Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Machine Learning Approach Towards Emotion Recognition from EEG Recordings

Project Report for ECE-GY 6123 "Introduction to Machine Learning"

Course Instructor: **Dr. Sundeep Rangan**
Mentor and PI: **Dr. Mohamad Eid**
Co-mentor: **Dr. Wanjoo Park**

VAHAN BABUSHKIN
[1]Applied Interactive Multimedia (AIM) Laboratory, New York University Abu Dhabi, Abu Dhabi, 129188, UAE (e-mail: vahan.babushkin@nyu.edu)

This work was supported by the Abu Dhabi Award for Research Excellence (AARE-2019).

ABSTRACT Classifying human emotional responses from bio-sensing signals is getting more interest in applications for Human Computer Interaction. Traditionally most of the studies focus on human emotion recognition from visual stimuli. Advancements in haptics made it possible to study emotions elicited by tactile modality. In this work, we address the emotion classification from the EEG recordings of neural activations elicited by visual stimuli. The aim is to establish foundations for interpreting the emotional response from the neural activations in the presence and absence of tactile simulation by relating it to the observed visual cue which elicited emotional response is already known. We compare the performance of SVM and CNN classifiers on the recorded neural activations from visual responses initially categorized into four groups according to the circumplex model. We use subject rankings to define subcategories of four emotional categories, resulting in 9 and 12 classes. The analysis demonstrated that in general, SVM performs slightly better than CNN and both classifiers are capable of achieving high accuracies.

INDEX TERMS Machine Learning, AI, Haptics and Haptic Interfaces, Sensorimotor Learning, Virtual Reality and Interfaces

I. INTRODUCTION

HUMAN emotions play an immense role in human-computer interaction (HCI). It is absolutely essential to be able to detect and classify the emotions humans demonstrate while interacting with a virtual environment. The stimuli that elicit the emotions are multimodal and each modality contributes to the experienced emotional state. In the past the researchers were limited by only visual and audio modalities to study human emotions, however, the developments in haptics allowed to study the effect of tactile modality on human's emotional response. Since all the emotional states are controlled by the brain, the general approach towards studying emotions elicited by the tactile simulations is to analyze the Electroencephalography (EEG) signals associated with haptics interaction. The main advantage of EEG signals for emotion recognition is that it records the neural activity generated in the brain as a reaction to the external stimuli and can be directly associated with the emotional state [1]. The experiment considered here attempts to relate the EEG recordings of participants' emotional state elicited by visual stimuli to the emotional changes caused by haptic stimuli.

There is a wide range of studies investigating emotion and attention that rely on the International Affective Picture System (IAPS). The IAPS was developed at The University of Florida, by P. J. Lang et al. and represents a set of pictures with average ratings of the emotions elicited by each picture [2], [3]. The pictures are ranked according to the emotion elicitation levels quantified in terms of valence and arousal defined in the circumplex model [4]. The schematic representation of the circumplex model is given in Figure

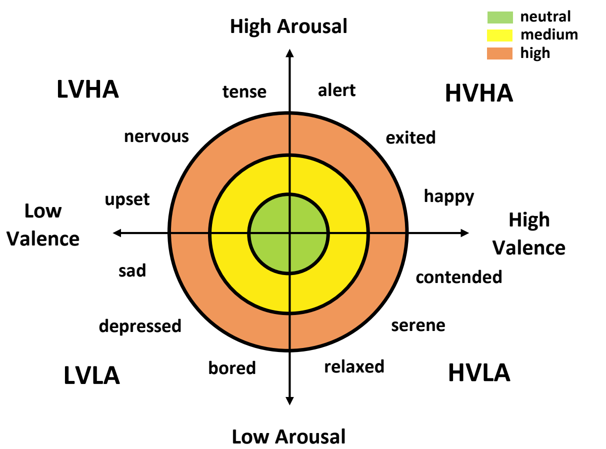


FIGURE 1: Circumplex model with 12 different classes used in this study. In this project subcategories of emotions with neutral rating have postfix 0 (HVHA_0, LVHA_0, LVLA_0, HVLA_0), with medium ranting – postfix 1 (HVHA_1, LVHA_1, LVLA_1, HVLA_1) and with high rating – postfix 2 (HVHA_2, LVHA_2, LVLA_2, HVLA_2)

