2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)
Virtual Conference, May 4-6, 2021

EEG-based Classification of the Intensity of Emotional Responses

Vahan Babushkin[1], Wanjoo Park[2], Muhammad Hassan Jamil[2], Haneen Alsuradi[1], and Mohamad Eid[2]

***Abstract*— Considering the fact that emotional experiences are stored in the brain, classifying emotion from brain activity measured using electroencephalography has become a trend. In most of the previous studies, the user's emotions were classified based on a stimulus. In this paper, we present a model that can classify the emotion intensity by the participants' self-report. Two machine learning classifiers are considered: support vector machine (SVM) and convolutional neural networks (CNN). Results demonstrated that both SVM and CNN models perform well with four classes of emotions (positive/negative valence high/low arousal combination) where SVM achieved an accuracy of 85% whereas CNN achieved 81%. Considering 12 classes of emotional responses (low, medium, and high intensity for positive/negative valence high/low arousal combination) by the participants' self report resulted in an accuracy of 70% for SVM and 69% for CNN. The proposed model excels in classifying emotional intensity and provides superior performance compared to the state-of-the-art emotion classification systems.**

***Index Terms*— Emotion recognition, Affective computing, Machine Learning, Biomedical signal processing**

I. INTRODUCTION

EMOTIONS are a fundamental human experience; commonly associated with decision making, perception and cognition, human interaction, and performance and intelligence [1]. Affective computing is a rising topic in human-computer interaction to recognize and/or influence human emotion. When defining emotions, an explicit separation is made between physiological arousal, the behavioral expression of emotion (affect), and the conscious experience of an emotion (feeling) [2]. Physiological arousal is measured using physiological signals such as the user's heart rate, skin conductance, and pupil dilation. Human behavior such as facial expressions, voice, and body gestures concern the expression of emotion. The mental experience of emotion can be subjectively measured using self-reporting such as the self-assessment manikin [3] or directly tapping into the brain activity using scanning techniques such as electroencephalography (EEG) [4].

Early research in the field of EEG-based emotion recognition demonstrated distinguished brain activation associated with emotional responses [5]. As for valence, it was shown that happy emotions result in a higher frontal coherence in alpha, and higher right parietal beta power, compared to negative emotions. On the other hand, excitation resulted in a higher beta power and coherence in the parietal lobe, in addition to lower alpha activity.

[1]V. Babushkin and H. Alsuradi are with Tandon School of Engineering, New York University, NYC, 11201, USA

[2]W. Park, M. H. Jamil, and M. Eid are with the Engineering Division, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, 129188, United Arab Emirates mohamad.eid@nyu.edu

Machine learning systems are gaining great popularity in neuroscience, particularly for emotions classification using EEG data. Different machine learning algorithms such as K-nearest neighbor, Bayesian network, artificial neural network, and support vector machine (SVM) are applied to the recorded EEG data for extracting the emotional levels [6]. Machine learning approaches are used not only for emotions classification but for revealing the emotional activation mechanism [7]. In general, the accuracy of emotion recognition depends on several factors such as different experiment environments, pre-processing techniques, feature selection, etc. [6].

II. RELATED WORK

SVM remains one of the most popular classifiers for the supervised multi-class recognition of emotions from raw EEG data. An SVM classifier for discerning the valence and arousal in EEG data was able to achieve accuracy of 32% and 37% in [8]. Preprocessing EEG data to extract important statistical features that are used to classify emotions using SVM resulted in a mean accuracy of 85.17% for six emotions [9]. A subsequent study utilized independent component analysis before classifying the EEG recordings into seven emotion classes with SVM and LDA classifiers, with accuracies of 74.13% and 66.50% respectively [4]. Real-time classification of five emotions using SVM was possible with an average accuracy of 70.5% [10]. All these studies are limited to four to six classes of emotions.

Meanwhile, the growing popularity of deep learning resulted in a number of studies using convolutional neural network (CNN) for emotion recognition from the EEG data [11]. Since CNNs are capable of learning hidden dependencies in raw data, there is no need to engineer new features, which might significantly reduce the preprocessing stage. A CNN model used the partial structure of AlexNet [12] to recognize two classes of emotions: arousal (with an accuracy of 87.30%) and valence (85.50%) [13]. A similar approach has been adopted by Liu et al. [14] using ResNets for emotion classification from raw EEG data, achieving almost an accuracy of 90% for high/low valence and 58.03% for high/low arousal. A CNN-based model achieved an accuracy of 85% for classifying valence and arousal, 77% for classifying positive, neutral, and negative valence/arousal, and 61% for classifying four categories of emotional states, namely low arousal/low valence (LALV), low arousal/high valence (LAHV), high arousal/low valence (HALV) and high arousal/high valence (HAHV) [15]. In order to improve classification accuracy, the authors in [16] utilized the EEG spectrogram and wavelet transformed galvanic skin

