Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks

Fang Wang[1], Sheng-hua Zhong[1(✉)], Jianfeng Peng[1], Jianmin Jiang[1], and Yan Liu[2]

[1] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518000, People's Republic of China
{wangfang20161, pengjianfeng2017}@email.szu.edu.cn, {csshzhong, jianmin.jiang}@szu.edu.cn

[2] Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
csyliu@comp.polyu.edu.hk

Abstract. Emotion recognition is the task of recognizing a person's emotional state. EEG, as a physiological signal, can provide more detailed and complex information for emotion recognition task. Meanwhile, EEG can't be changed and hidden intentionally makes EEG-based emotion recognition achieve more effective and reliable result. Unfortunately, due to the cost of data collection, most EEG datasets have small number of EEG data. The lack of data makes it difficult to predict the emotion states with the deep models, which requires enough number of training data. In this paper, we propose to use a simple data augmentation method to address the issue of data shortage in EEG-based emotion recognition. In experiments, we explore the performance of emotion recognition with the shallow and deep computational models before and after data augmentation on two standard EEG-based emotion datasets. Our experimental results show that the simple data augmentation method can improve the performance of emotion recognition based on deep models effectively.

Keywords: Emotion recognition · Data augmentation · EEG

1 Introduction

Emotional recognition is the process of identifying human emotional state. As an interdisciplinary field, the research of emotional recognition is benefited from the development of psychology, modern neuroscience, cognitive science, and computer science as well [1]. For computer science, the emotion recognition based on computer system aspires to enhance human-machine interaction across a wide range of application domains including clinical, industrial, military, gaming and so on [2].

Various approaches have been proposed for emotional recognition and can be divided into two categories [2]. The first category is using the features of emotional behavior, such as facial expression, the tone of voice, body gestures and so on, to detect

K. Schoeffmann et al. (Eds.): MMM 2018, Part II, LNCS 10705, pp. 82–93, 2018.
https://doi.org/10.1007/978-3-319-73600-6_8

a specific emotion. The second category is taking advantage of the physiological signals to recognize emotions. These physiological signals include the electroencephalograph (EEG), electrocardiograph (ECG), pulse rate, respiration signals, etc. Compared with the former, the physiological signals not only provide more detailed and complex information for estimating emotional states but also yield more effective and reliable recognition results.

Up to now, a number of EEG-based emotion recognition methods have been studied. Emotional recognition mainly includes the two modules: feature extraction and emotion classification. As a time-series signal, the EEG original signal has potential information and can't be directly used for emotion recognition [3]. In order to fully exploit the potential information, the classifier uses the feature data as input to identify the emotional states. In the traditional EEG-based emotion recognition methods, many studies explored the effects of emotional recognition using the traditional machine learning models as the classifiers, such as Support Vector Machine [4–6], K Nearest Neighbors [7–9], and so on.

In recent years, deep learning techniques have received widespread attention and achieved remarkable results in many fields. Using the deep learning model as a classifier to complete the classification task, especially in the image recognition task, show an appreciable performance [10, 11]. Recently, the deep learning method has been applied to the study of the EEG-based emotional recognition [12–14]. Zheng and Lu compared the performance of the deep learning models with the feature-based shallow models on EEG-based emotion recognition and showed the superior performance of the deep learning models [1]. Based on the above achievements, it is necessary for us to promote exploration on the EEG-based emotion recognition with deep learning models. However, compared with the shallow models, the deep learning models, for example, the deep convolution neural network, have more model parameters, which makes the training of deep learning models requires a large number of labeled training samples. But due to the expensive cost, for EEG-related task, only a limited amount of labeled data can be obtained to train the model. For most public EEG datasets, they only have a few hundred samples, which are collected for different tasks with different specifications. Thus, if we try to further explore deep learning models on the task of EEG-based emotional recognition, getting sufficient and effective labeled training data based on existing datasets is the primary issue needs to be addressed.

In this paper, we focus on applying simple data augmentation method to generate more EEG training samples. After that, we compare the performance of EEG-based emotion recognition with the shallow and the deep models before and after data augmentation.