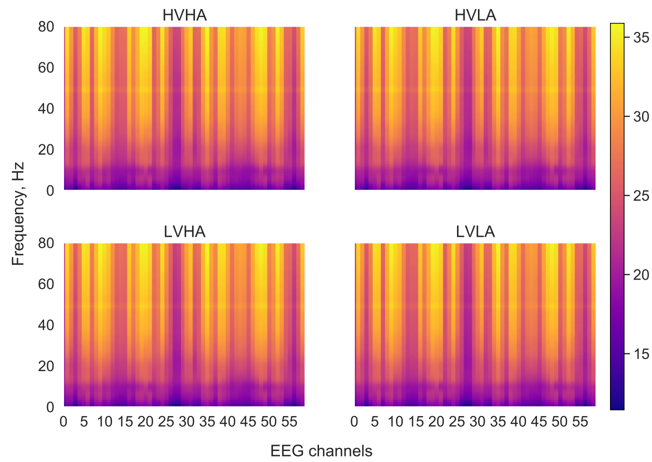


FIGURE 2: spectrogram of EEG recordings for each of the four categories, averaged over all trials and all timepoints

1, where 12 emotions are projected to the valence-arousal coordinates. The valence quantifies the extent of the pleasantness of perceiving the stimulus and the arousal represents the degree of awakeness induced by the stimulus. In most of the affective computing studies the valence and arousal are assigned high/low levels, thus distinguishing between high valence high arousal (HWHA), high valence low arousal (HWLA), low valence low arousal (LWLA) and low valence high arousal (LWHA) [5], [6], [7].

In this work, we focus on the comparative analysis of SVM and CNN classifiers for supervised classification of emotional states from the EEG recordings of neural activations evoked by the visual cues. Apart from the four emotional states according to the circumplex model we also used 9 and 12 emotional subcategories defined from the ranking of the subjects collected during the experiment. Despite heavily biased datasets in the case of 9 and 12 classes, the SVM classifier performed better than CNN, demonstrating high robustness to the classification of misrepresented instances. In the future SVM, combined with visual cues for emotional response labeling, can be used for multimodal classification of emotional states elicited by the presence and absence of tactile stimulation.

II. RELATED WORK

The study of emotions classification from EEG data is receiving growing attention from the academic community. Different machine learning algorithms such as K-Nearest Neighbor (KNN), Bayesian Network (BN), Artificial Neural Network (ANN), and Support Vector Machine (SVM) were applied to the recorded EEG data for extracting the emotional levels. In general, the accuracy of emotion recognition depends on several factors such as different experiment environments, preprocessing techniques, feature selection, etc [8]. Originally designed for binary classification problems, the Support Vector Machine (SVM) remains one of the most popular classifiers for the supervised multi-class recognition of emotions from raw EEG data. The two approaches are used while classifying emotions coming from multiple classes with SVM: one-against-one scheme, i.e. constructing the hierarchical binary classifiers each recognizing only one class of emotions and all-together scheme which perform collective optimization using all classes [9]. For example, in [9] authors used SVM for recognition of four music-induced emotional responses from EEG recordings. They were able to achieve an accuracy of 90.52% using the one-against-one scheme of a multi-class SVM classifier. In the meantime the all-together scheme achieves the accuracy of 82.37%, suggesting that the one-against-one scheme of multi-class SVM classifier should be used for emotion recognition.

