EEG based Emotional State Estimation using 2-D Deep Learning Technique

Mehmet Akif OZDEMIR[1], Murside DEGIRMENCI[1], Onan GUREN[2] and Aydin AKAN[1]
[1]Electrophysiological Signals Laboratory, [2]Medical Electronic Laboratory
Department of Biomedical Engineering, Izmir Katip Celebi University
Izmir, Turkey
makif.ozdemir@ikc.edu.tr, mrsddgrmnc.09@gmail.com, {onan.guren, aydin.akan}@ikc.edu.tr

***Abstract*—Emotion detection is very crucial role on diagnosis of brain disorders and psychological disorders. Electroencephalogram (EEG) is useful tool that obtain complex physiological brain signals from human. In this paper, we proposed a novel approach for emotional state estimation using convolutional neural network (CNN) based deep learning technique from EEG signals. Firstly, we convert 32 lead EEG signals to 2D EEG images with Azimuthal Equidistant Projection (AEP) technique. Then, 2D images that represented measurements of EEG signals sent to CNN based deep neural network for classification. In this study, we have achieved accuracy of 95.96% two classes as negative and positive valence, 96.09% two classes as high and low arousal and 95.90% two classes as high and low arousal dominance.**

Keywords—Convolutional Neural Network; EEG Images; Electroencephalogram; Emotion Detection Topographic EEG Maps.

I. INTRODUCTION

Emotion is the most important feature of being human. Emotions have very important effects on different human status such as learning, decision-making, prediction and communication between people [1].

Emotion research is multidisciplinary and complex field which is inclusive medicine, neuroscience, psychology and biomedical engineering [2]. Emotional state estimation is very important for recognize human status and physiological disorders in human-computer interaction (HCI) systems and medicine [3]. Physiological signal such as EEG [4] and galvanic skin response (GSR) [5], facial expression [6], speech [7], visual scanning behavior [8] and multimodal [9] based too many emotion recognition approaches has been published in recently.

EEG based emotional state estimation is noninvasive and effective method for identification brain disorders like as Parkinson's disease [10], and psychological disorders [11], and human status for using in HCI [12].

Emotions is estimated by representing a 2-dimensional axis as a combination of arousal and valence values, or by representing arousal, valence and dominance values in a 3-dimensional plane. This planar model is the Circumplex model called as bipolar model [13]. In the 3D model, the horizontal axis represented by valence values is a scale of emotion between pleasant and unpleasant. The vertical axis represented by the arousal values which is the expression of the degree of emotion. The diagonal axis represented by the dominance values is an indicator of domination in the state of emotion. Also, on the 3D axis, the combination of arousal, valence and dominance values can be assigned to a feeling. The sides where the valence values are positive represent feelings such as happiness and content, negative sides represent feelings such as tired and angry. The sides where the arousal values are high representing feelings such as excited and tense, low sides represent feelings such as tired and angry. The sides where the dominance values are high representing feelings self-control low sides represent feelings influenced [14].

In recently, amount of EEG signals based emotional state estimation approaches have been proposed. Machine learning-based approaches have achieved successful results in automatic emotion recognition. In machine learning based approaches, is generally used two estimation, i.e., EEG feature extraction section and emotion classification approximation section [15].

Nie et al. [16] have achieved average test accuracy of 87.53% that was obtained by using some features and linear dynamic system approach to smooth these features with a support vector machine (SVM).

Tong et al. [17] used SVM with features which are, EEG nonlinear features, power spectrum, entropy and correlation dimension and achieved an average accuracy of 82.22%.

In recent years, deep learning-based approaches have gained much popularity. Deep learning approaches are also basically divided into two different approaches. One is a model of deep neural networks (DNNs) which in EEG signals are sent to deep networks as a single attribute that represent 1D data, the other is convolutional neural networks (CNNs) models in which EEG signals are sent to networks in the form of attributes that represent 2 or 3-dimensional images.

Song et al. [18] proposed a dynamical graph CNNs which is achieved accuracies of 86.23%, 84.54% and 85.02% respectively for valence, arousal and dominance classifications on the DREAMER database.

