Emotion Recognition of EEG Signals Based on Channel Attention Convolution Neural Network

Xiu Zhang, Xun Pei, and Xin Zhang(✉)

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
ecemark@tjnu.edu.cn

Abstract. Electroencephalography (EEG) is more likely to respond to emotional changes compared to facial expressions and voice because it is not subject to external interference and is not easily disguised. By classifying the emotions of EEG signals, it can provide an aid for future treatment of depression, epilepsy and other diseases. Therefore, this article uses the publicly available EEG emotion dataset and uses a ratio of 80%–20% to divide the training set and test set. The features of EEG data are extracted using Fast fourier transform (FFT), and the Convolutional neural network (CNN) in deep learning is used as the basic premise to incorporate the channel attention mechanism. An EEG emotion recognition model combining CNN and channel attention mechanism is designed, which includes three convolutional layers, three channel attention blocks, three maximum pooling layers, three fully connected layers, a dense layer and a softmax layer, and batch normalization is used to suppress the overfitting of the model. Experimental results show that the sentiment recognition accuracy of the DEAP dataset reaches 90%, which achieves a significant improvement over existing emotion recognition models.

Keywords: Convolutional neural networks · Deep learning · EEG · Emotion recognition

1 Introduction

Nowadays, as the research of artificial intelligence continues to advance, more and more researchers are focusing on human emotional intelligence. As an advanced stage in the field of artificial intelligence, emotion analysis is an important research object in the fields of human-computer interaction and anthropomorphic control theory [1]. It has been found that positive emotions reflect pleasant psychological states and contribute to human health and attitudes [2]. With the growth of information through social channels, emotion computing emerged under the need for deeper understanding and rational use of emotions [3, 4]. Emotion recognition research has not only improved human-computer interaction systems, but is also an important topic in clinical research, such as the treatment of psychological disorders and epilepsy [5, 6]. Therefore, emotion recognition based on

Q. Liang et al. (Eds.): CSPS 2022, LNEE 874, pp. 90–97, 2023.
https://doi.org/10.1007/978-981-99-2362-5_12

EEG signals has received wide attention from an increasing number of researchers in the field of affective computing.

Deep learning has achieved remarkable results in the fields of computer vision and natural language processing, and researchers have started to apply deep learning to the field of EEG emotion recognition. The literature [7] proposed a convolutional recurrent neural network combining CNN and Recurrent neural network (RNN), and its binary recognition rate in DEAP dataset [8] was 72.06% and 74.12% in validity and arousal dimension, respectively. The literature [9] compared the performance of deep learning models, i.e., Long Short-Term Memory (LSTM) and CNN on training and test set splits of 80–20 and 75–25. The EEG emotion recognition accuracy of LSTM and CNN was 88.6% and 87.72%, respectively. The literature [10] proposed the application of Hierarchical convolutional neural networks (HCNN) to classify emotions as positive, neutral and negative. Their method preserves the spatial topology of the electrodes. The EEG signal is converted into an image by interpolation of the electrode positions obtained from the 2D projection technique and the frequency bands of the EEG signal. Their results have an accuracy of 86.2% in the β-band and 88.2% in the γ-band. In the literature [11] CNN are built based on convolutional operations to extract information features by fusing spatial and channel information within the local sensory field. They focused on channel relations and proposed a new architectural unit, which we call "Squeeze-and-Excitation" (SE) block, to improve the performance of the CNN model significantly.

Some approaches also use brain activity for subject identification or authentication. For example, a cascaded deep learning using a combination of CNN and RNN was proposed in the literature [12] and evaluated on the sentiment dataset DEAP. The results show that CNN-Gate recurrent unit (GRU) and CNN-LSTM can perform person recognition from different sentiment states with an average correct recognition rate of 99.90–100%. The literature [13] proposed imagined speech for user recognition, achieving 99.76% accuracy in six different subjects.

The use of brain activity for sleep monitoring is an emerging field and can be very helpful in the adjunctive treatment of apnea. The literature [14] proposed a single-channel EEG sleep staging model, SleepStageNet, which extracts sleep EEG features through a multiscale CNN and then infers the type of sleep stage by capturing contextual information between adjacent stages using RNN and Conditional random fields (CRF). Their model was validated on Obstructive sleep apnea (OSA) patients with mean accuracies of (F4-M1, 0.80) and (F4-M1, 0.67). To address the problem that existing methods rarely consider local features in feature extraction and cannot distinguish the importance of critical and non-critical local features.