Before classifying the raw EEG data some preprocessing is necessary to reduce the noise and dimensionality. In [1] the Independent Component Analysis (ICA) is applied before classifying the EEG recordings into seven emotion classes with SVM and LDA classifiers. The achieved accuracies after 4-fold cross-validation are 74.13% and 66.50% respectively [1]. Another preprocessing approach aiming to extract important statistical features by applying to the EEG signal the Empirical Mode Decomposition and then Genetic Algorithms before classifying emotions with SVM is considered in [10]. In total six classes were considered in this study, corresponding to happiness, surprise, anger, fear, disgust, and sadness. The mean accuracy for SVM has been around 85.17%. Anh et al also used SVM to classify five basic states of human emotion in real-time within the circumplex model framework [11] and reported average accuracy of 70.5%. The SVM classifier for discerning the valence and arousal in EEG data was able to achieve accuracy of 32% and 37% in [12]. A slightly higher average accuracy of 55.72% for valence and 60.23% for arousal has been achieved in

[13] applying SVM to five frequency features that were extracted from each channel of the EEG signals. Combining these features with the audio/visual features, extracted from video stimulus it is possible to achieve accuracy for 58.16% valence and 61.35% for arousal [13]. However, the accuracy of the recognition depends on the length of the dataset, for example, on large datasets SVM usually achieves 33.3% and 25% accuracy for three and four emotion categories [14]. Also, the high accuracy with the SVM classifier is achieved while recognizing emotions from the offline data. However, for real-time emotion recognition, the classification accuracy is usually low.

The growing popularity of Deep Learning resulted in a number of studies using Convolutional Neural Network for emotion recognition from the EEG data. The hidden (subsampling) layers of CNN encode primitive features in lower layers, gradually progressing towards encoding more complex features in higher layers [15]. 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. CNNs can also capture local trend and scale-invariant features for closely interrelated neighboring data points, e.g. frequency variations patterns in nearby electrodes, thus learning the relationship between emotional states and the EEG data. However, the CNN haven't achieved that wide implementation in emotion recognition from EEG data as for image recognition, due to the costly process of data collection, inconsistency and non-uniformity of the data coming from the different bio-sensing devices, as well as the need for preprocessing of the data [7].

Several studies focusing on emotion recognition with CNN apart of the EEG data use other modalities, coming from peripheral physiological signals such as GSR, electrooculogram (EOG), respiration amplitude, electrocardiogram, skin temperature, blood volume by plethysmograph and electromyogram(EMG) [16]. The authors used the partial structure of AlexNet [17] consisting of 8 parameterized layers (5 convolutional layers, 1 fully connected layer, and 1 softmax layer), to recognize two classes of emotions: arousal (with an accuracy of 87.30%) and valence (85.50%) [16]. In some sense a similar approach has been adopted by Liu et al [18] where authors used ResNets for emotion classification from raw EEG data, achieving almost 90% of accuracy with 10-fold cross-validation for classifying high/low valence and high/low arousal with an average accuracy of 58.03%. Instead of raw EEG data features, related to neurobiological processes as connectivity can be extracted and used as inputs for 2D CNN as it is done in [19] for 2,3 and 4 -classification tasks. As a rule, the lower the number of emotional states to recognize the higher the average accuracy. For example, in [19] the CNN consisting of two convolutional layers two max-pooling layers and one fully connected (dense) layer achieved the accuracy of around 85% for classifying valence and arousal, the accuracy of almost 77% for classifying positive, neutral, and negative valence/arousal and about 61% accuracy on 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) [19]. In [20] authors used CNN to recognize the same four categories as Mei et al using the EEG spectrogram and wavelet transformed galvanic skin response (GSR), achieving the accuracy of 73.43%. In this work, we use a similar approach first to recognize the four emotional states with a simple CNN model of two convolutional layers two max-pooling layers, and one fully connected (dense) layer and then extend this approach for recognizing 12 and 9 emotional states.