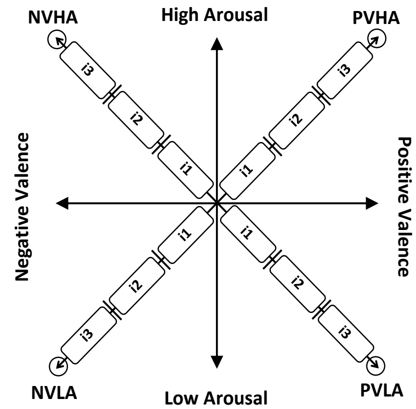

Fig. 1: 4 Emotion classes: Positive Valence High Arousal (PVHA), Negative Valence High Arousal (NVHA), Negative Valence Low Arousal (NVLA), Positive Valence Low Arousal (PVLA). Emotion intensity levels: i1 – i3 correspond to low, medium and high intensities respectively

response (GSR) to recognize the same four categories as in [15], achieving an accuracy of 73.43%. In an attempt to incorporate the spatial arrangement of EEG electrodes information, a 3D CNN model is developed to learn the positional and temporal features from the raw EEG data, thus achieving an accuracy of 87.44% for valence and 88.49% for arousal. A recent study [17] reported a 95.20% accuracy achieved with a CNN-based approach on DEAP dataset. Also, ensemble of CNNs, Sparse Autoencoder (SAE), and Deep Neural Network (DNN) demonstrates promising results with an accuracy of 89.49% on valence and 92.86% on arousal for DEAP and 96.77% for SEED datasets [18].

Previous emotional classification studies classified participants' emotional status by stimuli. However, the participants' emotions may be varied depending on individual differences, culture, or well-being. In other words, the participants' emotional state can be different even in the presence of the same stimuli, i.e. the same stimulus causes a similar sensibility; however, the intensity of the sensibility may differ. In this paper, we show how to detect the level of sensibility from participants' ratings (level of the particular emotion). Apart from the four emotional states according to the circumplex model, we also used 12 emotional subcategories defined from the ranking of the subjects collected during the experiment (high, medium, and low intensity for positive/negative valence and high/low arousal), such as Fig. 1. We conduct a comparative analysis of SVM and CNN for classification by the three levels of each emotional category and achieve a reasonable accuracy.

III. Methodology

A. Experimental setup and protocol

A total of 34 subjects were recruited for this study. Participants were briefed about the objective of the experiment, and a written informed consent was obtained prior to voluntary participation. The study was carried out with an approved protocol by New York University Abu Dhabi Institutional Review Board (#073-2017).

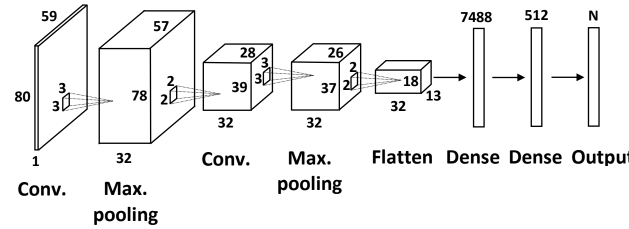

Fig. 2: Proposed CNN architecture ($N = 4$ or 12).

A total of 80 images were selected from the IAPS dataset [19] to represent the four classes of emotional responses (20 images each), namely positive valence high arousal (PVHA), positive valence low arousal (PVLA), negative valence low arousal (NVLA), and negative valence high arousal (NVHA). Note that images were selected based on the highest ratings for the respective emotional response.

An application was developed using the Presentation software (by Neurobehavioral Systems, Albany, CA, USA) in order to control the visual and auditory cues and synchronize these cues with the stimuli displayed on the monitor. The application recorded the participants' response and event trigger information in the EEG system. The experimental setup consisted of a monitor to display the wash-off video and selected IAPS images to elicit emotions and a numerical keypad to collect rating response from the participants. Participants were also asked to wear a pair of earphones with a white noise playing in the background to minimize the external auditory interference.

All 80 trials were breakdown into 4 runs (20 images each), representing the four classes of emotional responses (PVLA, PVHA, NVLA and NVHA, respectively). All runs were started/ended with a 20 seconds wash-off video to neutralize the emotional state of the participants'. Each image stayed on the display for a duration of 8 seconds before asking participants' for the rating on the keypad, followed by a 50 millisecond break period before loading the next image.

