5	2	4	0	0	3	2	5	0	2	0	8	0	3	2
23	0	1	0	2	0	3	2	5	8	0	0	3	7	2	5	2
24	2	1	2	2	4	1	9	0	3	1	4	0	7	0	2	2
25	8	1	1	1	4	0	1	3	6	0	3	0	4	1	6	1
26	4	3	1	2	2	0	1	1	2	1	5	2	4	1	7	4
27	2	0	1	2	0	0	2	3	8	0	3	1	7	0	4	7
28	3	3	4	1	2	0	2	0	3	1	2	3	7	2	6	1
29	2	3	1	2	4	2	3	0	3	0	6	0	9	0	4	1
30	0	0	0	1	2	3	6	1	12	1	5	0	7	0	2	0
31	5	1	0	0	1	1	3	3	8	0	7	3	3	1	3	1
32	2	0	0	1	2	5	8	2	2	1	4	1	5	2	5	0

In addition, the EEG data of subjects has the inherent problem of large difference in characteristic values in and between subjects, normalization of EEG data is an important part. In this paper, min-max normalization [31] is adopted, and its formula is as follows.

$$x^{*}(t) = \frac{x(t) - \min(x)}{\max(x) - \min(x)} \tag{1}$$

where $x(t)$ is the raw EEG data(essentially a time series), $min(x)$ and $max(x)$ is the minimum and maximum value of $x(t)$ respectively, $x^{*}(t)$ is the normalized EEG data. Through the above process, the raw EEG signal value is mapped to the value from 0.0 to 1.0.

Then, we classify EEG signals into three classes(positive, neutral, negative) on two emotion dimensions(valence and arousal). The original data distribution of EEG signals on valence and arousal are shown in Table. 2 and Table. 3 respectively. The label on the top row represents the left end of the score interval. From the two tables, we can find that the scores of labels are real numbers from 1.0 to 9.0, but the distribution of labels is relatively random and irregular. Therefore, to balance the distribution in each class as far as possible, 4.0 and 6.0 are selected as the classification boundary threshold. In other words, we re-label the value of every class in these two emotion dimensions: the label of positive class is "2"(score from 6.0 to 9.0), the medium class is "1"(score from 4.0 to below 6.0), and the negative class is "0"(score from 1.0 to below 4.0).

C. FEATURE EXTRACTION

Feature represents the properties or characteristics of the signal. By feature extraction, a set of features that contains relevant information as much as possible is extracted from the raw data, which is significant for improving the performance of the model [32], [6].

Considering the different manifestations of EEG signals in different frequency bands and their non-stationarity, based on the STFT method, spectral power features of five frequency

TABLE 3. The original data distribution of EEG signals on arousal.

Subject	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
01	0	0	4	3	4	4	0	1	0	0	2	4	9	4	5	0
02	8	1	1	1	2	0	0	3	3	0	2	1	5	0	5	8
03	3	2	2	5	8	4	7	1	2	2	3	0	0	0	1	0
04	1	3	4	8	2	3	1	2	5	1	4	2	2	0	2	0
05	0	2	0	3	2	2	6	6	3	1	3	3	3	1	2	3
06	0	0	0	2	7	6	4	4	3	6	2	3	2	0	1	0
07	0	0	3	2	5	0	5	0	3	3	4	1	12	0	1	1
08	0	0	0	1	3	3	2	7	6	1	4	7	2	2	2	0
09	0	0	0	0	0	1	10	4	5	2	12	1	5	0	0	0
10	0	1	1	5	2	3	5	1	5	5	9	0	2	1	0	0
11	9	2	1	6	1	2	1	3	3	0	5	2	2	0	2	1
12	0	0	0	1	4	0	0	2	7	1	1	7	9	2	2	4
13	1	1	0	0	1	0	2	1	2	3	5	2	7	2	4	9
14	0	0	2	2	7	0	2	0	0	8	7	6	3	2	1	0
15	0	0	1	3	4	2	8	1	11	1	3	2	4	0	0	0
16	2	0	2	4	2	8	2	0	3	9	3	0	0	3	2	0
17	0	0	0	3	1	6	4	1	1	12	6	1	4	0	1	0
18	0	0	0	0	3	2	8	2	3	4	7	7	2	2	0	0
19	0	1	2	2	2	4	1	1	2	7	6	5	3	3	0	1
20	0	0	0	1	3	1	1	3	10	7	6	2	5	0	1	0
21	0	0	0	3	1	1	3	0	1	3	12	5	10	0	1	0
22	2	3	1	0	2	0	3	4	3	2	8	0	11	0	1	0
23	17	4	1	1	0	0	1	3	5	0	1	2	3	1	0	1
24	0	0	1	1	2	0	1	2	6	4	11	2	4	2	4	0
25	2	1	2	1	2	0	1	1	4	1	10	0	9	1	4	1
26	5	4	1	5	2	1	2	3	5	0	0	2	5	1	4	0
27	6	0	3	0	3	0	0	1	11	0	3	0	7	1	3	2
28	4	3	4	2	6	0	1	0	2	0	2	1	9	2	3	1
29	5	2	2	2	0	1	1	0	5	0	4	0	10	1	5	2
30	1	0	0	3	2	4	5	5	6	2	7	2	3	0	0	0
31	2	3	3	3	2	3	3	1	4	1	2	1	3	1	4	4
32	1	1	0	1	1	3	4	1	3	4	9	1	11	0	0	0

bands are extracted from the normalized EEG signals. First, we use bandpass IIR filters to obtain the five frequency band signals including theta(4-7 Hz), alpha(8-13 Hz), low-beta(13-16 Hz), high-beta(16-31 Hz) and gamma (31-45 Hz) in each channel. Then, we use a 2-second(i.e. the size of the sliding window is 256) sliding window with 1.835-second overlapping(i.e. the step size of the sliding window is 16), to capture more detailed feature information. Finally, we get $40 * 465 = 18600$ samples, where 40 represents the total number of videos, 465 represents the total number of segments that produced by the 2-second sliding window, which can be calculated by the formula as follows:

$$samples = 1 + \frac{total\_length - sliding\_window\_length}{sliding\_window\_length - overlap\_length} \tag{2}$$

For every sample, we extracted the characteristics of 5 different frequency bands from 32 EEG channels, thus, for every subject, the shape of data and labels are (18600, 160) and (18600, 4) respectively.