All previous approaches ignored the spatial arrangements of EEG cap electrodes, which is important to take into account for more accurate emotion recognition. The first attempt to incorporate topological position information into inputs to 3-dimensional CNN is done in [21], achieving recognition accuracies 87.44% and 88.49% for valence and arousal classes. The 3D CNN is capable of learning the positional and temporal features from the raw EEG data. Wang et al also preprocessed the EEG data by reshaping the 2-D matrices (channels × time samples) to 3-D tensors (2-D electrode topological structure × time samples) to provide the 3D CNN with the topological position of electrodes [22] It allowed to achieve classification rates of 73.3% and 72.1% for arousal and valence. Also, it could be useful to process the EEG data separately for the theta (4-8 Hz), slow alpha (8-10 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30+ Hz) frequency bands and consider the spectral power of all the symmetrical pairs of EEG electrodes within these five bands [23].

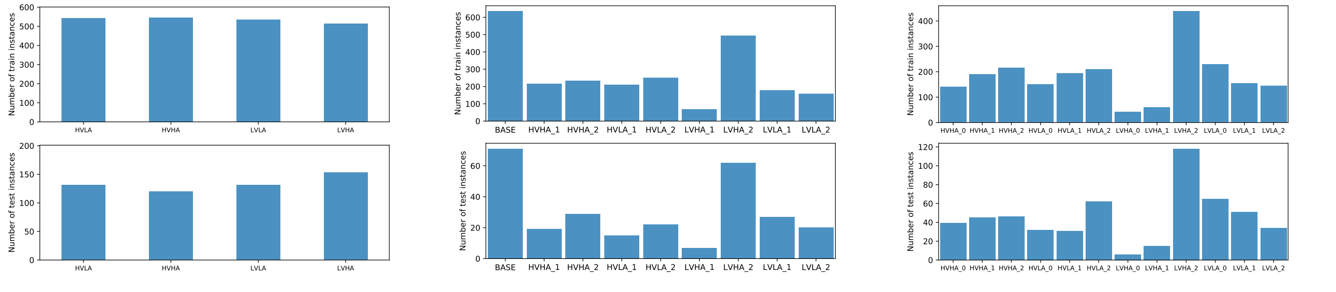
III. EXPERIEMNTAL SETUP

To infer emotional changes from haptic interaction at the beginning of the experiment the users were presented visual stimuli, consisting of IAPS images, which are able to elicit emotions, quantifiable in terms of the circumplex model. The neural response associated with these emotions was recorded with the EEG. Then the participant experienced a tactile stimulus through opening a "haptic zipper" task. Correlating the EEG recordings after the visual stimulus at the beginning of the trial and during the physical interaction (with or without tactile feedback) with a haptic touch screen allows to shed the light on the role haptic interaction changes human emotional state.