B. EEG Data Preprocessing

A 1000 Hz sampled EEG signal was recorded through a 64-channel Brain Product EEG system. First, outside channels of FT9, FT10, TP9, and TP10 were removed and then, a 0.1–85 Hz band-pass filter and 50 Hz notch filter, and a common average reference method were applied. After that, data was divided into two sets: the data set epoched according to the four stimulus events and the data set epoched according to the rating of the participant's emotional status after viewing the stimuli. For the two data sets, the power spectral density of 1–80 Hz frequency bins was extracted through Short-time Fourier transform with 500 ms window was shifted by 50 ms. Each frequency bin had baseline correction using the one second interval before image stimulation as a baseline.

C. Emotion Classification Methods

1) SVM: The processed EEG data represents power density from 2671 trials, each containing 160 timepoints

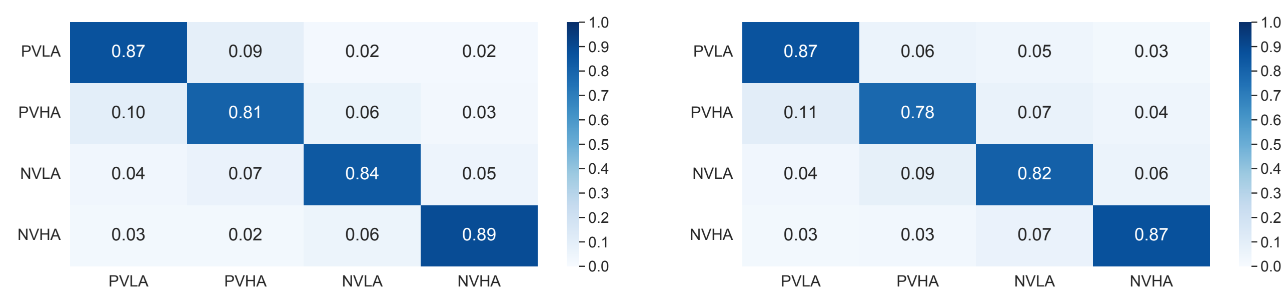

Fig. 3: Average of normalized confusion matrices over 10 folds for SVM, 4 classes (left), CNN, 4 classes (right)

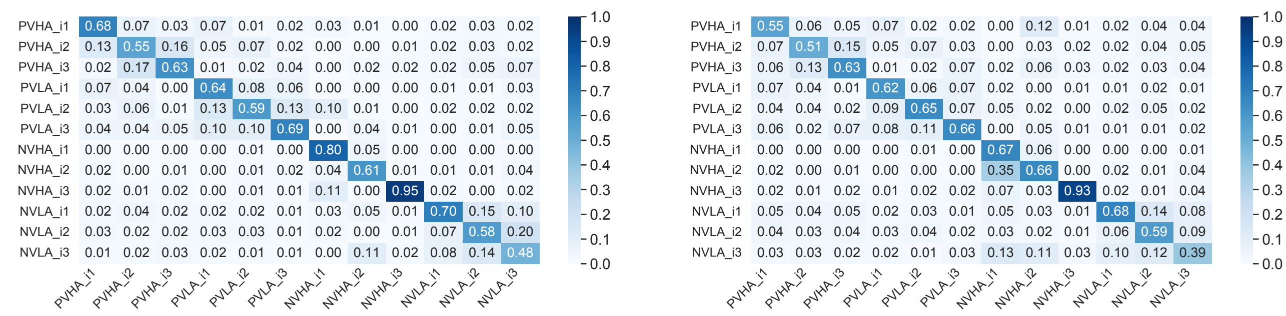

Fig. 4: Average of normalized confusion matrices over 10 folds for SVM, 12 classes (left), CNN, 12 classes (right)

recorded from 59 sensors and distributed across 80 frequency bins. Before classifying with SVM, the 80 $\times$ 160 $\times$ 59 $\times$ 2671 tensor was averaged across the time and three frequency bins, corresponding to beta, lower gamma, and higher gamma bands. Then the newly-obtained 3 $\times$ 59 $\times$ 2671 trials matrix was flattened to get a vector of length 177 for each trial. PCA analysis was conducted to identify the optimal number of principal components and the Radial Base Function kernel coefficient γ. The experiments with 4 and 12 classes demonstrated that the optimal number of principal components is 25 with $\gamma = 0.1$.