D. DATA AUGMENTATION

After the above steps of feature extraction, sufficient samples are obtained. Now, the data distribution of each class is shown in Table 4.

However, by analyzing the distributions of samples, we can find that these samples are imbalanced in different classes. Even in the worst case, the number of samples of majority class is 10 times or more than that of other minority class. Moreover, some samples are close to or at the classification boundary. These factors work together, increase the difficulty of the classification task and in turn influence the performance of model. Therefore, data augmentation should be a key factor that has to be considered.

Motivated by the typical applications of data augmentation algorithm Borderline-SMOTE in arrhythmia detection [33], legal judgment prediction [34], etc, in this paper, we use the Borderline-SMOTE algorithm for data augmentation. Borderline-SMOTE is an oversampling algorithm derived from SMOTE algorithm. SMOTE algorithm generates

TABLE 4. Data distribution of each class after feature extraction.

Subject	Valence			Arousal		
	Positive	Medium	Negative	Positive	Medium	Negative
01	8370	1860	8370	11160	465	6975
02	8835	5580	4185	9765	2790	6045
03	7440	8835	2325	1860	5580	11160
04	5580	3720	9300	4650	4185	9765
05	9300	4185	5115	6510	7440	4650
06	11160	5115	2325	3720	7905	6975
07	11625	2790	4185	8835	4650	5115
08	6975	7905	3720	7905	7440	3255
09	5115	7905	5580	8370	8835	1395
10	5580	6975	6045	4650	8370	5580
11	6975	6510	5115	5580	3265	9765
12	7905	3720	6975	11625	4650	2325
13	5580	6975	6045	13020	3720	1860
14	8835	930	8835	8835	4650	5115
15	7440	5115	6045	4185	9765	4650
16	4650	4650	9300	3720	6510	8370
17	7440	6510	4650	5580	7905	5115
18	7905	8370	2325	8370	7905	2325
19	7440	5580	5580	8370	5115	5115
20	7440	7905	3255	6510	9765	2325
21	7905	6045	4650	12555	3720	2325
22	6975	4185	7440	9300	5580	3720
23	8835	6975	2790	3720	4185	10695
24	6510	5115	6975	10230	6510	1860
25	6975	4650	6975	11160	3720	3720
26	10230	2325	6045	5580	4650	8370
27	10230	5580	2970	7440	5580	5580
28	9765	2790	6045	8370	1395	8835
29	9300	2790	6210	10230	2790	5580
30	6510	9300	2790	5580	7905	5115
31	8370	6510	3720	6975	3720	7905
32	7440	6510	4650	9300	6045	3255

synthetic data of a minority class by using the nearest neighbor of minority class data samples. But it does not consider the location of adjacent majority class data while synthesize the data of minority class, so the class samples can be overlapped [35], [36]. To address this limitation, we pay attention to and only over-sample or strengthen the borderline and its nearby points of the minority class.

Suppose the minority class has been labeled with red '+' (shown in Fig. 3(a)). First, we extract the m nearest neighbors from the samples of the minority class. Then, more than half of the m construct the set by selecting the samples of minority class corresponding to the majority class. This set is called *DANGER* which is denoted by solid squares(shown in Fig. 3(b)). Next, we select the s nearest neighbors of the samples in the *DANGER* set, multiply the distance between the sample and the nearest neighbor by a random number between 0 and 1 and add them to the sample. Finally, the synthetic samples, which are shown in Fig. 3(c) with hollow squares, are generated:

$$synthetic_j = p_j + r_j^* dif_j, \quad j = 1, 2, \ldots, s \tag{3}$$

where p_j represents the instance in the *DANGER* set, r_j represents a random number $\in [0, 1]$, dif_j represents the distance between the s nearest neighbors of samples.

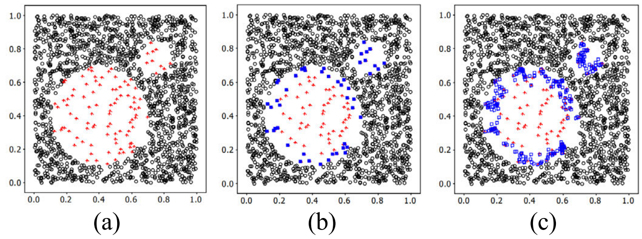

FIGURE 3. (a) The original distribution of data set. (b) The borderline minority samples(solid squares). (c) The borderline synthetic minority samples(hollow squares).

After data augmentation, for each subject, the number of samples of each class is close to the number of samples of the majority class. Therefore, the problem of data imbalance can be solved, and then facilitate the process of classification.

E. CLASSIFICATION

1) THE CNN ARCHITECTURES

Based on the fact that EEG signal is a one dimensional time series [37], we train a one dimensional convolutional neural network(1DCNN) for three classification. The 1DCNN model contains an input layer(L_0), an output layer(L_{13}), and multiple hidden layers($L_1 \sim L_{12}$), which contain many parts such as convolutional layer, batch normalization layer, pooling layer, flatten layer and fully-connected layer. The structure of the 1DCNN model is shown in Fig. 4.

L_0, L_3, L_6: The convolutional layers. The filters are set to 128, 128 and 64 respectively, the kernel_size is set to 3, and the padding is set to "same", to feed the features into the deep levels of network and carry out the convolution operation.

L_1, L_4: The batch normalization layers. By setting after convolutional layers, they can help to solve the problem of difficulty in convergence in the process of network training, and delay the occurrence of over-fitting to a certain extent.

L_2, L_5, L_7: The pooling layers. The pool size is set to 2. They are sandwiched between the prior convolutional layers, and they use the maximum value from each of a cluster of neurons at the prior layers. This will speed up the computation of networks.

L_8: The flatten layer. In the transition from the convolutional layers to the fully-connected layers, the two-dimensional inputs are transferred into one-dimensional inputs, which facilitate subsequent processing of the fully-connected layers.

L_9 - L_{12}: The fully-connected layers, which are fully connected to the flatten layer. Three dense layers are cross-connected to two dropout layers with the units are set to 32, 16 respectively, the two activation function are set to 'tanh' and 'relu' respectively, and the dropout rate is set to 0.4 to prevent overfitting effectively.

