Emotion Prediction of EEG Signals based on 1D Convolutional Neural Network

Jiaman Gao1, Chengyi Yang1,2, Feng Liu2,3,* and Jiayin Qi2,*

1 School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai, China
2 Institute of Artificial Intelligence and Change Management, Shanghai University of International Business and Economics, Shanghai, China
3 School of Computer Science and Technology, East China Normal University, Shanghai, China
* E-mail: lsttoy@163.com; ai@suibe.edu.cn

Abstract. Artificial intelligence has been widely used in the field of biological signal recognition. However, most researches use deep learning to classify emotions, which has limitations in its application in the medical field. To this end, this paper proposes a one-dimensional convolutional neural network (1D-CNN) model for regression tasks. After we standardize, transform and slice the data, we divide the training set, validation set, and test set at a ratio of 8:1:1, and feed the data into the neural network for training to achieve emotion prediction. Experiments on the DEAP dataset show that the model we built has good performance for emotion prediction, which provides new insights for the medical field. The source codes are available at https://github.com/gjm-web/1D-CNN.

1. Introduction
With the improvement of human living standards and the progress of medical technology, mental health problems have attracted increasing attention. Emotions have a great influence on behavioral expression. In daily life, people generally express their emotions through non-physiological signals such as body and language; but some emotions cannot be expressed through non-physiological signals. However, physiological signals can directly reflect show people’s emotions, such as electrocardiograms, brain electrical signals, etc. Therefore, predicting emotions based on physiological signals has practical significance.

In the early days, the research on electroencephalography (EEG) signals mainly focused on machine learning algorithms based on signal processing methods and statistical theory. Dornhege G et al. [1] used feature combination and multiclass paradigms to enhance the information gain of multi-channel EEG. Based on the recursive feature elimination of Support Vector Machine (SVM), the most discriminating feature can be selected during the feature selection process of EEG signal analysis [2].

As deep learning [3] has achieved remarkable results in the field of computer vision, this method has gradually been used in pattern recognition of brain signals. In the early days, the multi-layer perceptron (MLP) method was introduced into the field of brain signal recognition as a feature extraction method [4], and it made a breakthrough in the classification of medical patients. MLP has the disadvantages of high computational cost and large amount of parameters. However, CNN can reduce the amount of parameters through weight sharing and convolution of data, which has attracted widespread attention.
With the proposal of a 3D-CNN model, it can learn useful features from the data and automatically process multi-channel data, which contributes to the automatic detection of epileptic seizures [5]. The Multi-task deep recurrent neural network proposed by Chen et al. [6] can make full use of the correlation between EEG signals to recognize human intentions of EEG signals with different frequencies. Since the traditional DNN network cannot represent the relational data of fixed points and edges, Graph Neural Networks (GNN) emerged to solve the problem of graph data representation. Demir et al. [7] extract features from EEG based on GNN for classification tasks, which is essential to reduce computational cost. However, GNN cannot strengthen the network by deepening the number of network layers, thus Venuto et al. [8] combined the symbolized EEG signal with the automatic coding convolutional neural network and proposed the P300 detector, which can maintain high recognition accuracy while maintaining high recognition accuracy. Improve the information conversion rate of the brain-computer interface.

The proposal of 1D-CNN provides a powerful tool for feature extraction of time series, which has extended to the application of audio signal recognition [9], behavior detection [10] and other fields. However, we note that most recent studies on emotion recognition through EEG signals have modeled it as a classification task, and the rough classification undoubtedly limits the comparability between different emotions. In this paper, we designed the regression task of identifying emotion through EEG signals based on 1D-CNN model, hoping to provide more fine-grained score predictions for human emotion. The main contributions are as follows:

1) We establish a regression task and recognize the level of emotions through EEG, which provides new insights into the field of biomedical signal processing.
2) We design a neural network framework containing 40 1D convolution kernels, and successfully processed high-dimensional tensor data formed by multiple environmental variables.
3) We conduct the experiments on large-scale public dataset DEAP, and the results show that our method designed in this paper has achieved good results in emotion prediction.

Our methodology that designing a large-scale neural network framework integrating 40 filters for the regression task of EEG signal analysis also opens up the new horizons for the direction of biological signal processing in the medical field.

2. Related work
With the continuous development of computer technology, the brain-computer interface (BCI) system has become an important part of computer research. A hybrid detection system based on decision tree classifiers and fast Fourier transform (FFT) is used to assist the diagnosis system in detecting epileptic seizures in EEG signals [11]. Mporas et al. [12] proposed SVM-HMM to detect sleep spindles in EEG signals.

Traditional machine learning methods involve many complicated manual steps, such as feature selection and classification. In order to overcome the limitations, many scholars have proposed methods based on deep learning. The fusion method of MCNN and CCNN proposed by Amin S U et al. [13] is superior to machine learning technology for EEG classification. Sheykhivand S et al. [14] proposed that the CNN-LSTM model can be applied to develop human-machine interface systems. Khare SK et al. [15] proposed the SPWVD-CNN model to help doctors detect and diagnose mental illness.