The rest of this paper is organized as followings. Section 2 briefly reviews the representative work on EEG-based emotion recognition. In Sect. 3, we introduce our proposed method in detail. In Sect. 4, we provide a series of experiments to validate the proposed method on the standard datasets. The paper is closed with the conclusion and future work.

2 Related Work

EEG-based emotion recognition, as an important branch of emotion recognition, has received much attention in the past decades. Davidson and Fox investigated that infants show greater activation of the left frontal than of the right frontal area in response to the happy segments [15]. Klimesch *et al*. proposed the beta frequency band of EEG reflects the emotional and cognitive patterns [16]. Li and Lu indicated that gamma band of EEG can be used for classifying happiness and sadness [17]. In the early 21st century, the rapid development of dry electrode technologies and wearable devices allows us to record and analyze the EEG signals from the laboratories to the real environment. At the same time, the application of EEG-based emotion recognition in real-world has also been further promoted. Sourina *et al*. proposed to combine music therapy process with the real-time EEG-based human emotion recognition algorithm [18]. Kothe and Makeig collected the real-time EEG signals to perform two-class discrimination task between "high" and "low" workload levels [19]. Shi *et al*. proposed to use the extreme learning machine to perform the EEG-based vigilance estimation [3].

As can be seen from the above, the study of EEG-based emotional recognition has never stopped. Recent years, various studies about using machine learning techniques to establish emotion recognition model have been presented. Chanel *et al*. used EEG time-frequency information as features and SVM as a classifier to distinguish three emotional states [20]. Heraz and Frasson reported using the amplitudes of four EEG components as features and KNN as a classifier to characterize EEG signals into eight emotional states [21]. Zheng *et al*. used some commonly used features obtained from various frequency bands and discriminative Graph regularized Extreme Learning Machine as a classifier to build EEG-based emotion recognition systems [22]. Recently, deep learning models are applied to process EEG signals. Li *et al*. applied a DBN model to detect the emotional states from EEG signals. They compared the result with five baselines and got the improvement from 11.5% to 24.4% [23]. Mao *et al*. proposed two new architectures of convolutional neural networks and proved the potential of deep learning methods for real-life EEG-based biometric identification [24]. Li *et al*. designed a hybrid deep learning model that combined the CNN and RNN for mining inter-channel correlation and contextual information from EEG frames [25]. So far, using deep learning methods to identify emotions from EEG signals is still in its infancy. Due to the limitation of the cost of data collection, the labeled EEG samples that can be used to train the deep learning models are significantly insufficient. Therefore, augmenting the labeled EEG samples is the key to promoting the use of deep learning techniques for EEG-based emotional recognition.

3 The EEG-Based Emotional Recognition Framework Based on Data Augmentation

In this paper, we propose to use data augmentation on EEG-based emotion recognition task. The whole process including training and testing stages is demonstrated in Fig. 1. The EEG signals are recorded while subjects are watching emotional videos. The differential entropy (DE) feature is extracted from the recorded EEG signal. In the training stage, we use the data augmentation method to generate more EEG data. These data are input to the classifier to train a learned model. And this model is used to predict the label of EEG data in the test stage.

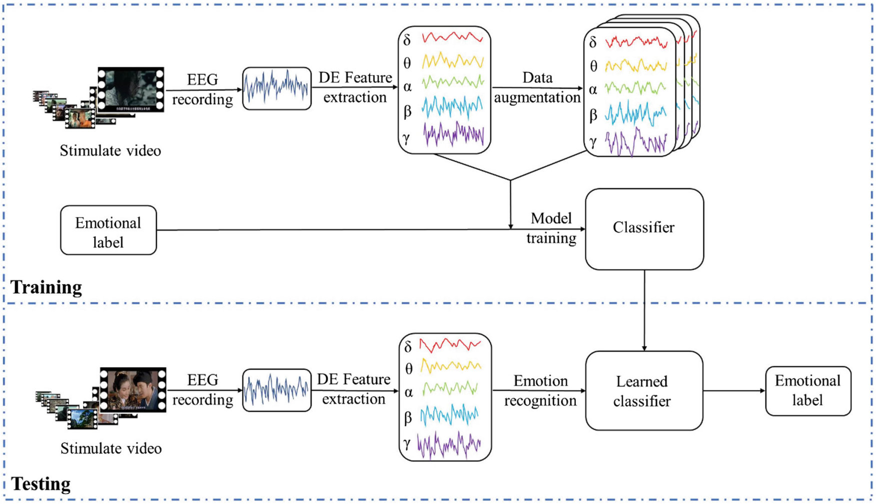

Fig. 1. The EEG-based emotional recognition framework based on data augmentation.