L_{13}: The output layer, with 3 units that represents the three classes positive(2), medium(1), and negative(0). The activation is set to softmax that represents the classification task is three-classification.

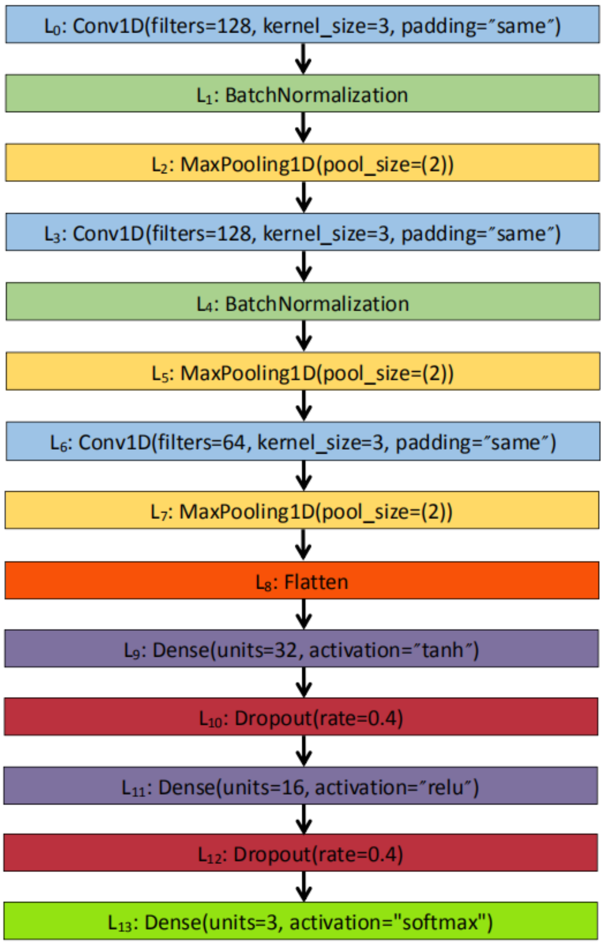

FIGURE 4. The structure of the 1DCNN model.

2) MODEL TRAINING, VALIDATING AND TESTING

The model of the proposed method is implemented by using *Keras* framework [38]. The learning rate is set to 1e-4(0.0001) with the optimizer Adam, the batch size for model training and testing is set to 128, and the epoch of model is set to 30.

All the data of 32 subjects before and after processing data augmentation algorithm Borderline-SMOTE are used for the 1DCNN model's training and testing. First, we randomly and independently split the data of every subject into 80% training set and 20% testing set. Then, the training set is sent to the 1DCNN for model training and the testing set for testing. During the training process, we use stratified-10-fold cross-validation. In every fold, for training set, we select 90% data for training and last 10% data for validation, and evaluate the performance of the model. "Stratification" means that the proportion relationship of each class in the original data is maintained in each fold. For example, there are three classes of the original data with a ratio of 1: 1.2: 1. If the stratified-10-fold cross-validation is adopted, then in the divided 10 folds, and the classes of data in each fold are maintained with a ratio of 1: 1.2: 1. By the kind of cross-valudation method, not only the number of samples of different classes reaches the balance situation, so does the number of every fold, which makes the results more credible than normal k-fold cross-validation [39].

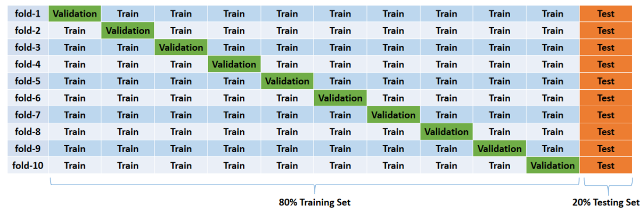

FIGURE 5. The schematic diagram of stratified-10-fold cross-validation.

The schematic diagram of the process is shown in Fig. 5.

3) MODEL EVALUATION

Under the three classification task, the following performance indicators are investigated: Accuracy, Macro-F1-Score, and Confusion matrix. Meanwhile, we draw the figures of results in order to evaluate the performance of the model from multiple angles.

Confusion matrix: There are many different combinations of predicted results and real results, and the corresponding matrix of these combination results is the confusion matrix. In binary classification task, there are 2 * 2 = 4 different combinations of the above results, which are denoted as TP(True Positive), FP(False Positive), FN(False Negative) and TN(True Negative). Thus, the confusion matrix as shown in Table 5 can be obtained.

TABLE 5. Confusion matrix.

True Labels	Predicted Labels	
	Negative(0)	Positive(1)
Negative(0)	TN	FP
Positive(1)	FN	TP

Extend to the three classification task in the proposed method. First, we calculate *TP*, *FP*, *FN*, *TN* of three classes respectively. Then, we can get the average value of *TP*, *FP*, *FN*, *TN* (denoted as $\overline{TP}$, $\overline{FP}$, $\overline{FN}$, $\overline{TN}$, respectively). Thus, like the performance indicators in binary classification task. Accuracy, Macro-Precision, Macro-Recall and Macro-F1-Score can be calculated as follows:

$$Accuracy = \frac{\overline{TP}+\overline{TN}}{\overline{TP}+\overline{TN}+\overline{FP}+\overline{FN}} \tag{4}$$

$$Macro-\Pr ecision = \sum_{i=1}^{3} \frac{TP_i}{TP_i+FP_i} \tag{5}$$

$$Macro-\mathrm{Re}\,call = \sum_{i=1}^{3} \frac{TP_i}{TP_i+FN_i} \tag{6}$$

$$Macro-F1-Score = \frac{2^*(Macro-\Pr ecision)^*(Macro-\mathrm{Re}\,call)}{(Macro-\Pr ecision)+(Macro-\mathrm{Re}\,call)} \tag{7}$$

The higher the performance indicators are, the better the performance of the model.

III. RESULTS AND DISCUSSION

Based on the above analysis, we carry out the main experiment(With using Borderline-SMOTE) and comparison experiment(Without using Borderline-SMOTE) of every subject on two emotional dimensions valence and arousal with using DEAP data set. Moreover, we compare our proposed method with some existing traditional machine learning methods(Decision Tree, KNN, SVM, XGBoost), to further verify the effect of data augmentation method Borderline-SMOTE on the performance of the model. The final results above are shown in Table. 6 and Fig. 6 ∼ Fig. 14.