3. Methodology
3.1. Problem description
EEG signal processing is a major branch of the field of biometrics. In early stage, EEG data analysis is considered as one-dimensional discrete signal processing task. With the complexity of experimental settings and environmental variables, EEG signals are gradually becoming complex, which undoubtedly increases the difficulty of medical analysis in this field. In this case, the steps of data preprocessing have increasingly become an inevitable part.
In order to match the EEG data with the input size of the proposed neural network, we added a dimension on the EEG signal data and performed Z-score standardization processing. We also performed 0-1 standardization processing on the label data. The standardized dataset was divided into the original data of size \(40 \times 32 \times 40 \times 8064\) according to the subjects. The training set and the test set were divided into \(40 \times 28 \times 40 \times 8064\) and \(40 \times 4 \times 40 \times 8064\) according to the ratio of 9:1. Then we divided the validation set as \(40 \times 3 \times 40 \times 8064\) from the training set at a ratio of 9:1, and the shape of the training set finally became \(40 \times 25 \times 40 \times 8064\). Among them, 40 in the first dimension represents the number of channels, and 25, 3, and 4 in the second dimension of the training set, validation set, and test set represent the number of subjects in each, namely, the number of samples. In order to explore whether music will affect the EEG signal of subjects, we designed a different filter (convolution kernel) for each music sample, namely, the quantity of music equals to the quantity of filters in the third dimension. The fourth dimension 8064 is the sample size of a batch.

3.2. Convolution Neural Network

The neural network algorithm CNN extracts data features through the alternate use of convolutional layers and pooling layers. When a one-dimensional convolution kernel is used, it can better capture the time series features of the data; therefore, this paper uses 1D-CNN to extract data features and the subject’s EEG signals are used to predict emotions.

The operation process of the convolutional layer can be described as follows:

\[ x_j^l = f \left( \sum_{i \in M_j} x_j^{l-1} \ast W_{ij}^l + b_j^l \right) \]

Where \(x_j^{l-1}\) is the feature vector of the \(j\)-th convolution kernel in the \((l-1)\)-th layer, and \(W_{ij}^l\) denotes the weight corresponding to the \(j\)-th convolution kernel of the \(l\)-th layer. \(b_j^l\) represents the bias term of the \(j\)-th convolution kernel in the \(l\)-th layer, and \(M_j\) represents the receptive field of the neuron. \(f(\cdot)\) is a non-linear activation function, and we select ReLU as the activation function in this paper.

In order to solve the problem of internal covariate shift, we normalized each layer of neural network, and the normalization operation can be described as follows:

\[ y = \frac{x - E(x)}{\sqrt{Var(x) + \epsilon}} \ast \gamma + \beta \]

Where \(\epsilon\) is a small constant whose function is to prevent the denominator from being 0. \(\gamma\) is a learnable parameter vector normally set to 1, while \(\beta\) is the learnable parameter vector normally set to 0. MSE loss function was applied in this paper to evaluate the neural network and conduct gradient descent operation. The specific formula can be expressed as follows:

\[ Loss = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 \]

Where \(y_i\) is the real data and \(\hat{y}_i\) is the predicted data. We developed a neural network framework containing 40 1D-CNN models in this paper, and designed structure is shown in Figure 1. It consists of an input layer, four convolutional layers, a fully connected layer and an output layer. The convolutional layer C1 uses a convolution kernel of size 5, and its input is a segment of 8064 sample points of the EEG signal. The convolutional layers C2, C3, and C4 all use a convolution kernel of size 3, and the finally obtained sample points are sent to the fully connected layer to output 4 neurons to predict the emotion of the subject.
4. Experiment

4.1. DEAP
In this paper, we conducted experiments on the large-scale public dataset called DEAP, which consists of two parts. One part records various physiological signals of 32 subjects, including 8 channels of peripheral physiological signals of 32 channels of EEG signals; these physiological signals are 40 bands viewed by each subject. Produced after music videos with different emotional tendencies. Another part of the data is the score of the subjects' self-assessment. After each video, the subjects used a continuous 9-point system to score valence, arousal, dominance, and degree of affection.

Before feeding the preprocessed data into the neural network, we first checked the EEG signals of the first channel and the second channel after the first subject watched the first music video, as shown in Figure 2. It can be seen from Figure 2 that after the first subject watched the first music video, the value range of the original EEG signal was from -4 to 4, and it was evenly distributed on both sides of 0.

![Figure 1. Structure of designed 1D-CNN model](image1)

![Figure 2. Distribution of raw data of EEG signals](image2)
4.2. Parameter adjustment
We fed the training set data into the neural network integrated with a large number of filters to learn the mapping from EEG signals to the score prediction of valence, arousal, dominance, and degree of affection. The test set was input into the trained model for emotion prediction, and the loss value used to evaluate the performance of the model was obtained. In order to show the fitting ability and generalization ability of the network constructed in this paper more clearly, we selected MAE and MSE as indicators to evaluate the network, as well as adjust the hyperparameters in the network to find out whether it has an impact on the test results. To express the experimental parameter settings concisely, we denoted the number of neurons in the first convolution layer as $N$.