2) CNN: A simple CNN model was utilized for predicting 4 and 12 categories of emotional responses from the EEG spectrogram, averaged over time. The model, presented on Fig. 2. consisted of two 2-D convolutional layers each with 32 filters and 3$\times$3 kernels with the valid mode of padding. As an activation, the ReLu function was used. Each convolutional layer was followed by a 2-D max pooling layer with 2$\times$2 pooling window and there was a flattening layer with two fully-connected (Dense) layers with ReLu and softmax activations at the output. After preprocessing, the EEG data contained 2671 trials, each trial consisting of 80 frequency bins, 160 timepoints, and 59 channels. The 4D tensor of 80 $\times$ 160 $\times$ 59 $\times$ 2671 was averaged across the second dimension (time) and an additional singleton dimension was added to the newly-created 3-D tensor of 80 $\times$ 59 $\times$ 2671 to make it suitable for processing with CNN. A total of 100 epochs were used to train the CNN, even though the model achieved training accuracy close to 1 after 20-th epoch for all classes during the 10 folds cross-validation. The output of the CNN is encoded as binary labels, later de-binarised into the multiclass vector.

IV. Results

In the case of 4 classes where equal number of samples per class were available, both SVM and CNN classifiers performed well, achieving average accuracy of 0.85 for SVM (Fig. 3, left) and 0.81 for CNN (Fig. 3, right). Increasing the number of classes introduces a bias in the number of trials, with the most underrepresented class containing only 8.6% of the number of trials of the most over-represented one. The presence of heavy imbalance in the data leads to a drop in accuracy of classifiers. For example, both CNN and SVM achieve accuracy of 0.70 and 0.69 for 12 classes correspondingly. The precision and recall of the SVM classifier for 12 classes exceeds 0.9 for negative valence high arousal of high-intensity level (NVHA_i3), which is also noticeable from the confusion matrix in Fig. 4, (left). Interestingly, the precision of 0.81 with recall of 0.80 were achieved for the most underrepresented class for negative valence high arousal of low intensity (NVHA_i1). It could be explained by the low number of samples to test, and the misclassification error, in this case, is also low. The low number of samples for low-intensity level might be attributed to the subjects' obscurity to differentiate between low and medium intensities of the elicited emotion for some categories. In the case of 12 classes, CNNs also achieved the highest accuracy for high-intensity of negative valence high arousal (NVHA_i3, see Fig. 4, right), with precision of 0.88 and recall of 0.90. However, performance metrics for NVHA_i1 were close to the other classes. In general, the heavy bias in the dataset of 12 classes affects the accuracy of both classifiers. Overall, both classifiers achieved an average accuracy above 0.8 for the dataset of 4 classes with evenly

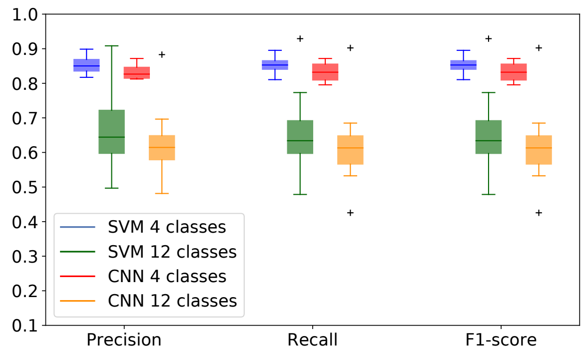

Fig. 5: Average precision, recall and F1-score after 10-fold cross validation for SVM and CNN classifiers.

distributed samples and average accuracy around 0.7 on 12 classes containing heavily underrepresented instances. SVM also achieved slightly better precision and recall averaged over 10 folds with four classes of emotions (an average precision of 0.85 vs. 0.83 for CNN and the same for recall). The same metrics are also slightly higher for the SVM classifier in the case of 12 classes (an average precision of 0.67 vs. 0.63 for CNN and the average recall of 0.65 vs. 0.62 for CNN, see Fig. 5). The SVM demonstrated a slightly better performance than CNN; it trains faster than CNN and easy to implement. The better performance of the SVM can be attributed to the averaging the data over beta, lower gamma, and higher gamma bands at the preprocessing stage – due to the time insensitivity of the task, an averaged activation within a frequency band should be representative of the emotional state of the participant. The spatial positioning of the electrode could also be taken into account while preprocessing the data for the SVM. Also, at the preprocessing stage, additional features can be engineered and used for more accurate classification with SVM. Furthermore, using PCA for dimensionality reduction before classifying with SVM helps to remove noise and thus, avoid overfitting. Despite the CNN ability of inferring hidden dependencies, inclusion of manually-engineered features reflecting the neurobiological processes is more preferential than relying on CNN's automatically generated feature maps, which makes the SVM classifier a preferred choice for emotion recognition from EEG recordings in the presence of biased datasets.

