Emotion recognition based on the human's brain

Liqi Tang

School of Computer Science, University of Nottingham Malaysia, Jalan Broga, Semenyih Selangor Darul Ehsan 43500, Malaysia

Liqitang127@gmail.com

Abstract. There are several different algorithms to train models in machine learning and for purpose of enhancing the accuracy of the models for emotion recognition systems by using EEG datasets, we can compare different algorithms and methods. To experience more intuitively the different impacts caused by different algorithms and methods, 'Emotion classification by applying 'Convolutional Neural Networks, Sparse Auto-encoder, Deep Neural Network', and 'ZTW-based epoch selection algorithm', these three models of emotion recognition will be mentioned and compared in this essay to find out a way to improve the accuracy of the emotion recognition models based on EEG datasets, which also include the process and final results of these models.

Keywords: emotion recognition, EEG, human Brain-based, artificial intelligence, machine learning.

1. Introduction

Due to the coronavirus disease 2019 and the fast pace of life, almost all people nowadays live under enormous life pressure, such as students who are under great study pressure, adults who are under mortgage and car loan pressure, and even the rising cost of family living, which is creating an unemployment situation. According to the World Health Organization (WHO), depression has become a prevalent condition globally, affecting over 350 million people, with the number of sufferers increasing by nearly 18% in the last decade. Each year, it is estimated that 5% of individuals worldwide suffer from depression. Since the COVID-19 pandemic, the prevalence of depression has risen dramatically. According to a World Health Organization (WHO) research brief issued in March 2022, the COVID-19 pandemic has resulted in a 28% rise in the global incidence of depression [1].

Emotions are ultimate what shape and define the typical human experience. It is the main means of communication in interpersonal interactions and the inspiration for both the noblest and most monstrous actions of people. In ways that are neither well recognized nor have been properly examined, emotions significantly impact reason and aid in anchoring people's beliefs [2].

Recent technology advancements have enabled humans to engage with computers in previously inconceivable ways. New types of human-computer interaction, for instance, speech, menu driven and graphical user interface (GUI), are emerging beyond the restrictions of the keyboard and mouse. Despite significant advances, one critical component of natural contact remains missing: emotion, which is produced during the process of seeing the external environment have a direct impact on daily activities, for example, social interaction, work productivity, and sound in body and mind [3]. Emotions are

produced by how people interpret their environment, and they have a direct impact on how people live their daily lives [4].

This essay will primarily discuss emotion recognition based on the human brain using EEG datasets and the advantages and disadvantages of two existing models of trained emotion recognition systems that can automatically recognize human emotions using EEG signals.

2. Electroencephalogram

An electroencephalogram is a recording of brain activity (EEG). In this painless test, small sensors are put on the scalp to detect electrical impulses generated by the brain. The pulses are recorded by a machine and then checked by a doctor. Clinical neurophysiologists typically utilize it to do EEG examinations during brief hospital visits [5].

The electroencephalography (EEG) signature of brain activity is typically the integral of the active potential induced by the brain at each moment with varying latency durations and populations. Brain activity modeling can be more difficult than mimicking the functions of other organs. More study in this research area is now feasible because Time-frequency (TF) domain analysis is used to detect neonatal epilepsy. EEG signals may be regarded as the output of the nonlinear system that can be precisely characterized. The change in signal distribution over time can be used to quantify non-stationary signals. The synchronous compression wavelet transform is used to enhance the TF spectrum generated by the wavelet transform. Empirical mode decomposition (EMD) is an adaptive spatiotemporal analysis technique for non-stationary nonlinear time series [6].

EEG datasets are typically seen to be much more trustworthy since they are based on unconscious changes in the body regulated by the sympathetic nervous system which makes them harder to manipulate. Recognizing emotions from brain signals is an important step toward developing emotional intelligence since EEG readings may directly detect brain dynamics in response to a variety of emotional states [7].