Zheng et al. [19] achieved 86.08% average accuracy using DNNs with creating their own 62-channel EEG databases.

978-1-7281-2420-9/19/$31.00 ©2019 IEEE

Li et al. [20] proposed a hybrid deep learning structure which is consist of CNN and recurrent neural network (RNN). They used this model on Dataset for Emotional State Analysis and Using Physiological signals (DEAP) for emotion recognition and achieved average accuracy of 73.09%.

In this paper, we proposed a new method for EEG based emotion estimation with using CNN. EEG signals used in this study which were obtained from DEAP. Firstly, 32 channel EEG signals converted to 2D topographic EEG images using different EEG frequency bands (alpha, beta and gamma). Then, the dataset of 2D EEG images was sent to the proposed CNN architecture for classification. The aim of this study is to discover new deep learning techniques to achieve a higher success rate with lower cost for emotional state estimation.

II. METHODS

A. DEAP Dataset and Preprocessing

The publicly available an international DEAP [21] were used to test the efficiency of our network. DEAP includes the 32 channel EEG and 8 peripheral physiological signals of 32 participants when they watched 40 one-minute music video clips. According to 10-20 electrode positioning system, 32 electrodes of the system was placed and EEG signals recorded. Before the each of the original EEG recording, 3 seconds baseline determined and recorded then 60 seconds EEG signals recorded. In this study, preprocessed version of the DEAP dataset was used for experimental studies. In this version sample rate defined as 128 Hz, band pass filter that range 4-45 Hz was applied to eliminate noise and EOG artefacts were extracted. In preprocessing section 3 second baseline time was removed. At the end of each music video clips, volunteers rated video clips in terms of their levels of five different status, namely, arousal, valence, dominance, liking and familiarity. Self-assessment manikins (SAM) were assigned values from 1 to 9 which were used to define the emotional label of the video clips . According to mean value of participants' self-assessment, if mean of rating values <5, the label assigned as low for three emotional state and if mean of rating values ≥5 the label of three emotional state is high [22].

Arousal Valence Dominance (AVD) model was used to represent emotional states. Arousal is represented in vertical axis and scale of it ranges from inactive or calm to active or excited. Valence is represented in horizontal axis and scale of it ranges from negative or unhappy to positive or happy. Another diagonal axis dominance known as capacity of being in control of person's emotions and scale of it ranges from submissive to dominant. The axis of the three emotion states ranges from low emotional state to high emotional state for emotional state estimation. By determining low and high state of the arousal valence and dominance value of defined emotions, emotions can be recognized as the different eight class that consists of three axes as shown in Fig. 1.

Multi-channel EEG time series from each video trial of 60 seconds were sliced into 15 seconds pieces and an image was created over each 15 seconds time frames to obtain four frames for input data of training.

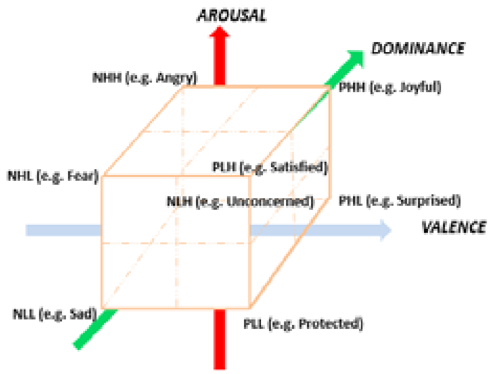

Fig. 1. 3D AVD model of emotional state estimation.

B. Images from Multi-channel EEG Time Series

EEG time series decomposed into five specific frequency bands that are known as 0.5-4 Hz (Delta frequency band), 4-8 Hz (Theta frequency band), 8-13 Hz (Alpha frequency band), 14-30 Hz (Beta frequency band), 30-60 Hz (Gamma frequency band). According to neural activity of brain and bands of EEG signals, frequency characteristics of EEG signals change which is related to emotional state estimation. Zhuang et al. [22] showed that beta and gamma which are known as higher frequency band of EEG signals outperforms delta and theta band which are known as lower frequency band of EEG. The impact of valence state value was observed on gamma frequency band of EEG [23]. The feeling of being present linked to alpha band of EEG signals that obtained from frontal side of brain. The important frontal changes in power on alpha range is related with changes on arousal emotional state. Fast Fourier Transform (FFT) is used to obtain the power spectrum of each frame (15 sec).