3.1 Feature Extraction

In our method, we extract the differential entropy (DE) feature from the recorded EEG signal segment. It has been shown that DE features can obtain the superior performance in comparison with other commonly used features [1, 22]. For the time series X obeying the Gauss distribution $N(\mu, \sigma^2)$, its DE can be defined as the following formulation:

$$h(X) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx = \frac{1}{2}\log 2\pi e\sigma^2 \quad (1)$$

where μ is the mean value and σ is the standard deviation. For a fix length EEG signal, DE feature is equivalent to the logarithm energy spectrum in a certain frequency band [2, 3]. Assume that the EEG signal is recorded with n-channel and the length of an EEG signal segment is l s. First, we apply the band pass filter to obtain the 5 frequency bands of each channel. The ranges of the filter are delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), gamma (31–50 Hz). Then, we use a 256-point short-time Fourier transform with a non-overlapped Hamming window of 1 s to obtain the energy spectrum of each specified frequency band. After that, we extract the DE feature for each frequency band by calculating the logarithm energy spectrum. Thus, the DE feature for each second consists of n-channel DE calculated across 5 frequency bands. After feature extraction, the size of each EEG feature sample is $n \times l \times 5$.

3.2 Data Augmentation

Data augmentation is the process of generating new samples by transforming training data, with the aim of improving the accuracy and robustness of classifiers [26]. Improper data augmentation methods are likely to increase the amount of the samples that are not informative enough, which will do not even reduce the recognition accuracy and robustness of recognition systems. In our work, we consider using the commonly used data augmentation methods in images processing to increase the number of EEG samples.

There are two basic data augmentation approaches used in images processing: geometric transformation and noise addition. Geometric transformations, including shift, scale, rotation/reflection and etc., are not directly suitable for enlarging our EEG data. Compared with the image, the EEG signal is a continuous signal that changes over time. Even if the feature extraction is performed, its features are still a time series. Thus, if we rotate or shift the EEG data, the feature of time domain will be destroyed. To avoid this problem, we further consider using the noise addition method to augment the EEG samples. In theory, there are many ways for us to add noise (Gaussian, Poisson, Salt, Pepper, etc.) into the EEG data. But EEG signal has a strong randomness and non-stationarity. If we randomly add some local noises, such as Poisson noise, Salt noise, or Pepper noise, which will change the features of EEG data locally.

Based on these considerations, in our work, we focus on adding Gaussian noise to each feature sample of the original training data to obtain new training samples. The probability density function P of a Gaussian random variable z is defined by:

$$P_G(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \tag{2}$$

where z represents the grey level, μ is the mean value and σ is the standard deviation. In our work, in order to ensure that the amplitude value of the sample will not be changed with the addition of noise, we generate the Gaussian noise with $\mu = 0$. Under this premise, in order to explore the effect of different noise intensity (standard deviation) on the work of EEG data augmentation, the value of σ is set to 0.001, 0.01, 0.02, 0.1, 0.2, and 0.5. We use m to represent the augmented multiple. When $m = 1$, it means there is no data augmentation stage, and the train samples input into the classifier only include the original training data.

3.3 Machine Learning Models

In this paper, we use three machine learning models as the classifiers to recognize EEG-based emotion category. These three models are support vector machine (SVM), LeNet, and ResNet.