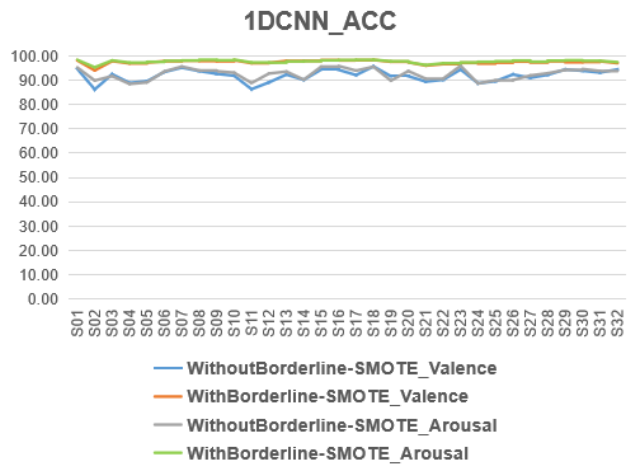

FIGURE 6. The accuracy line diagram of 1DCNN.

A. RESULTS OF THE PROPOSED METHOD

The accuracy line diagram of the proposed method is shown in Fig. 6.

In the figure above, we show the results of 1DCNN on two emotional dimensions valence and arousal, without and with using data augmentation method Borderline-SMOTE. We can find that the overall performance of the model is ideal. First, the performance on arousal is slightly better than that on valence. In accuracy, there is an improvement of about 1% ∼ 2%. Second, in different subjects, the performance of the model fluctuates to some extent, but the overall variance is small, which indicates that the model has a good generalization ability and can adapt to the situation of different sample data, so as to automatically learn from the features and carry out effective classification. Third, as the most important point, by using data augmentation method Borderline-Smote, the performance of the model on both dimensions valence and arousal are significantly improved. In accuracy, there is an average improvement of about 5%. Therefore, the problem of data imbalance can be addressed, so that the data distribution is no longer dominated by the majority of samples, then the model becomes more effective and the results become more convincing.

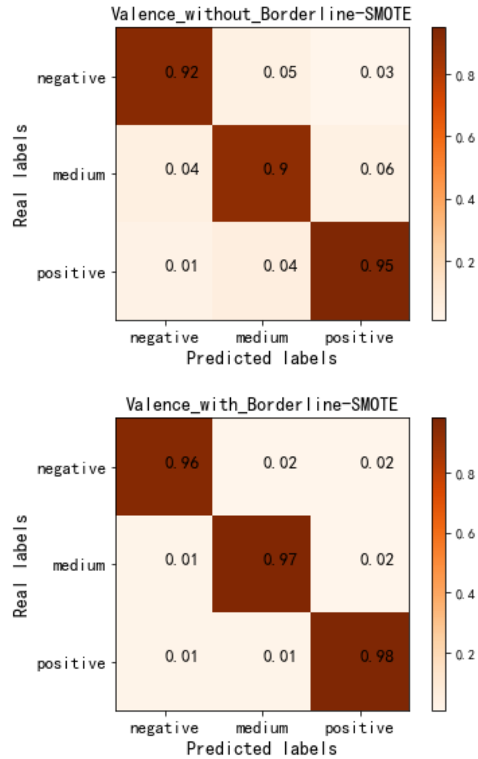

FIGURE 7. The confusion matrices of 1DCNN on valence.

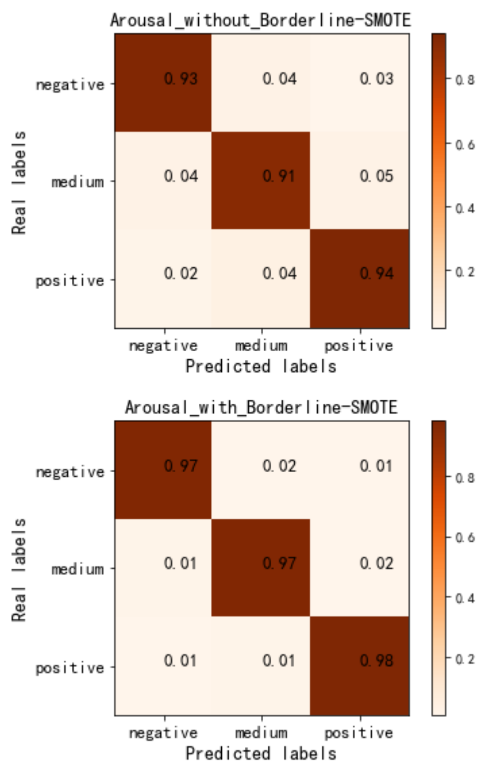

FIGURE 8. The confusion matrices of 1DCNN on arousal.

The confusion matrices of the proposed method are shown in Fig 7 ∼ Fig 8. The rows of the matrices represent real labels(from top to bottom are negative, medium, positive)

TABLE 6. Results of 1DCNN model in every subject, on valence and arousal, without and with Borderline-SMOTE.