In EEG processing of this study, the total frequency spectrum was subtracted into three frequency sub-bands: alpha, beta, and gamma. According to effectiveness of frequency bands, three frequency bands selected and their features collected. For each frame, mean power spectrum of three frequency band was calculated by using magnitude of FFT values and these are considered as feature (32 channel x 3 frequency bands) of EEG time series. The EEG time series are obtained from the measurements of spatial cortex locations of electrodes that are scattered in a three-dimensional space. In this study, power spectrum values of each frequency band and 2D electrode locations were computed to obtain 2D images transformation of EEG measurements so spatial dimension of EEG time series was preserved. Three different colored frequency band channels represent the spectral dimension of the EEG time series. The projection of electrode locations from three dimensional to two-dimensional space is needed to preserve the spatial structure of EEG. To accomplish this, the AEP that also known as polar projection technique is used due to characteristic of preserving relative distances of electrode locations. In 2D electrode locations, horizontal and vertical dimensions of the colored image assigned to spatially scattered activity values over the surface of brain.

One of the most important steps of this procedure is interpolating of the three different colored frequency band over the cortex for assigning values between electrodes over 32x32 mesh. Clough Tocher scheme was used for this purpose. This procedure applied for each of three frequency band and obtained into three maps of topographical activity for each of the frequency band. The three topographical activity maps merged together to obtain the three channeled color images which is used as an input image data to CNN structure. Image construction from EEG signal is illustrated in Fig. 2.

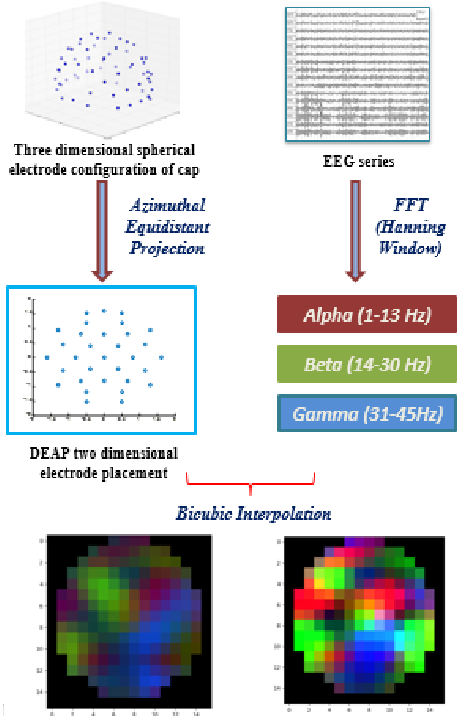

Fig. 2. The illustration of image construction from EEG time series.

C. Convolutional Neural Network Architecture

The CNN were used by virtue of the fact that its characteristic of learning 2D input data of EEG time series with size of 16x16. This structure processed the alterations in space and frequency domain of 2D image. The proposed CNN structure is summarized in Fig. 3. The network mimics the VGG structure which is used in classification of 2D data and includes the three convolutional layers, three max-pooling layers, and one fully connected layer. The convolutional layers with small kernel size of 3x3 are stacked together which are followed by max-pooling layer with small kernel size of 2x2 and stride of 2. After all operations of convolutional layers and max-pooling layers, each frame feeds to the fully connected layers and prediction of frames was processed with Softmax classifier as low or high emotional states (arousal, valence, dominance).

D. Network Training

In training of network, test size determined as 0.20 and another part of data used as train data after experimental results of different test sizes. Batch size has been set as 64 and epoch number was found as 500 to converge parameters of network. Learning rate defined as 10^{-3}. Kernel size defined as 3x3 for convolutional layers and 2x2 with stride of 2 for max-pooling layers respectively. Number of three stacked convolutional layers represented as 16, 32, 64 respectively.