2 DEAP Dataset

The DEAP dataset [8] is multichannel data collected experimentally by Koelstra et al. to study human emotional states. The dataset is based on physiological signals generated by music video material evoked stimuli, recorded by 32 subjects watching 40 min of music videos (1 min per music video). The original 512 Hz EEG signal was downsampled to remove artifacts and pre-processed to 128 Hz. Of the 40 physiological signal channels collected, the first 32 channels were collected as EEG signals. EEG channels were selected according to the international system of 32 channels, the locations of which are shown in Fig. 1.

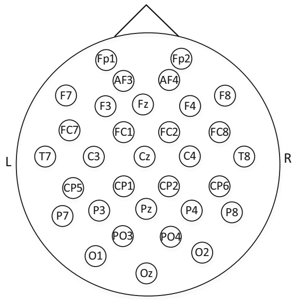

Fig. 1. 32 electrode channel positions

3 Method

3.1 Feature Extraction

FFT is performed on the feature extraction to reduce the dimensionality, thus having a faster training speed as well as better accuracy. These extracted features include five frequency bands. Delta-δ (1–4 Hz), theta-θ (4–8 Hz), alpha-α (8–14 Hz), beta-β (14–31 Hz), and gamma-γ (31–50 Hz). The FFT used in this paper extracts by time, thus changing a signal from the time domain to the frequency domain.

The DFT is given by:

$$X(k) = DFT[x(n)] = \sum_{n=0}^{N-1} x(n) W_N^{kn}, k = 0, 1, \ldots, N-1 \tag{1}$$

in which,

$$W_N = e^{-j\frac{2\pi}{N}} \tag{2}$$

Using the periodicity and symmetry of W_N^{kn} in DFT, the whole DFT calculation becomes a series of iterative operations, which can greatly improve the operation process and the number of operations, which is the FFT algorithm.

As shown in Table 1, 14 channels and 5 frequency bands are selected for our model. The window size chosen is 256, averaging the band power over 2 s. The step size is 16, which means that every 0.125 s the update.

3.2 Channel Attention Mechanism

Adding attentional mechanisms to the channel dimension, the traditional channel selection applied to EEG emotion recognition usually manually picks the EEG channels

Table 1. Parameter settings of FFT

Parameters	Values
channel	1, 2, 3, 4, 6, 11, 13, 17, 19, 20, 21, 25, 29, 31
Bands	4, 8, 12, 16, 25, 45
Window size	256
Step size	16

associated with emotional information. However, this approach loses some important EEG information and requires the role of more manual factors. In this paper, we combine CNN with channel attention mechanism, which can automatically assign weights to the feature channels obtained by CNN extraction. That is, another new neural network is used to get to the importance of each channel, and then it is used to assign a weight value to each feature so that the neural network focuses on certain feature channels. Boosting feature channels that are useful for the current task and suppressing feature channels that are less useful for the current task.

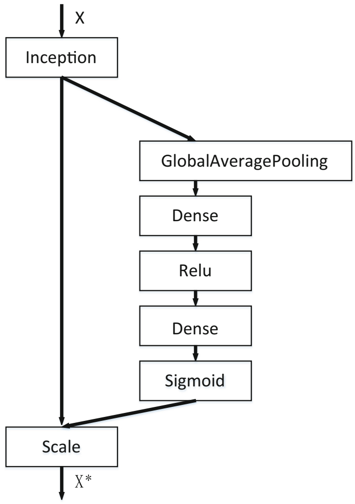

Fig. 2. Block diagram of channel attention mechanism

As shown in Fig. 2, the right-hand operation is to generate a weight value for each feature channel, where correlations between channels are constructed through two fully connected layers, with the same number of output weight values as the number of input channels.

Scale operation: The normalized weights obtained earlier are weighted to the features of each channel, and the Scale operation uses a multiplication method, multiplying the weight coefficients channel by channel.

3.3 Model Architecture

As shown in the model architecture in Fig. 3 below, there are three conv1D layers, three channel attention blocks, three fully connected layers and one dense layer, and finally softmax activation with 10 classes.