IV. METHODOLOGY

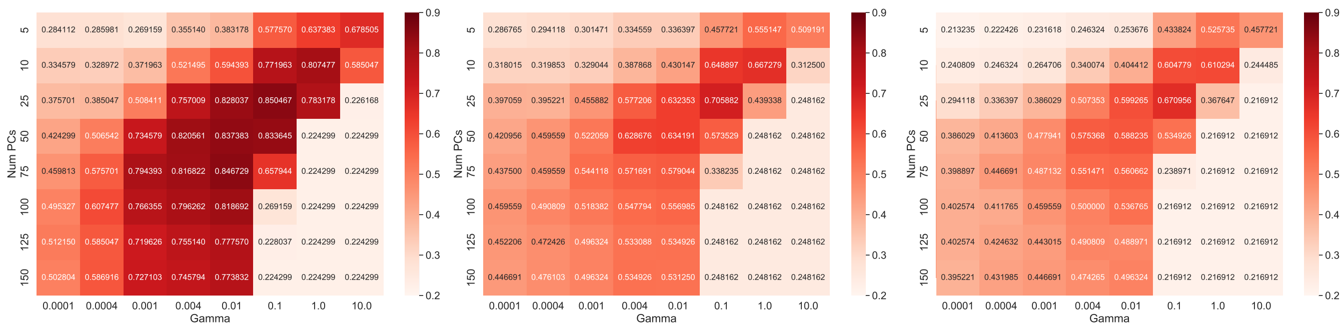
A. EXPERIMENT

For this experiment, 34 subjects were recruited. Each subject has been demonstrated 20 pictures from the IAPS database that are known to elicit positive and negative emotions among humans. These pictures were categorized according to the circumplex model, i.e. low arousal/low valence (LALV), low arousal/high valence (LAHV), high arousal/low valence (HALV), and high arousal/high valence (HAHV). The EEG response was recorded at the start of the experiment. Then, participants were asked to open a “haptic zipper” using fin-



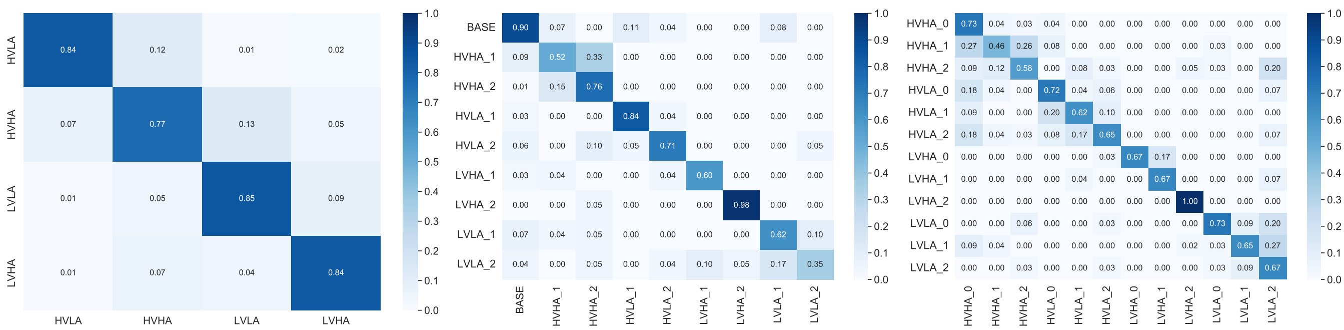
a.) b.) c.)

FIGURE 3: Distribution of instances across the classes in training and testing datasets. a.) 4 classes, b.) 9 classes, c.) 12 classes.



a.) b.) c.)

FIGURE 4: Optimal number of principal components and γ selection for SVM classifier in case of a.) 4 classes, b.) 9 classes, c.) 12 classes.



a.) b.) c.)

FIGURE 5: Average of normalized confusion matrices over 10 folds for SVM classifier in case of a.) 4 classes, b.) 9 classes, c.) 12 classes.

gertip in a controlled manner on a tactile touch screen device. After performing the haptic task, the correlation between the EEG response that was recorded at the start of the experiment to the EEG response during the physical interaction (with or without tactile feedback) with a haptic touch screen was measured. At the end of each trial the subjects were asked to rank their emotional state elicited by the visual stimulus on the scale from 0 to 9. The additional division of 4 emotional states into 12 substates was conducted based on these evaluations of the extent of the subjects' emotional states.

Initially, the experiment has been planned the way to collect an equal number of trials for each of the four categories of emotions according to the circumplex model thus each category contains approximately 680 trials. Therefore, the samples both in the training and testing datasets are equally distributed across all four categories (see Figure 3, a.)). However, when we introduce ratings for each of the category, the number of instances is no longer equally distributed between classes. The reason behind it is very simple – the subjects might find it difficult to differentiate between the neutral and medium level of some categories, for example in the case of Low Valence High Arousal it is hard to tell whether the subject is feeling neutral nervous about the scene on the picture or medium-nervous. Thus there is a non-uniform distribution of samples per each class in the case of 12 classes, varying from 48 to 556 (Figure 3, c.)). Combining all "neutral" samples (HVHA_0,HVLA_0,LVHA_0,LVLA_0) together in one class (BASE) also introduces a heavy bias to this class (Figure 3, b.)), which results in more samples classified as the instance of that class. Therefore, splitting four categories of the circumplex model into subcategories results in a heavy biased dataset.

