EEG based Valence Recognition using Convolutional Kernel on Time-Frequency axis

Jun-Su Kang and Minho Lee
School of Electronics Engineering, Kyungpook National University
Taegu, South Korea
E-mails: {wkjuns, mholee}@gmail.com

***Abstract*—In this study, we present an emotion recognition model, namely valence recognition, based on convolutional kernel on time-frequency axis. Our proposed model uses convolutional kernel on Time-Frequency axis for Convolutional Neural Network (CNN). In order to compare our results with previous ones, we use the DEAP dataset that represents the benchmark for emotion classification research. The preliminary results show that the proposed model has a potential for positive and negative emotion recognition when compared to conventional studies.**

Keywords-EEG; valence recognition; convolutional neural network; long-short term memory; spectrogram.

I. INTRODUCTION

Emotion is one of the most central and pervasive aspects of human experience. Normal people 'feel' and 'express' a wide range of emotions. While emotions deepen and enrich human experience, they also have profound affected on other cognitive functions such as decision-making, reasoning, language comprehension, etc.[1] It has been argued that cognition and emotion are complimentary to each other and one cannot be properly understood or modeled without understanding the other. Especially, at a time when technology is moving towards advanced intelligent systems and smarter robots, this fundamental aspect of human nature cannot be overlooked. Therefore, to develop 'human like intelligence' and/or for a 'qualitative human-machine interaction', it is important that machines also be trained to understand, if not feel, human emotions.

Human emotion is elicited by external stimuli such as image, sound, smell, texture, etc.[2] For understanding the human emotion, we need to catch the internal and external cues from human. For the external cues, there are facial expression, gesture, voice. For the internal cues, we can record the brain or bio signals by attaching sensors to the brain or body. Each modality has its own strength and weakness as shown in Table 1.

TABLE 1. PROS AND CONS FOR RECOGNIZING HUMAN EMOTION

Modality	Pros	Cons
Facial expression	Easy to use	Easy to deceive
Gesture	Easy to use	Easy to deceive
Speech/voice signal	Easy to use	Easy to deceive
Bio-signal	Hard to deceive	Hard to get good signal
Brain signal	Hard to deceive	Hard to get good signal

Among several modalities, we choose electroencephalography (EEG) signals to recognize human valence emotion as 2 classes (positive and negative). EEG signal is a measurement of the brain's electrical activity and provides good temporal resolution. Many prior studies show that EEG power and power asymmetry are related to the emotional valence[3]. The EEG signal with high alpha activity is shown to be an indication of low brain activity, whilst gamma band EEG is connected to high cognitive processes [4]. In this study, we use gamma band power spectrogram as the input of our proposed network. Because, the results of gamma band from the prior studies were better than the results of other frequency bands [5-7]. The rest of the paper is organized as follows: in the next section, the related works are presented. In the section 3, we explain our proposed method for preprocessing and model design. In section 4, the results of the experiment is given. In the last section, we conclude our results and present our future plans.

II. RELATED WORKS

There are many works in the field of emotion recognition. Many studies have tried to extract important emotional features from raw EEG data in the time and frequency domain. Table 2 shows the used feature set of the prior studies.

TABLE 2. FEATURES OF THE PRIOR STUDIES

Prior study	Features	Performance (%)
[8]	440 features (power spectral density(PSD), time domain feature), window size: 50[s], 10[s]	Valence: 78.75
[9]	425 features (PSD, time domain feature), window size: 60[s]	Valence: 76.02
	425 features, window size: 6[s]	Valence: 80.09
[10]	PSD	Valence: 57.60

For time domain features, entropy, kurtosis and zero crossing rate were well used to recognize human emotion. As for frequency domain features, power spectral density (PSD), power subtraction between left and right hemisphere were well used. Feature-based analysis makes it easy to interpret/understand the phenomenon, however, it takes a lot of effort to calculate all those features. Furthermore, the calculation method is very complex. In this study, we only use the spectrogram (gamma band) as the input data.

III. METHOD

We propose an emotion recognition model based on CNN, which attempts to determine the negative and positive emotional state from EEG signals.

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-609-5

A. Dataset description

We used the DEAP dataset, which represents the benchmark for emotion classification research. The data were taken from [10] and included responses of 32 participants (seventeen males, fifteen females). Their mean age was 27.197 years (SD = 4.446) and each subject watched 40 music videos of 60 seconds each with the goal of inducing positive/negative valence emotion. After watching the movie, the subjects were asked to rate the video on a continuous scale ranging from 1 to 9. If the response score of a subject is over 5, we consider it as positive status, otherwise, it's considered negative.