V. Conclusion and Future Work

In this study, we introduced a model to classify the intensity of emotional responses. We demonstrated that 12 classes of emotions considering the intensity of emotions through self-report of participants can be classified with high accuracy through EEG signals. The proposed model demonstrated a superior performance with SVM achieving an average accuracy of 70% and CNN of 69% for recognizing 12 classes of emotions after 10-fold cross-validation. This is less accurate than recognizing 4 emotion classes based on the stimulus, however considering the chance level, it does not mean less accurate.

Future work will be done to transform the classification model into a regression model where arousal and valence can be mapped into a quantitative number (positive and negative). Furthermore, we plan to extend the model to classify/measure emotions influenced using other modalities such as audio or touch. Finally, the trained model can be used in several applications where the intensity of emotion plays a crucial role, including health care, gaming and entertainment, and biofeedback.

References

[1] T. Brosch, K. R. Scherer, D. M. Grandjean, and D. Sander, "The impact of emotion on perception, attention, memory, and decision-making," *Swiss medical weekly*, vol. 143, p. w13786, 2013.

[2] L. F. Barrett, B. Mesquita, K. N. Ochsner, and J. J. Gross, "The experience of emotion," *Annu. Rev. Psychol.*, vol. 58, pp. 373–403, 2007.

[3] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[4] A. Bhardwaj, A. Gupta, P. Jain, A. Rani, and J. Yadav, "Classification of human emotions from EEG signals using SVM and LDA classifiers," in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 180–185.

[5] A. Choppin, "EEG-based human interface for disabled individuals: Emotion expression with neural networks," *Unpublished master's thesis*, 2000.

[6] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using EEG," in *Foundations of Augmented Cognition*. Berlin, Heidelberg: Springer, 2013, pp. 492–501.

[7] C. Qing, R. Qiao, X. Xu, and Y. Cheng, "Interpretable emotion recognition using eeg signals," *IEEE Access*, vol. 7, pp. 94 160–94 170, 2019.

[8] R. Horlings, D. Datcu, and L. Rothkrantz, "Emotion recognition using brain activity," 01 2008, p. 6.

[9] P. Petrantonakis and L. Hadjileontiadis, "Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis," *Affective Computing, IEEE Transactions on*, vol. 1, pp. 81–97, 07 2010.

[10] V. Anh, M. Van, B. Ha Bang, and T. Huynh Quyet, "A real-time model based support vector machine for emotion recognition through EEG," 11 2012, pp. 191–196.

[11] C. C. Aggarwal, *Neural Networks and Deep Learning*. Springer, 2018.

[12] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Neural Information Processing Systems*, vol. 25, 01 2012.

[13] L. Wenqian, C. Li, and S. Sun, *Deep Convolutional Neural Network for Emotion Recognition Using EEG and Peripheral Physiological Signal*, 12 2017, pp. 385–394.

[14] N. Liu, Y. Fang, L. Li, L. Hou, F. Yang, and Y. Guo, "Multiple feature fusion for automatic emotion recognition using EEG signals," 04 2018, pp. 896–900.

[15] H. Mei and X. Xu, "EEG-based emotion classification using convolutional neural network," in *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2017, pp. 130–135.

[16] Y.-H. Kwon, S.-B. Shin, and S. Kim, "Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system," *Sensors (Basel, Switzerland)*, vol. 18, 2018.

[17] R. Alhalaseh and S. Alasasfeh, "Machine-learning-based emotion recognition system using eeg signals," *Computers*, vol. 9, no. 4, 2020.

[18] J. Liu, G. Wu, Y. Luo, S. Qiu, S. Yang, W. Li, and Y. Bi, "EEG-based emotion classification using a deep neural network and sparse autoencoder," *Frontiers in Systems Neuroscience*, vol. 14, p. 43, 2020.

[19] P. Lang, M. M. Bradley, and B. Cuthbert, "International affective picture system (IAPS): affective ratings of pictures and instruction manual." University of Florida, Gainesville, Tech. Rep. A-8, 2008.

[20] V. Babushkin, "Code repository," https://github.com/vbabushkin/EEG-based-Classification-of-the-Intensity-of-Emotional-Responses, 2020.