III. RESULT AND DISCUSSION

The performance of CNN approaches depends on the multiple parameters so proposed network should be designed with optimal values of parameters. According to the trials of multiple training, the optimum values of image size, number of convolutional layers, number of max-pooling layers, kernel size, stride number, batch size, learning rate, epoch number, test size, determined for emotional state estimation. The results show that the proposed network with optimum parameters achieved average test accuracy of 96.09%, 95.96%, and 95.90%, for arousal, valence, and dominance emotional states as shown in Fig. 4, Fig. 5, and Fig. 6 respectively.

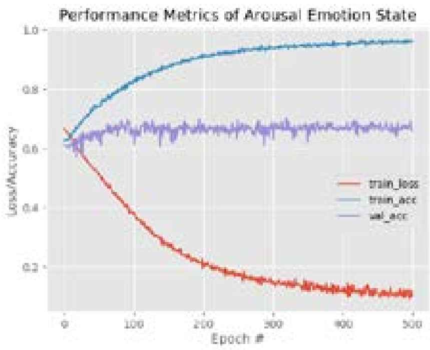

Fig. 4. Classification results of proposed network on Arousal.

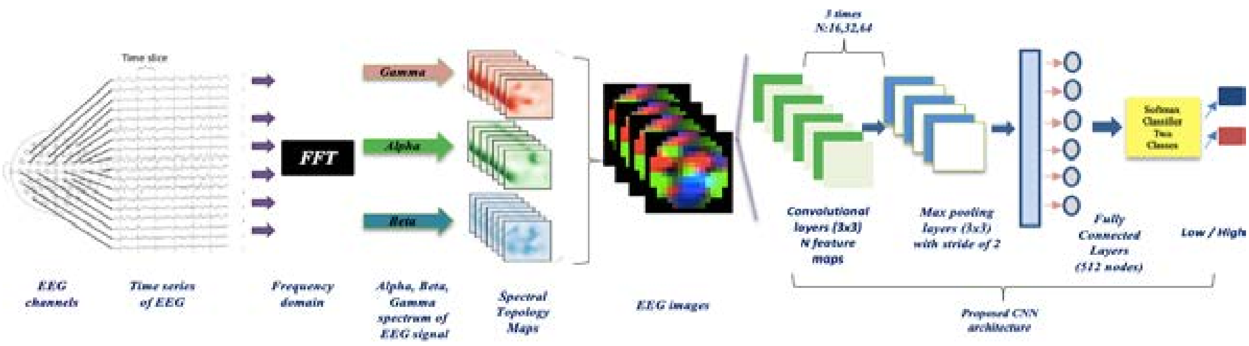

Fig. 3. Overview of the proposed approach.

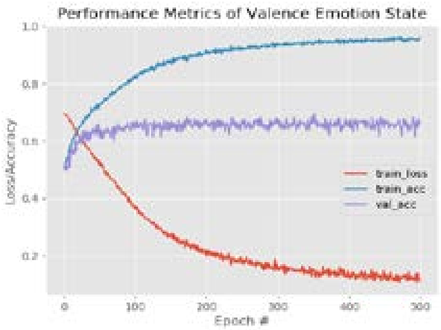

Fig. 5. Classification results of proposed network on Valence.

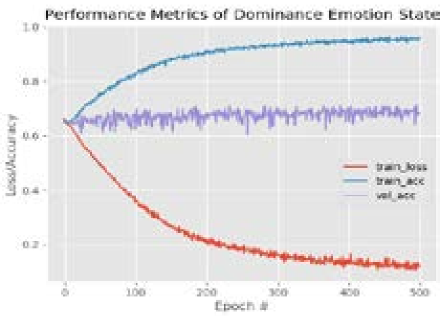

Fig. 6. Classification results of proposed network on Dominance.