Subject	Valence_ Accuracy (without Borderline-SMOTE)	Valence_ Macro-F1-Score (without Borderline-SMOTE)	Valence_ Accuracy (with Borderline-SMOTE)	Valence_ Macro-F1-Score (with Borderline-SMOTE)	Arousal_ Accuracy (without Borderline-SMOTE)	Arousal_ Macro-F1-Score (without Borderline-SMOTE)	Arousal_ Accuracy (with Borderline-SMOTE)	Arousal_ Macro-F1-Score (with Borderline-SMOTE)
S01	94.71	94.89	98.18	98.01	95.25	95.23	98.37	98.46
S02	86.22	87.43	94.05	95.47	89.89	90.07	95.33	96.44
S03	92.48	92.26	97.89	97.95	91.70	91.75	98.15	98.23
S04	88.93	89.09	96.94	96.80	88.48	88.83	97.26	97.39
S05	89.62	90.05	97.25	97.17	89.13	89.34	97.57	97.56
S06	93.41	93.66	97.83	97.88	93.63	94.01	97.98	97.90
S07	95.20	95.51	98.06	98.25	95.64	95.85	98.14	98.33
S08	93.79	93.44	97.91	97.77	94.02	93.96	98.43	98.41
S09	92.66	91.89	97.80	97.93	93.71	93.69	98.22	98.28
S10	91.87	92.55	98.25	98.39	93.13	93.24	98.50	98.56
S11	86.33	87.78	97.04	97.12	88.87	89.13	97.32	97.43
S12	89.12	88.66	97.17	97.07	92.76	92.86	97.26	97.15
S13	92.34	92.90	97.95	98.05	93.57	93.44	97.78	97.87
S14	90.17	90.53	97.99	98.16	90.25	90.68	97.86	97.93
S15	94.42	94.77	98.26	98.35	95.61	95.99	98.21	98.09
S16	94.28	94.55	98.17	97.99	95.86	96.22	98.32	98.44
S17	92.14	92.56	98.24	98.12	93.97	94.19	98.39	98.41
S18	95.68	95.55	98.36	98.20	95.55	95.47	98.43	98.46
S19	91.65	91.97	97.63	97.57	89.84	90.36	97.81	97.88
S20	91.68	92.13	97.41	97.16	93.78	93.99	97.44	97.36
S21	89.44	88.79	96.03	96.43	90.61	91.13	96.26	96.81
S22	90.19	89.97	96.66	96.79	90.68	90.64	96.99	97.05
S23	94.42	94.78	97.15	97.53	95.96	95.89	97.33	97.26
S24	88.75	88.58	96.87	97.18	88.68	88.43	97.58	97.77
S25	89.61	89.44	97.23	97.62	89.99	90.54	97.82	97.79
S26	92.38	91.21	97.69	97.88	90.07	90.80	98.11	98.05
S27	91.01	91.76	97.30	97.36	91.96	92.32	97.62	97.70
S28	92.23	92.66	97.81	97.61	92.88	93.09	98.07	98.16
S29	94.45	94.30	97.41	97.25	94.24	94.46	98.22	98.25
S30	93.90	93.95	97.53	97.38	94.62	94.53	98.01	98.13
S31	93.17	93.46	97.82	97.88	93.79	93.86	97.98	98.05
S32	94.48	94.14	97.04	96.96	94.23	93.99	97.45	97.39
Overall	91.90	92.04	97.47	97.54	92.57	92.75	97.76	97.84

while the columns represent predicted labels(from left to right are negative, medium, positive). The values in the matrices' cells represent the average accuracy of the three emotions on the test set data of all 32 subjects. Compared with the method without Borderline-SMOTE, the average accuracy that predicted correctly of all three kinds emotions have been improved, which also vividly indicates that when facing multi-classification tasks with imbalanced classes, data augmentation methods are a kind of significant tool to deal with different data samples, so that enhance the performance of model.

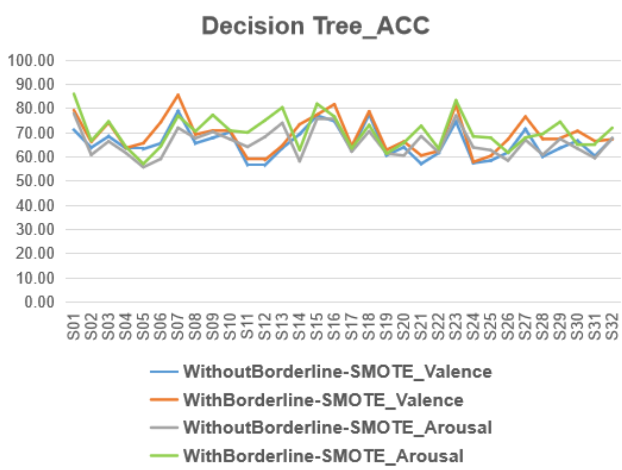

FIGURE 9. The accuracy line diagram of decision tree.

B. COMPARISON BETWEEN THE PROPOSED METHOD AND SOME EXISTING METHODS

Besides the proposed method, we also compare it with some existing traditional machine learning methods(Decision Tree, KNN, SVM, XGBoost). The results of them are shown in Fig. 9 ∼ Fig. 13.

Fig. 9 ∼ Fig. 12 show the accuracy line diagram of Decision Tree, KNN, SVM, XGBoost in every subject respectively. Fig. 13 ∼ Fig. 14 show the average accuracy and average macro-f1-score of all models respectively. We can find that models based on traditional machine learning method also have good performance. Here, SVM and XGBoost(based on emsemble learning) is better than decision tree and KNN model. However, when comparing with models based on deep learning method(such as the 1DCNN model proposed in this paper), based on the deep learning method, models based on traditional machine learning

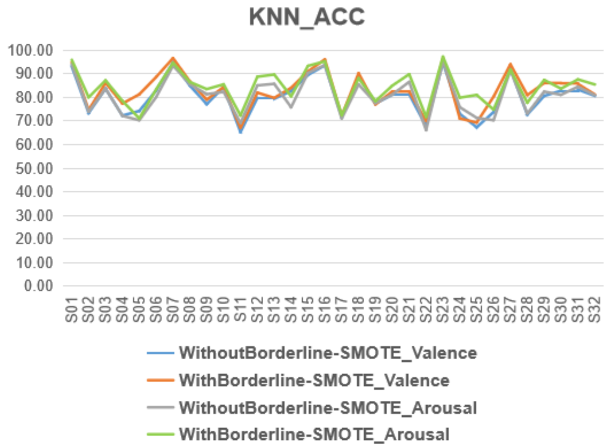

FIGURE 10. The accuracy line diagram of KNN.

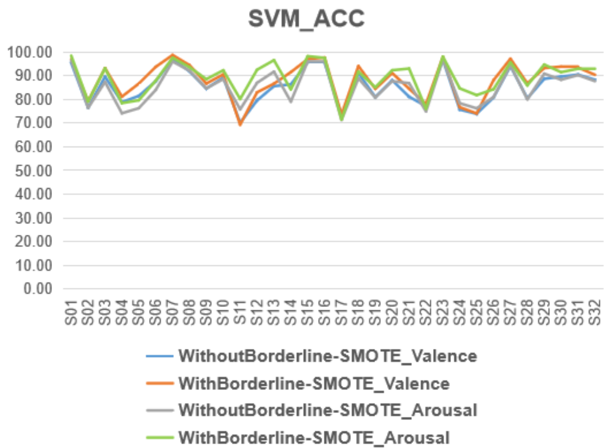

FIGURE 11. The accuracy line diagram of SVM.