SVM. As a commonly used machine learning model, the basic idea of SVM is to map the input data to a high-dimensional feature space via a kernel transfer function, in this new space, these input data will be easier to be separated than that in the original feature space [1]. SVM is often used in EEG-based emotional recognition tasks [27–29]. In our work, the input of SVM model is the DE feature of EEG signal segment with duration of 1s. In order to meet the demand of SVM for the input, we transform the size of DE feature sample from $n \times 1 \times 5$ to $1 \times (n \times 5)$. Each EEG signal segment will be divided into l samples. In training, all these l samples of each EEG signal segment have the same emotional labels with the corresponding EEG signal segment. In test, we obtain the predictive emotional label of an EEG signal segment by counting the frequency of the predictive labels of the l test samples corresponding to this EEG segment. The most frequently emotional predictive label will be identified as the predictive label for this EEG segment.

LeNet. LeNet is a typical convolutional network for recognizing characters. It can learn complex, high-dimensional, nonlinear mappings from large collections of examples [10]. In our work, the architecture of LeNet is C32@32 × 32- P32@16 × 16- C32@16 × 16- P32@8 × 8- C64@8 × 8- P64@4 × 4- C64@1 × 1- F3, which contains four convolutional layers (denoted by C with the number and output size of the feature map), three pooling layers (denoted by P with the number and output size of the feature map) and one fully-connected layer (denoted by F with the number of output). The input sample of this model is the DE feature of EEG signal segment with duration of 62 s. The size of an input sample is $n \times 62 \times 5$. Each EEG signal segment will be divided into $l/62$ samples. We use the same method of counting the frequency with SVM to obtain the predictive label for each EEG test segment.

ResNet. ResNet can ease the training of deeper network with a residual learning framework. It has been proven that this network is easier to optimize, and can gain accuracy from considerably increased depth [30]. In our work, we use the ResNet to establish the EEG-based emotion recognition model. The architecture of this network is C64@62 × 62- P64@56 × 56- C64@56 × 56- C64@56 × 56- C256@56 × 56- C128@28 × 28- C128@28 × 28- C512@28 × 28- C256@14 × 14- C256@14 × 14- C1024@14 × 14- C512@7 × 7- C512@7 × 7- C2048@7 × 7- P2048@1 × 1- F3. It contains thirteen convolutional layers, two pooling layers, and one fully-connected layer. It has been proven that alpha, beta and gamma bands of EEG are more predictive to the emotional states compared with the delta and theta bands [1, 17]. In our experiment, we use the feature of the three high frequency bands (alpha, beta, and gamma) as the input of ResNet. The size of an input sample is $n \times 62 \times 3$. We use the same voting method with LeNet to obtain the predictive label for each EEG test segment.

4 Experiments

4.1 Experimental Setting

In this paper, we investigate the performance of EEG-based emotion recognition with the shallow and deep model before and after data augmentation on two standard EEG-based emotional datasets: SJTU Emotion EEG Dataset (SEED) [1] and MAHNOB-HCI dataset [31]. Two classical convolutional neural networks are utilized in this work, including LeNet and ResNet. SVM, which is a representative shallow model, is included in our comparison. We use LIBSVM Toolbox [32] to implement

SVM classifier and use the linear kernel and the grid optimization to find the optimal value of C in the parameter space $[2^{-10}:2^{10}]$ with a step of one. The LeNet and ResNet are implemented by the MATCONVNET Toolbox [33]. To the parameters for LeNet and ResNet, we follow the general setting in MATCONVNET. For example, the learning rate is set to 0.1, and the batch size is set to 100. All statistical experiments are repeated for five times, and the average results are reported.

4.2 Experimental Results on SEED Dataset

As a standard dataset for emotion recognition, the SEED dataset consists of 630 EEG segments of 14 subjects while watching 15 emotional film clips for three sessions, which were recorded by the EEG cap according to the international 10–20 system for 62 channels. For every EEG segment, the duration is about 4 min. After perform the processing, the raw EEG data is downsampled to 200 Hz sampling. Three basic emotion states are included in this dataset: positive, neutral and negative. Each emotional state is with an equal number of segments. In our experiment, we segment every original segment into three short EEG samples with the length of 62 s. Totally, there are 1890 samples for the experiments. For these samples, all the samples of the first nine segments in each session are utilized as the training data, the rest are used as the test data.