3. Emotion classification by using Convolutional Neural Network, Sparse Autoencoder and Deep Neural Network

In this machine learning model, Sparse Autoencoder (SAE) is the first place that receives the features that Convolutional Neural Network (CNN) has extracted for encoding and decoding. Then, for classification tasks, the Deep Neural Network (DNN) uses data characterized by reduced redundancy as input. Use DEAP and SEED's shared data sets for testing, where the greatest recognition accuracy for valence and arousal for the DEAP dataset is 89.49% and 92.86%, respectively. However, the best accuracy of emotion recognition for the SEED dataset is 96.77% [8].

To achieve good mood classification performance, a network model incorporating CNN, SAE, and DNN is presented in this network to transform EEG time series into 2D pictures. The EEG signals are grouped into various different bands. EEG data is processed to derive two-dimensional features which are on account of frequency, time, and location data. The convolution layer of CNN is then trained for additional feature extraction. Whereas DNN is used for classification, and SAE is used to reconstruct the data acquired from the convolutional layer. The neural network model makes use of the sparsity of SAE and the convolutional layer of CNN, resulting in good classification accuracy. The proposed method's workflow is outlined in Figure 1. The final bands of EEG signals classification results are retrieved after training and testing the pre-processed original EEG data and features are drawn by this model [8].
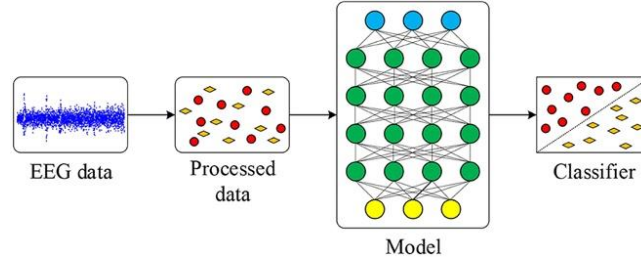
Figure 1. The application of emotion classification program in this work [8].

During the preprocessing of the raw EEG data, we can easily find that the EEG signal can be easily disrupted by lots of factors during the sampling process, such as influence from the environment and human emotional swings. In order to reduce the noise in the row EEG signal which significantly influences the result of the experiments and the required patterns of the brain, and also results in the training and testing accuracy of this network:

The CNN structure in this network consists of two convolution layers and one maximum pooling layer. Each convolution layer has a connection to drop out. An encoding layer, a concealment layer, and a decoding layer make up SAE. There are three completely connected classification layers in DNN. The input of SAE is also the output of the max-pooling layer given PCC and other characteristics of the network input. Finally, the SAE output is categorized as the DNN input.

All samples and characteristics are used in the training procedure to train CNN with a fully connected output layer. After training, the output layer is removed. Then, feed the features into the trained CNN to get the max-pooling layer's output. The output one-dimensional values are flattened and assigned to the SAE input. The data is recreated once the SAE has been unsupervised and evaluated. The recovered data was separated into DNN for training and testing. In order to train CNN and SAE independently. As a result, training in CNN and SAE may be thought of as progressive feature extraction prior to categorizing data. Note that the dataset DNN used for classification is not the same as the one fully connected output layer extracted from CNN in the first step. DNN will never train until being fed SAE output.

To fairly test the performance of the proposed network, it uses another CNN with the same characteristics and structure as the proposed network for comparison. The accuracy rate of this CNN does not increase as the number of layers increases, resulting in overfitting. The features employed by CNN in this experiment are separated into 80% training features and 20% test features.
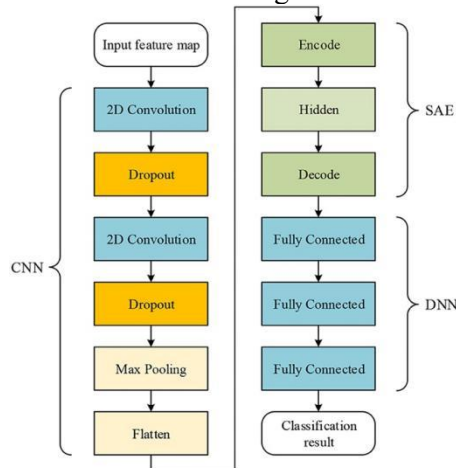

Figure 2. The structure of this network [8].