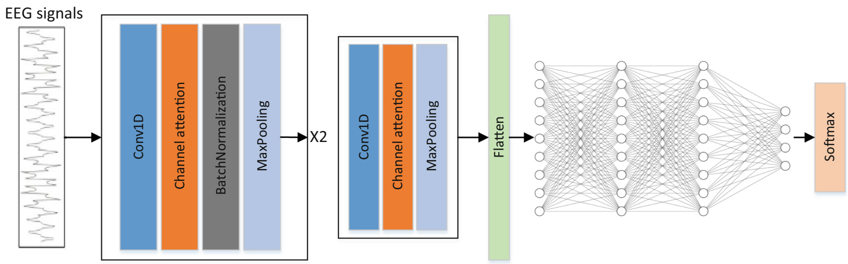

Fig. 3. Block diagram of EEG emotion recognition based on one-dimensional convolutional neural network and channel attention mechanism

The first convolutional layer uses Rectified linear unit (ReLU) as the activation function and 128 filters with a kernel size of 3. The input to the first layer passed to conv1D is the shape (70, 1) and uses the same padding in steps of 1. Afterwards, it is passed to the channel attention layer, which generates weights for each channel of the input and then weights the features of each channel.

Normalize the output of the first layer, i.e., use a batch normalization layer to make its mean 0 and variance 1. The next layer is maxpooling with a pool size of 2 and a window size of 2. Take the maximum value to downsample the input.

The next convolutional layer is the same as the first layer, followed by a batch normalization and maxpooling layer. The shape is then spread to form a one-dimensional layer and fed to a fully connected layer consisting of 64 neurons and hyperbolic tangent (tanh) as the activation function. A dropout on the output of the dense layer is used to reduce overfitting of the network with a dropout rate of 0.2. This is followed by a dense layer consisting of 32 neurons with tanh activation and a dropout rate of 0.2, and another consisting of 16 neurons with ReLU as the activation function and a dropout rate of 0.2.

Finally, a dense layer of 10 neurons (with an activation function of softmax) is used to give the output of the network. The softmax formula is shown in (3).

$$Softmax(x) = \frac{e^{x_i}}{\sum_i e^{x_i}} \tag{3}$$

The softmax function is different from the normal max function: while the max function only outputs the maximum value, the softmax function ensures that smaller values have a smaller probability and are not discarded outright. This means that the various probabilities it obtains are related to each other. Therefore it is more suitable for the multiclassification problem in this paper.

4 Overall Performance

The accuracy, loss and confusion matrices of the model are shown in Fig. 4, 5 and 6, respectively.

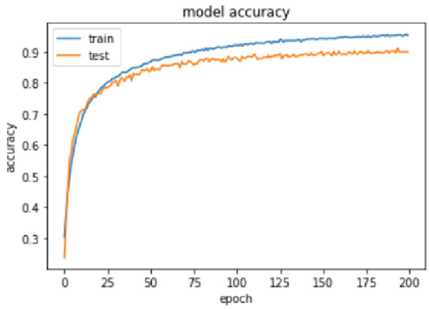

Fig. 4. Performance of the model

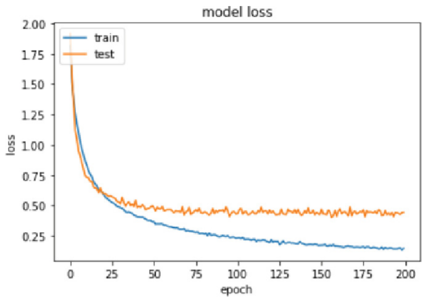

Fig. 5. Loss of model

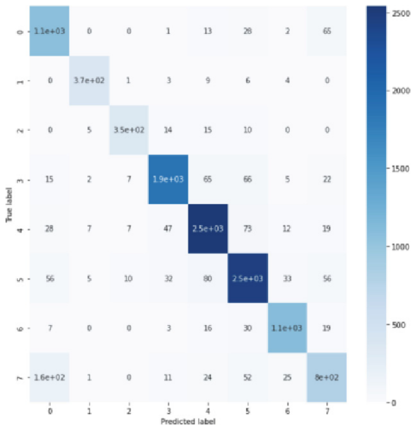

Fig. 6. Confusion matrix for Model

Combining the above model validation results, the one-dimensional channel attention CNN model was compared and analyzed with other models, and the results are shown in Table 2. it can be seen that the accuracy of our model is better than most EEG emotion recognition models.