In this section, we first compare the accuracies of EEG-based emotion recognition based on SVM, LeNet and ResNet models without data augmentation. In addition, we also explore the effect of dimensionality reduction on emotion recognition. In the experiment, we use the Principle component analysis (PCA) algorithm to reduce the original feature dimension. Table 1 shows the emotion recognition accuracies of these models without data augmentation. Among them, SVM is selected as the compared model because it is also applied as the classifier for emotion recognition on the same dataset [1, 22]. LeNet and ResNet for EEG-based emotion recognition are applied and implemented by ourselves.

Table 1. The average accuracy of three models on the original data without data augmentation.

Models	SVM [1]	PCA + SVM (95% energy)	PCA [22] + SVM (160 dimensions)	PCA [22] + SVM (210 dimensions)	LeNet	ResNet
Accuracy (%)	74.2	49.8	56.9	58.1	49.6	34.2

From these results, we can find that the accuracies of three models are different. For the SVM model, the accuracy is 74.2%, which is better than the performance of LeNet and ResNet. When we use PCA to reduce the original 310 dimensional feature to 160 dimensions, 210 dimensions and the number of PCA dimensions retaining the top 95% energy, the accuracies drop from 74.2% to 56.9%, 58.1% and 49.8%. We can see that dimensionality reduction will affect the accuracy of emotion recognition to a certain extent. That is because the complexity of EEG data makes the simple dimension reduction method may not be able to preserve the important discriminative information of original domain information. This result is also consistent with the conclusion of existing work [22]. However, it can be found that these results are still better than the results of two deep learning models. We believe that the consistency of the number of training samples and the number of free parameters in the model is the main factor that affects the experimental results. For the LeNet used in our work, there are about 4000 network parameters that need to be learned. For the ResNet, the number of the network parameters in it is more than 20000. Unfortunately, in this experiment, only 1134 training samples can be used in the training stage of the recognition model. The number of training samples can't meet the requirement of the deep learning models. Compared with these deep learning models, SVM shows better emotional accuracy because SVM is not sensitive to the number of sample.

Table 2. The average accuracy (%) of data augmentation on LeNet.

Augmented multiple (m)	Standard deviation of Gaussian noise (σ)					
	0.001	0.01	0.02	0.1	0.2	0.5
5	67.4	68.8	66.6	68.1	69.0	61.6
20	69.9	70.2	68.3	71.7	73.4	70.8
30	68.9	69.8	68.9	71.4	**74.3**	70.9

As we have described in Table 1, the lack of EEG training samples may affect the accuracy of emotional recognition of those deep models. In order to prove this conjecture and solve this problem, we use the proposed data augmentation method in Sect. 3.2 to increase the amount of training samples and test its effect on LeNet. Table 2 shows the recognition accuracies of LeNet trained by the augmented training data. From the table, we find that data augmentation can effectively improve the performance of LeNet. We can achieve the best accuracy of 74.3% when the standard deviation is 0.2 and the number of training samples is augmented to 30 times. The standard deviation of Gaussian noise used in data augmentation and the scale of the augmented training data will affect the performance. Too small or too large standard deviation can't generate the effective new samples.

We have already proved that data augmentation method can improve the performance of EEG-based emotion recognition with LeNet. In order to further explore whether this method is also effective for other machine learning models, we also conduct experiments on SVM and ResNet with the augmented data. In Fig. 2, the emotion recognition accuracies of SVM, LeNet and ResNet obtained before and after data augmentation are demonstrated. In experiments, we follow the best parameter setting of data augmentation obtained in Table 2. For each model, the mean value μ is set to be 0, and the standard deviation is set to be 0.2. For LeNet and ResNet, the training-data is augmented to 30 times. Owing to the extremely high computational complexity, the LIBSVM Toolbox only permits the number of training samples to be augmented 5 times. Hence, in this experiment, for SVM, the training data is only augmented to 5 times.