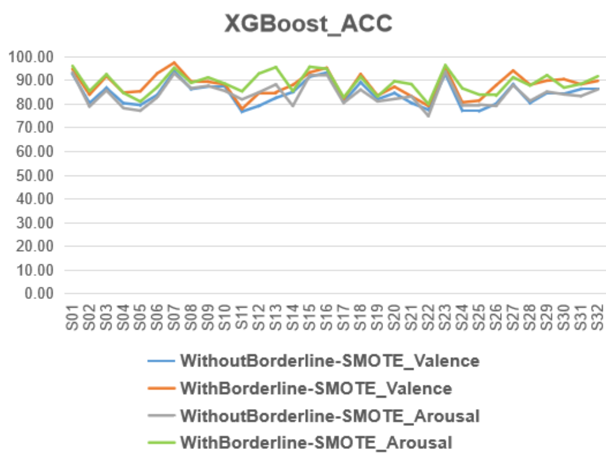

FIGURE 12. The accuracy line diagram of XGBoost.

method have a common problem: the automatic learning ability is slightly poor and is greatly affected by data distribution and well-designed features. For example, in subjects 1, 7, 15, 23 and 27, the performance of the four models is better than that in subjects 2, 11, 21 and 24. Therefore, the variance of the results is relative large, which indicates that there are some problems in stability of performance.

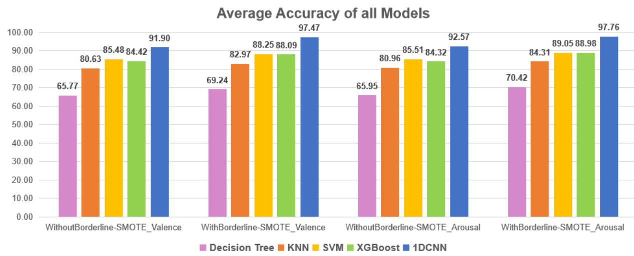

FIGURE 13. The average accuracy of all models.

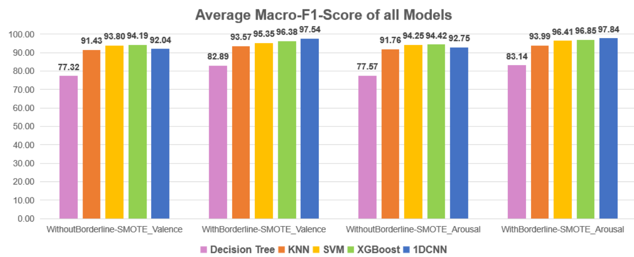

FIGURE 14. The average macro-f1-score of all models.

However, by using data augmentation method Borderline-SMOTE method, this problem has been slightly improved. On both emotional dimensions valence and arousal, the variance of the results in different subjects decreased, and the average accuracy of the above five models has been improved by $3\% \sim 5\%$, especially some individual samples, the increase could even be as much as $6\% \sim 8\%$. Therefore, in the traditional machine learning methods, the effectiveness of Borderline-SMOTE data augmentation method can be proved, and the idea of "data augmentation + classification" can be effectively promoted in many classification problems.

Meanwhile, the results of the proposed method has also been compared with some different existing methods by other researchers that used DEAP data set, as shown in Table. 7. Our proposed method represents a significant improvement over the existing methods on both emotional dimension valence and arousal, which further verifies the value of the proposed method of "data augmentation + classification" in this article.

TABLE 7. Comparison of the proposed method with some existing methods by other researchers.

Reference	Class	Classifier	Accuracy	
			Valence	Arousal
Yoon and Chung[17]	Binary	Bayesian	70.90%	70.10%
	Three		53.40%	53.10%
Parui et al.[19]	Binary	XGBoost	75.97%	74.20%
Alhagry et al.[23]	Binary	LSTM	85.65%	85.45%
Yi Zhan et al.[24]	Binary	CNN	82.95%	84.07%
Yang et al.[25]	Binary	2DCNN-LSTM	90.80%	91.03%
Our proposed method	**Three**	**1DCNN without Borderline-SMOTE**	**91.90%**	**92.57%**
		1DCNN with Borderline-SMOTE	**97.47%**	**97.76%**

C. MERITS AND LIMITATIONS OF THE PROPOSED METHOD

The model of the proposed method is designed based on 1DCNN, and the features in frequency domain, which makes the model simple and intuitive with low space cost. Meanwhile, the idea "data augmentation + classification" are validated on traditional machine learning and modern deep learning methods, which helps to enhance the adaptive ability and the performance of models.

However, the personalized emotion recognition model also has some limitations. On the one hand, the brain itself is a complex nonlinear system, 1-dimensional features still cannot fully reflect the characteristic information of EEG signals, although frequency domain feature is a kind of typical EEG signals. On the other hand, in terms of computational complexity, the process of data augmentation involves over sampling of a few types of samples, When the sample size is large or the imbalance is serious, a large time cost will be created. These limitations need to be improved in the future.

IV. CONCLUSION

To sum up, in this paper, we propose a method of emotion recognition of EEG signals based on 1DCNN with data augmentation method Borderline-SMOTE. Under the three classification task, after data pre-processing and feature extraction, this method pays close attention to the problem of data imbalance and augments the samples at and near the classification boundary to make the samples of different classes more balanced. Then, the samples are sent to the well-designed 1DCNN for three classification. Finally, the model is evaluated on DEAP data set on two emotion dimensions valence and arousal, with stratified-10-fold cross-validation. Meanwhile, we compare the model of proposed method with four traditional machine learning models(Decision Tree, KNN, SVM, XGBoost) and some existing methods by other researchers. It can be proved that the proposed method is effective and shows its advantage in emotion recognition. By comparison, the effectiveness of idea "data augmentation + classification" has been proved, which helps to enhance the adaptive ability and the performance of models. Especially when the data imbalance is very serious, for instance, the number of samples of majority classes is 10 times or more than that of other minority classes, data augmentation should be a key factor that has to be considered.