IV. CONCLUSION

In this study, a new approach CNN based emotional state estimation proposed and trained on EEG signals of DEAP dataset. EEG time series converted to 2D images to give the training network as input data. In steps of making images from EEG signals, 2D electrode locations that processed with AEP technique were used and these location values interpolated with topographical activity maps of three frequency bands. The 2D three colored images represented the EEG signals and they used as features in classification. These images preserve the spectral, temporal and spatial attributes of EEG time series in their structure to handle original characteristic of EEG signal. The image construction technique implemented with new CNN model, which in structure is different from earlier deep learning-based attempts. The training results revealed that the proposed network demonstrated the important improvements on separating high and low class of the three emotional states by giving high success rates.

REFERENCES

[1] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *International Journal of Advanced Computer Science and Applications(IJACSA),* vol. 8, no. 10, pp. 355-358, 2017.

[2] W. Liu, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using multimodal deep learning," in *International conference on neural information processing*, 2016: Springer, pp. 521-529.

[3] Y. Gao, H. J. Lee, and R. M. Mehmood, "Deep learninig of EEG signals for emotion recognition," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2015: IEEE, pp. 1-5.

[4] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing,* 2018.

[5] Q. Zhang, X. Lai, and G. Liu, "Emotion recognition of GSR based on an improved quantum neural network," in *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2016, vol. 1: IEEE, pp. 488-492.

[6] L. van Zonneveld, L. de Sonneville, S. Van Goozen, and H. Swaab, "Recognition of facial emotion and affective prosody in children at high risk of criminal behavior," *Journal of the International Neuropsychological Society,* vol. 25, no. 1, pp. 57-64, 2019.

[7] K.-Y. Huang, C.-H. Wu, Q.-B. Hong, M.-H. Su, and Y.-H. Chen, "Speech Emotion Recognition Using Deep Neural Network Considering Verbal and Nonverbal Speech Sounds," in *ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019: IEEE, pp. 5866-5870.

[8] L. Dinkler *et al.*, "Visual scanning during emotion recognition in long‐term recovered anorexia nervosa: An eye‐tracking study," *International Journal of Eating Disorders,* 2019.

[9] X. Huang *et al.*, "Multi-modal emotion analysis from facial expressions and electroencephalogram," *Computer Vision and Image Understanding,* vol. 147, pp. 114-124, 2016.

[10] M. H. Black *et al.*, "Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography," *Neuroscience and Biobehavioral Reviews,* vol. 80, pp. 488-515, 2017.

[11] H. Cai *et al.*, "A pervasive approach to EEG-based depression detection," *Complexity,* vol. 2018, 2018.

[12] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Systems with Applications,* vol. 47, pp. 35-41, 2016.

[13] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology,* vol. 39, no. 6, p. 1161, 1980.

[14] G. K. Verma and U. S. Tiwary, "Affect representation and recognition in 3D continuous valence–arousal–dominance space," *Multimedia Tools and Applications,* journal article vol. 76, no. 2, pp. 2159-2183, January 01 2017, doi: 10.1007/s11042-015-3119-y.

[15] H. U. Amin *et al.*, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australasian physical & engineering sciences in medicine,* vol. 38, no. 1, pp. 139-149, 2015.

[16] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "EEG-based emotion recognition during watching movies," in *2011 5th International IEEE/EMBS Conference on Neural Engineering*, 2011: IEEE, pp. 667-670.

[17] J. Tong *et al.*, "EEG-based emotion recognition using nonlinear feature," in *2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, 2017: IEEE, pp. 55-59.

[18] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks," *IEEE Transactions on Affective Computing,* pp. 1-1, 2018, doi: 10.1109/TAFFC.2018.2817622.

[19] W. Zheng and B. Lu, "Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks," *IEEE Transactions on Autonomous Mental Development,* vol. 7, no. 3, pp. 162-175, 2015, doi: 10.1109/TAMD.2015.2431497.

[20] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, "Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network," in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2016: IEEE, pp. 352-359.

[21] S. Koelstra *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing,* vol. 3, no. 1, pp. 18-31, 2011.

[22] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *BioMed research international,* vol. 2017, 2017.

[23] O. Bălan, G. Moise, A. Moldoveanu, M. Leordeanu, and F. Moldoveanu, "Fear Level Classification Based on Emotional Dimensions and Machine Learning Techniques," *Sensors,* vol. 19, no. 7, p. 1738, 2019.