From Fig. 2, we can see that increasing the amount of input training data cannot improve the recognition accuracy of SVM (74.2% vs. 73.4%). As we known, SVM is insensitive to the number of training data [34]. As with LeNet, data augmentation can effectively improve the accuracy of ResNet. Compared with LeNet, the input sample of ResNet only includes the DE feature of three frequency bands of EEG signals. Under this situation, the accuracy of ResNet is improved from 34.2% to 75.0%, better than the LeNet (from 49.6% to 74.3%). These results also evidence that data augmentation is a useful method to address the issue of the lack of training samples.

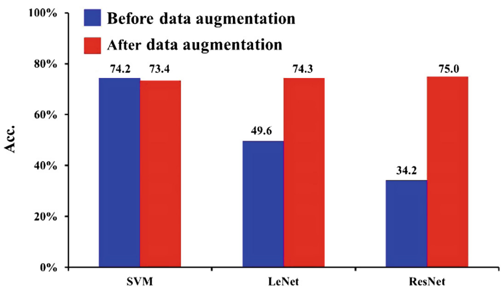

Fig. 2. The average accuracy of three models trained before and after data augmentation.

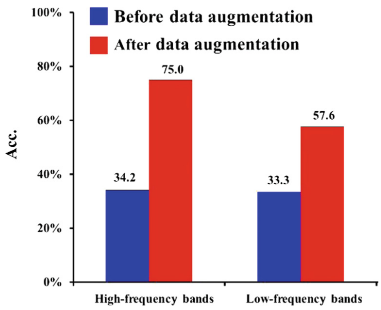

Fig. 3. The accuracies of the high and low frequency bands on ResNet.

In the last experiment of this dataset, we have explored the contribution of low and high frequency bands for emotion recognition. In detail, the high frequency bands include: alpha, beta, and gamma. The low frequency bands include: delta, theta, and alpha. The emotion recognition accuracies of the low and high frequency bands with ResNet before and after data augmentation are shown in Fig. 3. From the figure, we can find that the accuracy of the three low frequency bands before and after data augmentation is 33.3% and 57.6%, respectively, both of which are lower than the high-frequency bands.

4.3 Experimental Results on MAHNOB-HCI Dataset

The MAHNOB-HCI dataset, a multi-model dataset for affect recognition and implicit tagging, includes the EEG data recorded by the EEG cap according to the international 10–20 system for 32 channels. The length of the emotional video as the stimulus is between 34.9 and 117 s. Before the feature extraction, we remove the artifacts from the original EEG data and use the average reference as the virtual reference. In our experiments, we divide the EEG segments into three classes according the self-reported feeling of the subjects in valence space, negative (1–3), neutral (4–6) and positive (7–9). Totally, there are 188 negative, 208 neutral samples and 131 positive samples.

In Sect. 4.2, we have explored the effect of data augmentation on the SEED dataset. In this section, in order to verify the universality of our proposed data augmentation method, we continue to perform experiments on the MAHNOB-HCI dataset and compare the performance before and after data augmentation on shallow and deep models. In Table 3, we show the best emotion recognition results obtained by different models on the MAHNOB-HCI dataset. For SVM, the training data is augmented to 30 times, the mean value μ is set to be 0, and the standard deviation is set to be 0.2. For ResNet, the training data is augmented to 30 times, the mean value μ is set to be 0, and the standard deviation is set to be 0.01. In the confusion matrices of Table 3, the row represents the predicted label, the column represents the ground truth, and the number denotes the recognition accuracy in percentage.

Table 3. The confusion matrices for recognition accuracy (%) of different models on the MAHNOB-HCI dataset before and after data augmentation (row: predicted label; column: ground truth)

	SVM						ResNet					
	Before			After			Before			After		
	Neg.	Neu.	Pos.	Neg.	Neu.	Pos.	Neg.	Neu.	Pos.	Neg.	Neu.	Pos.
Neg.	**0**	0	0	**0**	0	0	**8.1**	4.3	9.3	**56.5**	44.9	41.9
Neu.	93.5	**95.7**	81.4	96.8	**98.6**	79.1	91.9	**95.7**	90.7	16.1	**34.8**	11.6
Pos.	6.5	4.3	**18.6**	3.2	1.4	**20.9**	0	0	**0**	27.4	20.3	**46.5**
Acc.	42.5			44.3			40.8			45.4		