Three distinct features are derived from the proposed network for classification purposes. Results indicated that by utilizing PCC-based characteristics, the proposed network's average recognition accuracy could increase to 96.77% for SEED and DEAP, respectively, and 89.49% on valence and

92.86% on arousal. The SEED dataset results showed that data of 8 s with an overlap of 4 s can get the best outcome. Overlap length also affects performance. Additionally, the SAE's ability to extract characteristics from EEG data is demonstrated by the ease with which the proposed network can categorize the data it has processed [8].

4. ZTW-based epoch selection (ZTWBES) algorithm

The main objective of this study is to improve the performance of emotion detection using brain signals by adopting a new adaptive channel selection technique that identifies brain activity as having unique behaviors that change across people as well as emotional situations. Furthermore, the accuracy of the system was improved in this study by locating the epoch at the peak of excitation (only EEG channel voting that matches the location of the selected epoch is considered, other channels are ignored), and by using the molecular group delay (NGD) function to extract instantaneous spectral information with a zero-time window (ZTW) method, thus precisely identifying the epoch in each emotional state. Various categorization systems were designed and tested using QDC and RNN on the DEAP database [9].

In the time domain, ZTW and NGD are used to combine the selection process with spectral epoch selection. Channel selection is the key stage of emotion classification using the quadratic discriminant classifier (QDC) and RNN. The strong electrodes in each emotional state were filtered out of the specified channels using frequency power fluctuation [9].
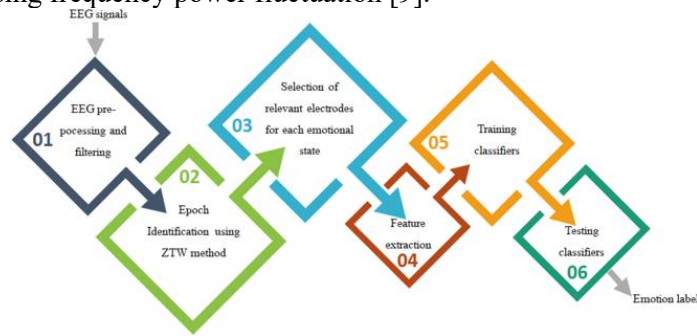

Figure 3. A block diagram of the proposed method [9].

Since EEG signals that have been recorded are typically depicted in a time domain. To begin with, to pre-process and filter the EEG row data, the group-delay function converted EEG signals are converted from the time domain to the frequency domain, allowing the high-resolution spectra to be extracted and their formant components to be highlighted. To trace and extract spectral characteristics from short EEG test fragments, the ZTW technique was used using a window function that has a form similar to a zero-frequency resonator's frequency response to multiply the brief length of each trial at each electrode. For the extraction part, choose the important epoch in each frequency band using the ZTW-based epoch selection (ZTWBES) method. ZTWBES works by breaking down each trial into a group of brief EEG signal segments and locating the frame that contains the epoch. To classify the extracted features, using a two-step classification method using the Quadratic classifier (QDC), and two different RNN-based classification techniques. The generalized form of the linear classifier is a quadratic classifier. To differentiate between the different problem classes, a quadratic decision surface is used. By using QDC, the goal of the first stage was to forecast which of the several emotional states a particular trial may fall under. And the second stage sought to determine the precise emotional state to which the trial t is most likely to belong. RNN, on the other hand, uses time-based sequential data while working with a deep-learning neural network. It is an enhancement over the convolutional neural network, which has fixed inputs, outputs, and hidden layer data flow limitations.
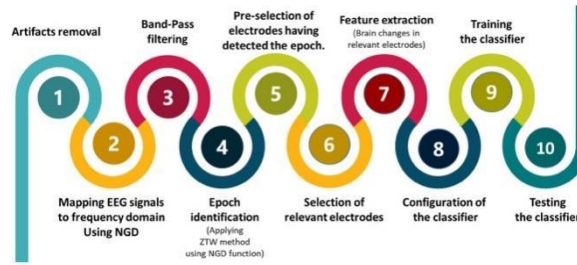
Figure 4. An experimental diagram of this method [9].