In the future, we will focus on the limitations of the model, continue tuning the parameter and optimizing the model structure, and use the model to carry out some relevant works that worth researching, like identification of depression [40], [41] and dementia stage classification [42], [43].

REFERENCES

[1] M. Soleymani, S. Asghari-Esfeden, M. Pantic, and Y. Fu, "Continuous emotion detection using EEG signals and facial expressions," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, vol. 1, Jul. 2014, pp. 1–6.

[2] S. Gallagher, "Philosophical conceptions of the self: Implications for cognitive science," *Trends Cognit. Sci.*, vol. 4, no. 1, pp. 14–21, Jan. 2000.

[3] L.-F. Rodriguez, F. Ramos, and Y. Wang, "Cognitive computational models of emotions," in *Proc. IEEE 10th Int. Conf. Cognit. Inf. Cognit. Comput. (ICCI-CC)*, Aug. 2011, pp. 75–84.

[4] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.

[5] E. Andre, "Experimental methodology in emotion-oriented computing," *IEEE Pervas. Comput.*, vol. 10, no. 3, pp. 54–57, Jul. 2011.

[6] F. P. Shah and V. Patel, "A review on feature selection and feature extraction for text classification," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Chennai, India, Mar. 2016, pp. 2264–2268.

[7] J. Deng, S. Frühholz, Z. Zhang, and B. Schuller, "Recognizing emotions from whispered speech based on acoustic feature transfer learning," *IEEE Access*, vol. 5, pp. 5235–5246, 2017.

[8] X. Cai, D. Dai, Z. Wu, X. Li, J. Li, and H. Meng, "Emotion controllable speech synthesis using emotion-unlabeled dataset with the assistance of cross-domain speech emotion recognition," 2020, *arXiv:2010.13350*. [Online]. Available: http://arxiv.org/abs/2010.13350

[9] S. Shojaeilangari, W.-Y. Yau, K. Nandakumar, J. Li, and E. K. Teoh, "Robust representation and recognition of facial emotions using extreme sparse learning," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2140–2152, Jul. 2015.

[10] P. Tahghighi, A. Koochari, and M. Jalali, "Deformable convolutional LSTM for human body emotion recognition," 2020, *arXiv:2010.14607*. [Online]. Available: http://arxiv.org/abs/2010.14607

[11] Z. Yang and S. S. Narayanan, "Modeling dynamics of expressive body gestures in dyadic interactions," *IEEE Trans. Affect. Comput.*, vol. 8, no. 3, pp. 369–381, Mar. 2017.

[12] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[13] C. Joesph, A. Rajeswari, B. Premalatha, and C. Balapriya, "Implementation of physiological signal based emotion recognition algorithm," in *Proc. IEEE 36th Int. Conf. Data Eng. (ICDE)*, Dallas, TX, USA, Apr. 2020, pp. 2075–2079.

[14] W. A. Jang, S. M. Lee, and D. H. Lee, "Development BCI for individuals with severely disability using EMOTIV EEG headset and robot," in *Proc. Int. Winter Workshop Brain-Comput. Interface (BCI)*, Jeongsun-kun, South Korea, 2014, pp. 1–3.

[15] T. M. Vaughan, D. J. Mcfarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw, "The wadsworth BCI research and development program: At home with BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 229–233, Jun. 2006.

[16] S. Hwang, M. Ki, K. Hong, and H. Byun, "Subject-independent EEG-based emotion recognition using adversarial learning," in *Proc. 8th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Gangwon, South Korea, Feb. 2020, pp. 1–4.

[17] P. Lahane and M. Thirugnanam, "A novel approach for analyzing human emotions based on electroencephalography (EEG)," in *Proc. Innov. Power Adv. Comput. Technol. (I-PACT)*, Vellore, India, 2017, pp. 1–6.

[18] M. R. Islam and M. Ahmad, "Wavelet analysis based classification of emotion from EEG signal," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2019, pp. 1–6.

[19] I. Wichakam and P. Vateekul, "An evaluation of feature extraction in EEG-based emotion prediction with support vector machines," in *Proc. 11th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Chon Buri, Thailand, May 2014, pp. 106–110.

[20] H. J. Yoon and S. Y. Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2230–2237, Dec. 2013.

[21] S. Parui, A. K. Roshan Bajiya, D. Samanta, and N. Chakravorty, "Emotion recognition from EEG signal using XGBoost algorithm," in *Proc. IEEE 16th India Council Int. Conf. (INDICON)*, Rajkot, India, Dec. 2019, pp. 1–4.

[22] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Jun. 2019, Art. no. 031001.

[23] S. K. Goh, H. A. Abbass, K. C. Tan, A. Al-Mamun, N. Thakor, A. Bezerianos, and J. Li, "Spatio–Spectral representation learning for electroencephalographic gait-pattern classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1858–1867, Sep. 2018.

[24] X. Xing, Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu, "SAE+LSTM: A new framework for emotion recognition from multi-channel EEG," *Frontiers Neurorobot.*, vol. 13, p. 37, Jun. 2019.

[25] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 355–358, 2017.

[26] Y. Zhan, M. I. Vai, S. Barma, S. H. Pun, J. W. Li, and P. U. Mak, "A computation resource friendly convolutional neural network engine for EEG-based emotion recognition," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environ. Meas. Syst. Appl. (CIVEMSA)*, Tianjin, China, Jun. 2019, pp. 1–6.

[27] Y. Yang, Q. Wu, M. Qiu, Y. Wang, and X. Chen, "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–7.

[28] O. E. Par, E. Akcapinar Sezer, and H. Sever, "Small and unbalanced data set problem in classification," in *Proc. 27th Signal Process. Commun. Appl. Conf. (SIU)*, Sivas, Turkey, Apr. 2019, pp. 1–4.

[29] Y. Hu, X. Bai, P. Zhou, F. Shang, and S. Shen, "Data augmentation imbalance for imbalanced attribute classification," 2020, *arXiv:2004.13628*. [Online]. Available: http://arxiv.org/abs/2004.13628

[30] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jun. 2012.

[31] S. Kundu and S. Ari, "Score normalization of ensemble SVMs for brain-computer interface P300 speller," in *Proc. 8th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Delhi, India, Jul. 2017, pp. 1–5.