From Table 3, we can learn the details of SVM and ResNet before and after data augmentation on the MAHNOB-HCI dataset. For SVM, we can see that the average accuracy is improved from 42.5% to 44.3% after data augmentation. Among the three categories, the neutral emotion can be recognized with high accuracy, and after data augmentation, the classification performance of the neutral emotion and positive emotion are improved. However, for the shallow model, SVM, the proposed data augmentation method does not improve the recognition accuracy of negative emotions. For ResNet, we can find that data augmentation can also improve the performance of emotion recognition (from 40.8% to 45.4%). Before data augmentation, ResNet confuses the positive emotion with the neutral and negative emotions. Data augmentation can significantly improve the accuracies of positive and negative emotion for ResNet.

5 Conclusion

In this paper, we propose to use the data augmentation method to solve the problem that the amount of EEG data can't meet the needs of EEG-based emotional recognition with deep models. In order to explore the effect of data augmentation, we conduct a series of experiments on two standard EEG-based emotion recognition datasets. On the SJTU dataset, we first compare the accuracy of shallow (SVM) and deep models (LeNet, ResNet) without data augmentation. We find that the shallow model can achieve better performance than the deep model. Then, we use the proposed data augmentation method to generate new training samples. By analyzing the experimental result, we find that the data augmentation method can effectively improve the performance of deep models. In order to further confirm the effectiveness and universality of our proposed method, we validate the experiment on the MAHNOB-HCI dataset. Our results show that the data augmentation is a useful method to address the issue of the lack of training samples in EEG data for deep learning models. In future, we will seek to use other data augmentation methods, such as generative adversarial networks, to generate more effective samples of EEG data and improve the performance of EEG-based emotion recognition.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (No. 61502311, No. 61373122, No. 61620106008), the Natural Science Foundation of Guangdong Province (No. 2016A030310053, No. 2016A030310039), the Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) (No. U1501501), Shenzhen Emerging Industries of the Strategic Basic Research Project under Grant (No. JCYJ20160226191842793), the Shenzhen high-level overseas talents program, and the Tencent "Rhinoceros Birds" - Scientific Research Foundation for Young Teachers of Shenzhen University (2015, 2016).

References

1. Zheng, W.L., Lu, B.L.: Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Trans. Auton. Ment. Dev. **7**(3), 1 (2015)
2. Duan, R.N., Zhu, J.Y., et al.: Differential entropy feature for EEG-based emotion classification. In: International IEEE/EMBS Conference on Neural Engineering, vol. 8588, pp. 81–84 (2013)
3. Shi, L.C., Jiao, Y.Y., et al.: Differential entropy feature for EEG-based vigilance estimation. In: International Conference of the IEEE Engineering in Medicine & Biology Society, p. 6627 (2013)
4. Lin, Y.P., Wang, C.H., et al.: EEG-based emotion recognition in music listening. IEEE Trans. Biomed. Eng. **57**(7), 1798 (2010)
5. Petrantonakis, P.C., Hadjileontiadis, L.J.: A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. IEEE Trans. Inf. Technol. Biomed. **15**(5), 737 (2011). A Publication of the IEEE Engineering in Medicine & Biology Society
6. Wang, X.W., Nie, D., et al.: Emotional state classification from EEG data using machine learning approach. Neurocomputing **129**(4), 94–106 (2014)
7. Heraz, A., Frasson, C.: Predicting the three major dimensions of the learner's emotions from brainwaves. Int. J. Comput. Sci. **2**(3), 187–193 (2007)
8. Murugappan, M., Ramachandran, N., et al.: Classification of human emotion from EEG using discrete wavelet transform. J. Biomed. Sci. Eng. **03**(4), 390–396 (2010)
9. Brown, L., Grundlehner, B., et al.: Towards wireless emotional valence detection from EEG. In: International Conference of the IEEE Engineering in Medicine & Biology Society, pp. 2188–2191 (2011)
10. Lecun, Y., Bottou, L., et al.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
11. Krizhevsky, A., Sutskever, I., et al.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105 (2012)
12. Soleymani, M., Esfeden, S.A., et al.: Analysis of EEG signals and facial expressions for continuous emotion detection. IEEE Trans. Affect. Comput. **7**(1), 17–28 (2017)