Table 2. Comparison with the existing studies

Method	Accuracy
C-RNN [15]	74.12%
CNN [16]	81.41%
RNN [17]	84.00%
DBN [18]	88.33%
Ours	**90.00%**

5 Conclusion

In this paper, an EEG emotion recognition method based on channel attention convolution is proposed. Experimental results on the DEAP dataset show that the CNN incorporating the channel attention mechanism has a significant performance improvement over the traditional CNN model, with a test accuracy of 90%. The improved model proposed in this paper outperforms most current schemes. In future work, we intend to extract more EEG signal features and fuse them to further improve the performance of EEG emotion recognition through feature complementation.

References

1. Zhu, Y.H., Wei, J.H., Mao, J.X.: A review of artificial emotion research. J. Jiangnan Univ. (Nat. Sci. Ed.) **11**(04), 497–504 (2012)
2. Pressman, S.D., Cohen, S.: Does positive affect influence health? Psychol Bull. **131**(6), 925–971 (2005)
3. Picard, R.W.: Affective Computing. MIT Press, Cambridge (1997)
4. Calvo, R.A., D'Mello, S.: Affect detection: an interdisciplinary review of models, methods, and their applications. IEEE Trans. Affect. Comput. **1**(1), 18–37 (2010)
5. Ozdemir, M.A., Degirmenci, M., Guren, O., Akan, A.: EEG based emotional state estimation using 2-D deep learning technique. In: 2019 Medical Technologies Congress (TIPTEKNO), pp. 1–4. IEEE (2019). https://doi.org/10.1109/tiptekno.2019.8895158
6. Othman, M., Wahab, A., Karim, I., Dzulkifli, M.A., Alshaikli, I.F.T.: EEG emotion recognition based on the dimensional models of emotions. Procedia Soc Behav Sci **97**, 30–37 (2013)
7. Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., Hu, B.: Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. In: 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 352–359 (2016). https://doi.org/10.1109/BIBM.2016.7822545
8. Koelstra, S., et al.: DEAP: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2012). https://doi.org/10.1109/T-AFFC.2011.15
9. Acharya, D., et al.: Multi-class emotion classification using EEG signals. In: Garg, D., Wong, K., Sarangapani, J., Gupta, S.K. (eds.) IACC 2020. CCIS, vol. 1367, pp. 474–491. Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-0401-0_38
10. Kuanar, S., Athitsos, V., Pradhan, N., Mishra, A., Rao, K.R.: Cognitive analysis of working memory load from EEG, by a deeprecurrent neural network. In: 2018 IEEE International Conferenceon Acoustics, Speech and Signal Processing (ICASSP), pp. 2576–2580. IEEE (2018)

11. Jie, H., Shen, L., Albanie, S., Sun, G., Enhua, W.: Squeeze-and-excitation networks. IEEE Trans. Pattern Anal. Mach. Intell. **42**(8), 2011–2023 (2020). https://doi.org/10.1109/TPAMI.2019.2913372
12. Wilaiprasitporn, T., Ditthapron, A., Matchaparn, K., Tongbuasirilai, T., Banluesombatkul, N., Chuangsuwanich, E.: Affective EEG-based person identification using the deep learning approach. IEEE Trans. Cogn. Dev. Syst. **12**(3), 486–496 (2020). https://doi.org/10.1109/TCDS.2019.2924648
13. Brigham, K., Kumar, B.V. K. V.: Subject identification from electroencephalogram (EEG) signals during imagined speech. In: 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–8 (2010). https://doi.org/10.1109/BTAS.2010.5634515
14. Chen, K., Zhang, C., Ma, J., Wang, G., Zhang, J.: Sleep staging from single-channel EEG with multi-scale feature and contextual information. Sleep Breath. **23**(4), 1159–1167 (2019). https://doi.org/10.1007/s11325-019-01789-4
15. Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., Hu, B.: Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. In: IEEE International Conference on Bioinformatics and Biomedical (BIBM), Guang-dong, China, pp. 352–359 (2016)
16. Tripathi, S., Acharya, S., Sharma, R.D., Mittal, S., Bhattacharya, S.: Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset. In: AAAI Conference on Innovative Applications, San Francisco, California, USA, pp. 4746–4752 (2017)
17. Spampinato, C., Palazzo, S., Kavasidis, I., Giordano, D., Souly, N., Shah, M.: Deep learning human mind for automated visual classification. In: Proceedings - IEEE Conference on Computer Vision Pattern Recognitin, CVPR, pp. 4503–4511 (2017)
18. Xu, H., Plataniotis, K.N.: Affective states classification using EEG and semi-supervised deep learning approaches. In: IEEE Workshop Multimedia and Signal Processing, Montreal, Canada, USA, pp. 1–6 (2016)