Automatic Stress Recognition in EEG Using Deep Learning

Chen Zhang
Research School of Computer Science, Australian National University
u6745297@anu.edu.au

Abstract. Stress is a human response to changes in external environmental conditions, usually those negative ones. Research shows that if people are under pressure for a long time, their health will be affected. Therefore, how to identify pressure has become an important issue. Traditional detection methods all need to contact users, which means more time and equipment, and at the same time it is more likely to cause users' resistance. Therefore, it is better to use some non-contact methods, such as cameras. These systems mostly use only RGB or thermal image frames to recognize stress. In order to get better results, we need to use two modalities and a feature level fusion of the two modalities, which was mentioned in last paper. In addition, EEG, as an attribute closely related to human emotion, is often used to measure human emotional state. Therefore, based on the EEG signal data extracted from the same experiment, this paper will try to fold the EEG signal to make it become the image and train a CNN to recognize stress, which achieves 63% classification accuracy.

Keywords: Stress Recognition, Reduction, Optimization, Deep learning, CNN, EEG.

1 Introduction

Stress is a common problem that people have in modern society. Stress can be divided into many categories, the most common of which is time pressure. When stress is generated, people will have many psychological reactions, such as heart rate acceleration and respiratory acceleration. However, because stress is a subjective response, different people will show different psychological changes under stress. This makes it difficult to identify stress with a common set of rules[1].

Some other systems are based on physiological data monitored by wearable devices. In this case, if the subject forgets to wear the device or turn on the device, the study cannot be carried out at all. In order to solve the defects of the traditional stress cognitive system, including the inability to continuously monitor subjects [2]. Now people prefer to use non-contact methods, such as RGB and thermal cameras.

Considering that pressure is related to physical appearance, using some special physical characteristics to identify stress is also a potential way. A good case in point is a system that monitoring the subject's facial details including lips, mouth and eyebrows [3]. However, because these methods based on physical characteristics require a lot of equipment and time, and cannot obtain satisfactory results, they have not become the mainstream methods. On the contrary, nowadays imaging techniques like RGB video recorder or thermal imaging have been employed for contact-less measuring of physiological signals [4].

As neuron network become more and more popular and be applied in plenty of areas[5]. Irani, Nasrollahi, Dhall, Moeslund and Gedeon introduced a very efficient method to recognize stress base on RGB and thermal images, implemented by SVM and logical fusion[4]. Their theoretical basis is that In thermal imaging pictures, pixels with similar temperatures will have the same color (Figure 1.a, and the corresponding super-pixels is a group of adjacent pixels which have similar characteristic and special information (Figure 1.b).

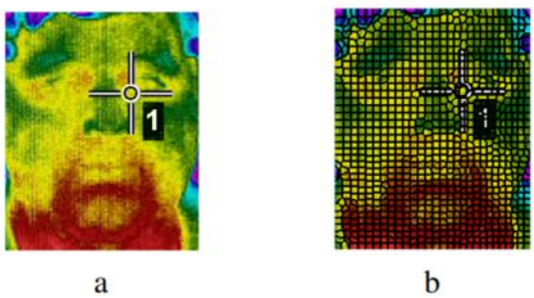

Fig. 1. A typical facial Region (a) and its corresponding super-pixels (b)[4]

Next step is to collect data, implemented by camera and some face cropping technique extract features and construct a model to recognize whether the people is stressful or calm down. Figure 2 shows the overall sequence and structure of the strategy.

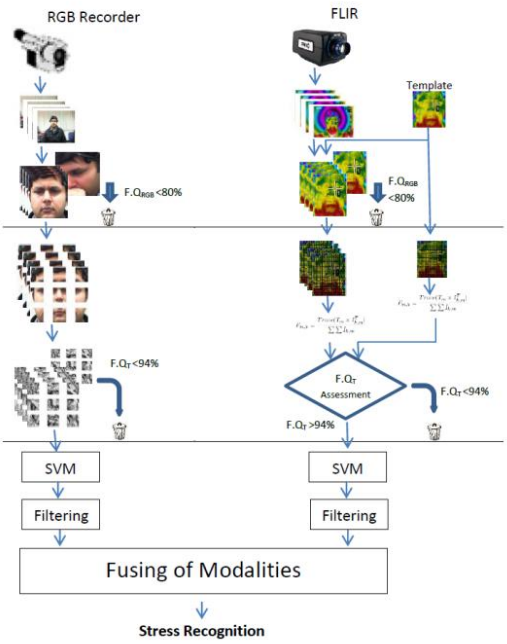

Fig. 2. The block diagram of proposed bimodal system[4]

However, it is obvious that this structure is very complex, It contains two parallel pipelines, which means two SVM, filtering and a fusion. All of these requires more time and consumption. A complex model is not conducive to future expansion or recurrence for research and promotion. As the consequence, last paper introduces a neural network that could achieve similar accuracy but has a light weight structure. In addition, the researchers also used an off-the-shelf wearable EEG device to collect data. Considering that EEG signals can provide rich information about individuals' mental status related stress, many attempts have been made to recognize and measure subjects' stress using EEG signals[6], I believe there is a strong connection between EEG and stress in this experiment. Inspired by Gao Zhongke and his research[7], this paper will aim to train a CNN to recognize stress and achieve similar or higher accuracy to NN in last paper. The rest of the paper is organized as follows: Section II explains the details of the proposed system, Section III discusses the experimental results, and finally, Section IV concludes the paper.