Figure 3 shows the change of loss function of the neural network designed in this paper during training process. The horizontal ordinates represent the number of training rounds and the vertical axis represent the loss value. It can be seen that with the increase of training epochs, the MAE and MSE loss can drop to the level close to 0 when the value of $N$ is different, which implies that the neural network model designed for emotion score prediction in this paper has good fitting ability. In addition, we note that with the increase of the number $n$ of neurons in the first layer, the loss value of the neural network can reach a stable level at a faster speed, which means that the number of neurons in the first layer has a positive impact on the learning ability of the whole network.

![Figure 3. Changes in loss function in the training process](image)

(a) N=32 (b) N=128 (c) N=1024

5. Test results of different music
In order to analyze the influence of hyperparameters on the network, we carried out comparative experiments under different super parameter settings, and the results of test loss are shown in Figure 4. It can be seen that the discrimination of the model to music increases as $N$ increases. It can be seen from the results that when $N=1024$, the MSE of the 1th, 19th, 22th and 25th music is significantly smaller than that of other music, which implies these three pieces of music for emotion prediction will achieve better results.

![Figure 4. Test results of different music](image)

(a) N=32 (b) N=128 (c) N=1024

5. Conclusion
In this paper, we established a regression model for the task of sentiment prediction via EEG data, which provided new insights in the field of biological signals processing. We proposed a large neural network
framework with 40 filters integrated together, which could effectively process high-dimensional tensor data formed by multiple environmental variables. Experiments on the DEAP dataset showed that our method could effectively predict the change of emotion level. In the future, we plan to devote ourselves to the study of optimization problems when the 1D-CNN model is used in regression tasks, providing technological solutions for more difficult problems of medical data analysis.

References
[1] Dornhege, G., Blankertz, B., Curio, G., Muller, K.-R. (2004) Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. IEEE transactions on bio-medical engineering, 51:993-1002.
[2] Gysels, E., Renevey, P., Celka, P. (2005) SVM-based recursive feature elimination to compare phase synchronization computed from broadband and narrowband EEG signals in Brain–Computer Interfaces. Signal Processing, 85:2178-2189.
[3] Lecun, Y., Bottou, L., Bengio, Y. (1998) Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86:2278-2324.
[4] Sukanesh, R., Harikumar, R. (2007) A Comparison of Genetic Algorithm & Neural Network (MLP) In Patient Specific Classification of Epilepsy Risk Levels from EEG Signals. Engineering Letters, 14: 96-104.
[5] Wei, X., Zhou, L., Chen, Z., Zhang L.,Zhou, Y., (2018) Automatic seizure detection using three-dimensional CNN based on multi-channel EEG. Bmc Medical Informatics & Decision Making, 18: 71-80.
[6] Chen, W., Wang, S., Zhang, X., Yao, L., Yue, L., Qian, B., Li, X. (2018) EEG-based motion intention recognition via multi-task RNNs. In: Proceedings of the 2018 SIAM International Conference on Data Mining. San Diego. PP. 279-287.
[7] Demir, A., Koike-Akino, T., Wang, Y., Haruna, M., Erdogmus, D. (2021) EEG-GNN: Graph Neural Networks for Classification of Electroencephalogram (EEG) Signals. https://arxiv.org/abs/2106.09135.
[8] Venuto, D.D., Mezzina, G. (2021) A Single-Trial P300 Detector Based on Symbolized EEG and Autoencoded-(1D)CNN to Improve ITR Performance in BCIs. Sensors, 21:3961.
[9] Chowdhury, A., Ross, A. (2020) Fusing MFCC and LPC Features Using 1D Triplet CNN for Speaker Recognition in Severely Degraded Audio Signals. IEEE transactions on information forensics and security, 15:1616-1629.
[10] J Trelinski, Kwolek B. (2021) Embedded Features for 1D CNN-based Action Recognition on Depth Maps. In: 16th International Conference on Computer Vision Theory and Applications. Montreal. PP. 536-543.
[11] Polat, K., Güne, S. (2007) Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Applied Mathematics and Computation, 187:1017-1026.
[12] Mporas, I., Korvesis, P., Zacharakis, E I., Megalooikonomou, V. (2013) Sleep Spindle Detection in EEG Signals Combining HMMs and SVMs. Springer, Berlin, Heidelberg, (2):138-145.
[13] Amin, S.U., Alsulaiman, M., Muhammad, G., Amine, M., Hossain, M.S. (2019) Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion. Future Generation Computer Systems, 101:542-554.
[14] Sheykhand, S., Mousavi, Z., Rezaei, T.Y., Farzamnia, A. (2020) Recognizing Emotions Evoked by Music using CNN-LSTM Networks on EEG signals. IEEE Access, 8:139332-139345.
[15] Khare, S.K., Bajaj, V., Acharya, U.R., (2021) SPWVD-CNN for automated detection of Schizophrenia patients using EEG signals. IEEE Transactions on Instrumentation and Measurement, 70:1-9.