[32] Y. Zhang, S. Zhang, and X. Ji, "EEG-based classification of emotions using empirical mode decomposition and autoregressive model," *Multimedia Tools Appl.*, vol. 77, pp. 26697–26710, Oct. 2018.

[33] S. Reddy, S. B. Seshadri, G. Sankesh Bothra, T. G. Suhas, and S. C. Thundiyil, "Detection of arrhythmia in real-time using ECG signal analysis and convolutional neural networks," in *Proc. IEEE 21st Int. Conf. Comput. Problems Electr. Eng. (CPEE)*, Poland, Warsaw, Sep. 2020, pp. 1–4.

[34] J. V. F. Franca, J. M. C. Boaro, P. T. C. dos Santos, F. Henrique, V. Garcia, and C. Manfredini, "Legal judgment prediction in the context of energy market using gradient boosting," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Toronto, ON, Canada, Oct. 2020, pp. 875–880.

[35] T. Lee, M. Kim, and S.-P. Kim, "Data augmentation effects using borderline-SMOTE on classification of a P300-based BCI," in *Proc. 8th Int. Winter Conf. Brain-Computer Interface (BCI)*, Gangwon, South Korea, Feb. 2020, pp. 1–4.

[36] H. Hui, W.-Y. Wang, and B.-H. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning," in *Proc. Int. Conf. Intell. Comput.* Berlin, Germany: Springer, 2005, pp. 878–887.

[37] S. Sablok, G. Gururaj, N. Shaikh, I. Shiksha, and A. R. Choudhary, "Interictal spike detection in EEG using time series classification," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Madurai, India, May 2020, pp. 644–647.

[38] F. Chollet, "Keras: The Python deep learning library," in *Astrophysics Source*. ASCL.net, 2018.

[39] S. Yadav and S. Shukla, "Analysis of K-fold cross-validation over hold-out validation on colossal datasets for quality classification," in *Proc. IEEE 6th Int. Conf. Adv. Comput. (IACC)*, Bhimavaram, India, Feb. 2016, pp. 78–83.

[40] A. Ashraf, T. S. Gunawan, F. D. A. Rahman, M. Kartiwi, N. Ismail, and J. Ulfiah, "A summarization of the visual depression databases for depression detection," in *Proc. 6th Int. Conf. Wireless Telematics (ICWT)*, Yogyakarta, IN, USA, Sep. 2020, pp. 1–6.

[41] L. F. L. Villalba, M. A. M. Pelaez, and E. R. L. Arias, "Low-cost identification system of ICT for depression in children 8-11 years old: The case of the IED Edgardo Vives de Santa Marta, Colombia," in *Proc. 15th Iberian Conf. Inf. Syst. Technol. (CISTI)*, Sevilla FC, Spain, Jun. 2020, pp. 1–7.

[42] P. Durongbhan, Y. Zhao, L. Chen, P. Zis, M. De Marco, Z. C. Unwin, A. Venneri, X. He, S. Li, Y. Zhao, D. J. Blackburn, and P. G. Sarrigiannis, "A dementia classification framework using frequency and time-frequency features based on EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 826–835, May 2019.

[43] C. Wang, J. Xu, S. Zhao, and W. Lou, "Identification of early vascular dementia patients with EEG signal," *IEEE Access*, vol. 7, pp. 68618–68627, 2019.

YU CHEN was born in Yichun, Heilongjiang, in 1975. He received the bachelor's degree in computational mathematics and applied software from Heilongjiang University, in 1997, the master's degree in computer science and technology from the Harbin University of Science and Technology, in 2005, and the Ph.D. degree in test measurement and instrument from Heilongjiang University, in 2009.

He has presided more than ten projects, including the National Natural Science Foundation of China, the Provincial Post-doctoral Foundation, the Basic Scientific Research Business Fund of Central universities, the Natural Science Foundation of Heilongjiang Province, and the scientific and technological research project of the Provincial Education Department. He has published more than 80 articles and five textbooks. His research interests include mathematical physics inverse problems, detection and imaging, machine learning, natural language processing, enterprise search and algorithm optimization, and image processing.

He is currently an Expert Member of the China Discrete Mathematics Society, the northeast China of ACM College Student Programming Competition, the Heilongjiang Provincial Computer Society Competition Committee, and the Heilongjiang Provincial Information Science Olympic Competition. He is also a Reviewer of *Chinese Journal of Image and Graphics*, *Journal of Harbin University of Science and Technology*, and *Journal of System Simulation*.

RUI CHANG was born in Zibo, Shandong, in 1996. He received the B.E. degree in computer science and technology from Yantai University, in 2019. He is currently pursuing the master's degree in computer technology with Northeast Forestry University.

His research interests include machine learning, brain-computer interface technology, and emotion recognition.

During his undergraduate and postgraduate studies, he has won many honors and awards, such as mathematics contest, English contest, programming contest, various university-level scholarships, and provincial outstanding graduate.

JIFENG GUO was born in Yichun, Heilongjiang, in 1974. He received the bachelor's degree in electrical engineering from the Harbin University of Science and Technology, in 1995, the master's degree in computer application technology from Northeast Forestry University, in 2005, and the Ph.D. degree in electrical engineering and automation from the Harbin Institute of Technology, in 2011.

He has presided more than ten projects of Heilongjiang Provincial Natural Fund, the Harbin Innovation Fund, the Provincial Department of Education science and technology research project, and the central university basic Scientific research Business fund. He has participated in more than 20 projects of 863 Plan Project, 948 Project, the National Natural Science Fund, and the Heilongjiang Provincial Science and Technology tackling key problems. He has published more than 30 articles as the first or corresponding author, 20 of which are included in SCI/EI, Published two textbooks, authorized 12 invention/utility model patents, and five software copyrights. His research interests include brain-like computing, artificial intelligence, intelligent control, and electrical reliability theory.

He is currently a member of CCF and CCF Harbin Branch. He is also the Director of Computer Basic Education and Experiment Center; the Deputy Director of Forestry and Woodworking Mechanical Engineering Technology Center, Northeast Forestry University; and the Deputy Director of Applied Software Research Institute, Northeast Forestry University.