2 Method

2.1 Data preprocessing

Generally speaking, EEG is rarely used in large-scale research because the process of acquiring data is very fragile and easily affected by various factors. The EEG usually has a high degree of artifact contamination and the quality of EEG recordings often substantially differs between subjects[8]. In this case raw EEG data can not be used for training directly a variety of preprocessing methods to clean EEG are applied. Our data are EEG signals of 24 participants and a recording of film timing. Since these EEG data are unlabeled and not all frames of signals are used, I need to convert the film time into time stamps and then combine the data of 24 participants to extract the used signals and label them before formal data processing. According to Pedroni's research, the first step of preprocessing is to identify bad channel and then remove them. Through observation, I found the last column, which is "STI 014" is useless as all values are zero so just drop it. Moreover, the values of data in "CQ" areas is obviously smaller than that in other areas. In general, the value of "CQ" areas is only 1% of that of other areas. So to achieve consistency in dynamic range of various sources of data[9], normalization is necessary. What I choice is a simple but efficient method, which is to scale all values to(0,1) by following formula.

$$x_{normalization} = \frac{x - \mu}{\sigma}$$

After extracting and dropping out the number of valued EEG is around six hundred thousand. Due to the huge amount of data, even if only 5% of the data is used for testing, the test set has 30,000 data, which is enough to test the classification accuracy of the model, so we will use 95% of the data to train the model to improve the quality of the model as much as possible.

2.2 Neural network

At the beginning, I will try to train a traditional neural network, which is similar to last paper. Before we start training neural networks, the first problem we need to face is how to set hyperparameters. Specifically, the hyperparameters mainly include the number of neurons in hidden layers, learning rate and epoch. Scientists usually use the heuristic rules and trial-and-error approach, followed by manual fine-tuning of values by hand to determine the value of hyperparameters[10]. We have to admit that in some cases it could be expensive and it always require user to run the training many times and has good intuition. In last paper I trained three neural networks with different hyperparameters. The original one contains the biggest number of hidden neurons, the second and third one was optimized by T.D. Gedeon and D. Harris' method[11]. Considering that the performance of the neural network optimized twice has declined, I choose the neural network optimized once to train the new data. Cross-entropy loss and Adam was chosen as loss function and optimizer respectively. Figure 3 is the structure of neural network.

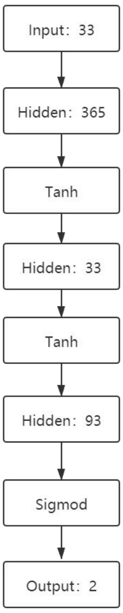

Fig. 3. Neural network structure[12]

2.3 Convolutional neural network

With the development and popularity of deep learning, many researchers have tried to classify EEG base on deep learning, especially CNN(convolutional neural network)[13]. Therefore this paper will also try to train a CNN to identify whether people is stressful or not by EEG. In most of studies EEG was recorded as images and CNN is a neural network specially used to train images. However, the EEG I have are not images, but simply record some information of different brain area. Therefore, before training I have to fold these information then turn them to images. According to figure 4, I can match each attribute to their corresponding position in this picture. Comparing our data with this picture, we can find that this experiment only recorded the information of these peripheral channels, and the information of the central part is missing. Therefore, in order to make the image structure more complete, I will use 0 to fill in some missing central information. Finally I transformed the data as 6*6 matrix and set up 20 3*3 convolutional kernel for feature extraction.