B. Preprocessing of EEG data

We downloaded the preprocessing EEG dataset from [10], but there still are noises such as eye movement. Because of that, we considered Independent Component Analysis (ICA) as a noise removing algorithm [11]. After removing the major noise components with ICA, Short-Time Fourier Transform (STFT) is applied to obtain the spectrogram. The STFT uses a 3 [sec] window to divide the time series data. After dividing the time series data, 1 [sec] window is applied to get the spectrogram from each 3 [sec] windows with overlap ratio of 87.5%.

C. Model design

In order to understand the human emotion from the spectrogram data, convolution kernels on time-frequency axis are used in the CNN structure. In the CNN structure, there are two sub-CNN networks (CNN with convolutional kernel on time axis and CNN with convolutional kernel on frequency axis). Both CNN networks have 6 convolutional layers with 4 drop connections from the 2nd to the 6th layer for increasing the generalization performance [12]. After the 6th convolutional layer, the feature moves to a fully connected layer. The fully connected layer consisted of 3 layers and the output of the 2nd fully connected layer (time/frequency network) is concatenated into one vector. This concatenated feature is moved to the softmax function. Figure 1 shows the part of the proposed CNN structure.

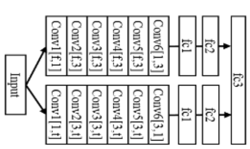

Figure 1. Part of proposed CNN structure

IV. RESULTS

When testing the proposed model, we preliminary test to the 9 subjects among 32 subjects. To increase the regularization performance, l2 regularization is used with weight decay factor of 1e-2, and dropout ratio of 50%. Learning rate is used with 6e-6. For reducing the data bias, 5-fold cross validation is applied (train: data of 32 movies, test: data of 8 movies, 1 fold: data of 8 movies). Figure 2 shows the 9 subject-wise average train and test performance for 5-fold cross validation. The proposed model shows 71.367%(±5.469%) average test accuracy for the 32 subjects.

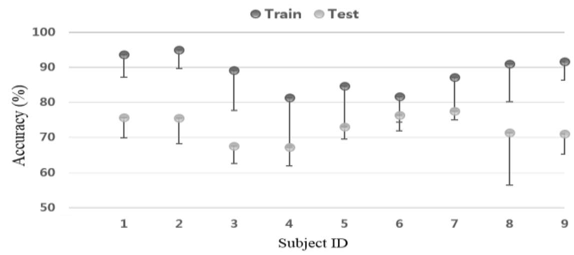

Figure 2. 9 Subject-wise train and test performance (dot means the average accuracy and vertical bar means the standard deviation)

V. CONCLUSION

The characteristics of EEG signal is needed to understand both the time-domain pattern and frequency-domain pattern. Because of that, we design the time-frequency axis convolution kernels for understanding the EEG spectrogram. For future work, the data for other subjects should be considered and also an expansion emotion's dimensions.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2016R1A2A2A05921679).

REFERENCES

[1] L. Pessoa, The cognitive-emotional brain: From interactions to integration: MIT press, 2013..

[2] H. Leventhal, "A perceptual-motor theory of emotion," Advances in experimental social psychology, vol. 17, pp. 117-182, 1984.

[3] Q. Zhang, S. Jeong, and M. Lee, "Autonomous emotion development using incremental modified adaptive neuro-fuzzy inference system", Neurocomputing, vol. 86, pp. 33-44, 2012.

[4] M. Li and B.-L. Lu, "Emotion classification based on gamma-band EEG," in Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, 2009, pp. 1223-1226.

[5] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Emotion classification using minimal EEG channels and frequency bands", in Proceedings of the 10th international joint conference on computer science and software engineering (JCSSE), 2013, pp. 21-24.

[6] W.-L, Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "EEG-based emotion classification using deep belief networks", in Proceeding of IEEE international conference on Multimedia and Expo (ICME), 2014, pp. 1-6.

[7] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification", in Proceeding of 6th international IEEE/EMBS conference on Neural Engineering (NER), 2013, pp. 81-84.

[8] Z. Yin, Y. Wang, L. Liu, W. Zhang, and J. Zhang, "Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination," Frontiers in neurorobotics, vol. 11, pp. 1-16, 2017.

[9] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," Computer Methods and Programs in Biomedicine, vol. 140, pp. 93-110, 2017.

[10] S. Koelstra, et al., "Deap: A database for emotion analysis; using physiological signals," IEEE Transactions on Affective Computing, vol. 3, pp. 18-31, 2012.

[11] M. Chaumon, D. V. Bishop, and N. A. Busch, "A practical guide to the selection of independent components of the electroencephalogram for artifact correction," Journal of neuroscience methods, vol. 250, pp. 47-63, 2015.

[12] L. Wan, M. Zeiler, S. Zhang, Y. L. Cun, and R. Fergus, "Regularization of neural networks using dropconnect," in Proceedings of the 30th international conference on machine learning (ICML-13), 2013, pp. 1058-1066.

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-609-5