The QDC attained an accuracy average of 82.37%, while in classification schemes 1 and 2 the RNN attained an accuracy average of 91.22% and 85.64%, respectively. The experimental findings demonstrated that, in comparison to earlier investigations of multi-class emotion identification, the suggested technique is quite competitive. Over 89% accuracy was achieved on average. The proposed algorithm increased the accuracy rate by 8% compared to existing algorithms dealing with emotions. Additionally, an experiment demonstrates that the suggested system outperforms comparable methods that simply discriminate between 3 and 4 emotions [9].

5. Comparison among the Above Two Algorithms

The accuracy of two algorithms based on the DEAP dataset is 89.49% and over 89% respectively. However, in the second algorithm, the accuracy of the first scheme using RNN is 91.22%, which is much higher than the first algorithm while using the same algorithm to classify the same dataset.

As we can easily find through the performance comparisons with other approaches from Algorithm 1, the average accuracy of using the DEAP datasets is much lower than using SEED datasets, and from the comparison between the suggested approach with related studies from Algorithm 2, we can also analyze that the more emotions used, the higher accuracy we get [8] [9]. Since the DNN also can be applied in the second algorithm, so basically the main difference between the two algorithms is the raw EEG data preprocessing methods, as the result showed, using ZTW and NGD methods are better than using CNN and SAE.

6. Conclusion

In conclusion, by comparing the above two algorithms, we can easily find that we may try to use the SEED dataset by using the ZTW-based epoch selection (ZTWBES) algorithm next time, which may improve the accuracy of the emotion recognition model. And since different of algorithms can result in different effects in training models, the more emotions we used to train the model, the higher accuracy we get.

We can easily find through this essay that row EEG datasets can significantly influence the accuracy of the training model, which also means the higher the quality and quantity of row EEG datasets, the higher the accuracy of the model will be, no matter which algorithm is used to train the emotion recognition model. Due to the properties of EEG, an EEG dataset, no matter how pure the signals it contains, can be more or less easily and severely affected by the external environment at the time of measurement. So maybe in future experiments, for statistics, we can improve the accuracy of EEG datasets, or use more accurate measurement methods than EEG to obtain data, whereas, for the training model, we can use a better and more suitable algorithm to reduce the influence of noise in EEG dataset on accuracy. In addition, we can also try more algorithms to train models to achieve more accurate emotion recognition systems.

References

[1] Psychology is a little spiritual. (2022, July 12). What's killing nearly 100 million people with depression? The 2022 National Blue Book on Depression. _ Ifeng.com. https://i.ifeng.com/c/8Hb7dhuxkoz

[2] Dolan, R. J. (2002, November 8). Emotion, Cognition, and Behavior. Science, 298(5596), 1191–

1194.

[3] Nicu Sebe (The Netherlands), Ira Cohen (USA), and Thomas S. Huang (USA)Handbook of Pattern Recognition and Computer Vision, MULTIMODAL EMOTION RECOGNITION 3rd. January 2005, 387-409

[4] Hu, W., Huang, G., Li, L., Zhang, L., Zhang, Z., & Liang, Z. (2020, September).Video‑triggered EEG‑emotion public databases and current methods: A survey. Brain Science Advances, 6(3), 255–287.

[5] NHS. (2022, January 5). Electroencephalogram (EEG). NHS choices. https://www.nhs.uk/conditions/electroencephalogram/

[6] Sanei, S., & Chambers, J. A. (2022). Chapter 4 Fundamentals of EEG Signal Processing. In EEG signal processing and machine learning (pp. 77–113). essay, Wiley.

[7] Topic, A., Russo, M., Stella, M., & Saric, M. (2022a, April 23). Emotion Recognition Using a Reduced Set of EEG Channels Based on Holographic Feature Maps. Sensors, 22(9), 3248.

[8] Liu, J., Wu, G., Luo, Y., Qiu, S., Yang, S., Li, W., & Bi, Y. (2020). EEG-based emotion classification using a deep neural network and sparse autoencoder. Frontiers in Systems Neuroscience, 14, 43. https://www.frontiersin.org/articles/10.3389/fnsys.2020.00043/full

[9] Gannouni, S., Aledaily, A., Belwafi, K., & Aboalsamh, H. (2021). Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification. Scientific Reports, 11. https://doi.org/10.1038/s41598-021-86345-5