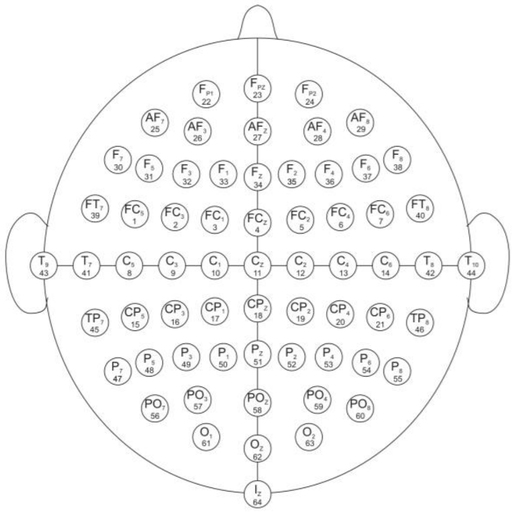

Fig. 4. This diagram illustrates electrode designations and channel assignment numbers[14]

The structure of CNN mainly refers to LeNet5 because it is mainly applied to gray-scale images . However, considering the size of the image is very small, only 6x6, so I deleted a convolution layer. In addition, because the raw data are not images, I added a full connection layer to ensure the ability to extract features. In order to avoid premature neuronal death, I chose RELU6 as the activation function. Negative log likelihood loss and Stochastic Gradient Descent(SGD) was chosen as loss function and optimizer respectively. Figure 5 displays the final structure of CNN.

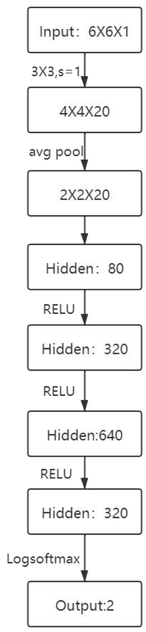

Fig. 4. CNN structure

3 Result and Discussion

3.1 Evaluation of neural network

The first step is to test the neural network mentioned in previous paper. However, the results were so bad that there was no significance for further research and optimization. Figure 5 show the result of neural network.

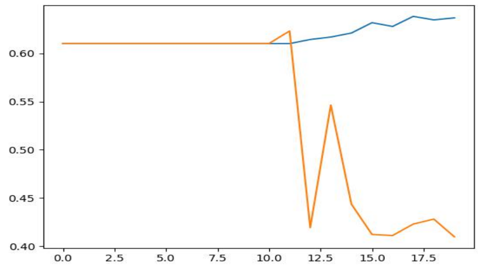

Fig. 4. Train and test accuracy for original model, blue is training, orange is testing.

From the figure, we can see that the training accuracy and testing accuracy have not changed for a long time after the training starts. After that, although the training accuracy increased slightly, the price was a sharp drop in testing accuracy, which is a trend of over-fitting. At the beginning I thought it may caused by RELU activation function because I try RELU to make it converge fastest. So I change RELU to Tanh and check the weights, but did not find a lot of negative weight neurons which may died because RELU ignores any negative value.After further analysis, I found that the straight line at the beginning of the training was caused by the neural network predicting all the data into one class. This shows that the neural network does not learn the features of the data at all, so data cannot be correctly classified. There are many reasons for this result, perhaps because the network structure is not complex, or the network structure is not suitable for learning such data. considering the current results are very bad, and similar research is usually implemented using CNN, I will also focus on CNN too.

3.2 Evaluation of Convolutional neural network

Considering the over-fitting occurred in training the neural network, in order to avoid this problem happening again I applied dropout to CNN. According to Pierre Baldi and Peter Sadowski ' s opinion[15], it does not mechanically inactivate neurons randomly and we should have some deeper thinking about it, such as is it necessary to set q=0.5? How to apply it to multiple layer? A good case in point is that when I apply dropout to all hidden layers, the gap between train and test accuracy decrease indeed, but cannot get a satisfied train accuracy. This may cause by wrong usage of dropout so I followed Baldi and Sadowski ' s suggestion and modify q from 0.5 to 0.2 then get an acceptable result.Figure 5 displays the result with original model.

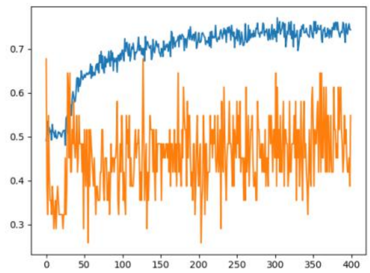

Fig. 5. Train and test accuracy for original model, blue is training, orange is testing.

Just as mentioned above, this result is only acceptable because both of train and test accuracy are not very high, the test accuracy is only achieve 60% and even the train accuracy is only little higher than 70% so I will optimize the network in terms of learning rate and number of neurons. Because EEG has many parameters and the numerical features are not obvious, I reduce learning rate and increase the number of neurons to maximize the feature extraction capability of the model. The full connection layers was expanded to 80-2048, 2048-1024, 1024-256, 256-2 from the first layer to the last layer. Figure 6 displays the result after optimization.

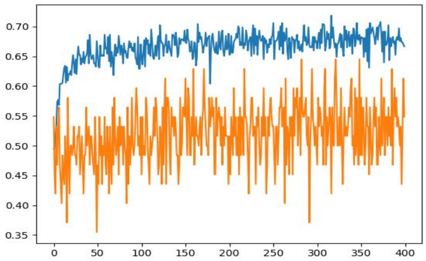

Fig. 6 Train and test accuracy after optimization, blue is training, orange is testing.