13. Li, M.G., Zhu, W., et al.: The novel recognition method with optimal wavelet packet and LSTM based recurrent neural network. In: IEEE International Conference on Mechatronics and Automation, pp. 584–589 (2017)
14. Kumar, S., Sharma, A., et al.: A deep learning approach for motor imagery EEG signal classification. In: Computer Science and Engineering, pp. 34–39 (2017)
15. Davidson, R.J., Fox, N.A.: Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. Science **218**(4578), 1235–1237 (1983)
16. Klimesch, W., Doppelmayr, M., et al.: Induced alpha band power changes in the human EEG and attention. Neurosci. Lett. **244**(2), 73 (1998)
17. Li, M., Lu, B.L.: Emotion classification based on gamma-band EEG. In: International Conference of the IEEE Engineering in Medicine & Biology Society, vol. 2009, pp. 1323–1326 (2009)
18. Sourina, O., Liu, Y., et al.: Real-time EEG-based emotion recognition for music therapy. J. Multimodal User Interfaces **5**(1–2), 27–35 (2011)
19. Kothe, C.A., Makeig, S.: Estimation of task workload from EEG data: new and current tools and perspectives. In: International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6547–6551 (2011)
20. Chanel, G., Kierkels, J.J.M., et al.: Short-term emotion assessment in a recall paradigm. Int. J. Hum.-Comput. Stud. **67**(8), 607–627 (2009)
21. Heraz, A., Frasson, C.: Predicting the three major dimensions of the learner's emotions from brainwaves. Int. J. Comput. Sci. **2**(3), 187–193 (2008)
22. Zheng, W.L., Zhu, J.Y., et al.: Identifying stable patterns over time for emotion recognition from EEG. IEEE Trans. Affect. Comput. **PP**(99), 1 (2016)
23. Li, K., Li, X., et al.: Affective state recognition from EEG with deep belief networks. In: IEEE International Conference on Bioinformatics and Biomedicine, pp. 305–310 (2013)
24. Mao, Z.J., Yao, W.X, et al.: EEG-based biometric identification with deep learning. In: International IEEE EMBS Conference on Neural Engineering, pp. 25–28 (2017)
25. Li, X., Song, D., et al.: Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. In: IEEE International Conference on Bioinformatics and Biomedicine, pp. 352–359 (2017)
26. Fawzi, A., Samulowitz, H., et al.: Adaptive data augmentation for image classification. In: IEEE International Conference on Image Processing, pp. 3688–3692 (2016)
27. Hadjidimitriou, S.K., Hadjileontiadis, L.J.: Toward an EEG-based recognition of music liking using time-frequency analysis. IEEE Trans. Biomed. Eng. **59**(12), 3498 (2012)
28. Tibdewal, M.N., Tale, S.A.: Multichannel detection of epilepsy using SVM classifier on EEG signal. In: International Conference on Computing Communication Control and Automation, pp. 1–6 (2016)
29. Chong, Y.S., Mokhtar, N., et al.: Automated classification and removal of EEG artifacts with SVM and wavelet-ICA. IEEE J. Biomed. Health Inform. **PP**(99), 1 (2017)
30. He, K., Zhang, X., et al.: Deep residual learning for image recognition. In: Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2015)
31. Soleymani, M., Lichtenauer, J., et al.: A multimodal database for affect recognition and implicit tagging. IEEE Trans. Affect. Comput. **3**(1), 42–55 (2014)
32. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**(3), 27 (2011)
33. Vedaldi, A., Lenc, K.: MatConvNet: convolutional neural networks for matlab. In: ACM International Conference on Multimedia, pp. 689–692 (2015)
34. Zhang, G., Ge, H.: Prediction of xylanase optimal temperature by support vector regression. Electron. J. Biotechnol. **15**(1), 7 (2012)