Although the training and testing accuracy in the figure is not very high, it can be seen that both are on the rise. Due to the limitation of computer performance, I did not train too much epochs but I believe that with the increase of epoch, the accuracy will be further improved. Table 1 shows the results of comparison of NN and CNN

Table 1. Comparing the result between NN and CNN

Methods	NN	CNN	CNN-2
Best accuracy	67%	60%	63%

4 Conclusion and Future Work

Through the previous comparative experiments, we can find that CNN has advantages over traditional NN in EEG feature extraction. Unfortunately, this paper did not get a very satisfactory result. There are two possible reasons. First of all, due to the limitation of computer performance, I cannot try more complex network structure or increase training epochs. Another reason is that the data may not be strongly related to human pressure. First of all, this experiment only recorded the EEG of some areas and channels, which probably do not reflect the stress state of human. A good case in point is that through additional data analysis, I found that the data of all channels and regions did not fluctuate by more than 15%, which seemed to indicate that they were not affected by participants' stress state. In addition, it is not clear whether the participants actually expressed their psychological state. For example, they may not feel the stress during the whole experiment but lied to the researchers. Therefore, two points should be ensured in future experiments. The first point is to ensure that EEG data are collected as completely as possible. The second point is to ensure that the data are labeled accurately.

References

[1] Lupien, S., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. Brain And Cognition, 65(3), 209-237. doi: 10.1016/j.bandc.2007.02.007

[2] Chen, T., Yuen, P., Richardson, M., Liu, G., & She, Z. (2014). Detection of Psychological Stress Using a Hyperspectral Imaging Technique. IEEE Transactions On Affective Computing, 5(4), 391-405. doi: 10.1109/taffc.2014.2362513

[3] Metaxas, D., Venkataraman, S., & Vogler, C. (2004). Image-based stress recognition using a model-based dynamic face tracking system. International Conference On Computational Science, ICCS 2004.

[4] Irani, R., Nasrollahi, K., Dhall, A., Moeslund, T., & Gedeon, T. (2016). Thermal Super-Pixels for Bimodal Stress Recognition. Sixth International Conference On Image Processing Theory, Tools And Applications (IPTA), 12/2016.

[5] Kwan, H., & Tang, C. (2002). Multiplierless multilayer feedforward neural network design using quantised neurons. Electronics Letters, 38(13), 645. doi: 10.1049/el:20020465

[6] Jebelli, H., Hwang, S., & Lee, S. (2018). EEG-based workers' stress recognition at construction sites. Automation In Construction, 93, 315-324. doi: 10.1016/j.autcon.2018.05.027

[7] Gao, Z., Li, Y., Yang, Y., Dong, N., Yang, X., & Grebogi, C. (2019). A Coincidence Filtering-based Approach for CNNs in EEG-based Recognition. IEEE Transactions On Industrial Informatics, 1-1. doi: 10.1109/tii.2019.2955447

[8] Pedroni, A., Bahreini, A., & Langer, N. (2019). Automagic: Standardized preprocessing of big EEG data. Neuroimage, 200, 460-473. doi: 10.1016/j.neuroimage.2019.06.046

[9] Wang, S., Tang, C., Sun, J., Yang, J., Huang, C., Phillips, P., & Zhang, Y. (2018). Multiple Sclerosis Identification by 14-Layer Convolutional Neural Network With Batch Normalization, Dropout, and Stochastic Pooling. Frontiers In Neuroscience, 12. doi: 10.3389/fnins.2018.00818

[10] Safarik, J., Jalowiczor, J., Gresak, E., & Rozhon, J. (2018). Genetic algorithm for automatic tuning of neural network hyperparameters. Autonomous Systems: Sensors, Vehicles, Security, And The Internet Of Everything, VOLUME 10643(15-19 APRIL 2018).

[11] Gedeon, T.D. and Harris, D. (1991) "Network Reduction Techniques," Proceedings International Conference on Neural Networks Methodologies and Applications, AMSE, San Diego, vol. 1: 119-126.

[12] Chen Zhang(2020) An Attempt to Simplify Neural Networks, Taking Thermal Stress Recognition as an Example

[13] X. Zhang and D. Wu, "On the Vulnerability of CNN Classifiers in EEG-Based BCIs," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 5, pp. 814-825, May 2019, doi: 10.1109/TNSRE.2019.2908955.

[14] Blankertz, B. (2004). Documentation Wadsworth BCI Dataset (P300 Evoked Potentials) [Ebook]. Retrieved from http://www.bbci.de/competition/iii/desc_II.pdf

[15] Baldi, P., & Sadowski, P. (2014). The dropout learning algorithm. Artificial Intelligence, 210, 78-122. doi: 10.1016/j.artint.2014.02.00