There is little that could be done to eliminate this bias experimentally since there is no way to control the subjects' ratings rather than introducing more trials and making the experiment long and tedious for the subject. And it is the main reason why all previous studies use only two or four classes to classify the emotional states. In this sense our work is novel, since it first ustilizes subjects ranking to obtain and investigate fine resolution of emotional states.

B. SVM

Before classifying with SVM we preprocessed the EEG data tensor containing 80 freq. bins$\times$ 160 timepoints $\times$59 channels $\times$ 2671 trials to average it across the time and three frequency bins, corresponding to beta, lower gamma, and higher gamma bands (suggested by Haneen Alsuradi). Then we flattened this newly-obtained 3 $\times$ 59 $\times$ 2671 trials matrix to get a vector of length 177 for each trial. PCA analysis has been conducted on the data to identify the optimal number of principal components and the Radial Base Function (RBF) kernel coefficient γ. The experiments with 4, 9, and 12 classes demonstrated that the optimal number of principal components is 25 with $\gamma = 0.1$ (see Figure 4).

C. CNN

We used a simple CNN model for predicting 4, 9, and 12 categories from an EEG spectrogram, averaged over all 160 timepoints. The model consists of two 2-D convolutional layers each with 32 filters and 3$\times$3 kernel size and valid mode of padding. As an activation, we used the ReLu function. Each convolutional layer is followed by a 2-D max pooling layer with 2$\times$2 pooling window and there is a flattening layer at the output. The EEG data containing 80 freq. bins$\times$ 160 timepoints $\times$59 channels $\times$ 2671 trials have been averaged across the time and an additional singleton dimension has been added to the 3-D tensor of 80 freq. bins $\times$59 channels $\times$ 2671 trials to make it suitable for processing with CNN. We used 100 epochs to train the CNN model, even though as plots of average training and testing accuracies after 10 folds cross-validation suggest that the model achieves training accuracies close to 1 after 20-th epoch for all classes (Figure 6). The output of the CNN is encoded as binary labels, therefore sklearn binarizer is used to obtain the multiclass vector. In general, in comparison with fully-connected neural networks, CNNs are quite fast – it takes about 10 seconds to train each epoch with approximately 2400 training trials during cross-validation.

V. DATA

After preprocessing there are 2671 trials, approximately 680 trials for each of the four categories, each trial containing 80 frequency bins, 160 timepoints, and 59 channels. Figure 2 shows the spectrogram for each of the four categories, averaged over all 680 trials and all time points. As we can see, the differences between categories of emotional states are almost unnoticeable with the naked eye.

VI. RESULTS

In the case of 4 categories, we have an almost equal number of samples per class (about 680). Therefore, both SVM and CNN classifiers perform well on the unbiased dataset, achieving average accuracies around 0.85 for SVM and 0.81 for CNN. However, with the increase in the number of classes that also introduces a bias in the number of trials, the accuracy drops. For example, both CNN and SVM achieves accuracies of 0.70 and 0.69 for 9 and 12 classes correspondingly (Figure 8). Notice that in the case of 9 classes SVM performs better for most overrepresented classes BASE and LVHA_2 (Figure 5,b.)). For 12 classes, as expected, we are getting the highest accuracy for LVHA_2 (Figure 5, c.)). A similar trend is also noticeable in Table 1. For 9 and 12 classes SVM achieves the precision and recall above 0.9 for LVHA_2. Interestingly, in the case of 12 classes the relatively high metrics achieved for the most underrepresented class (LVHA_0) (see also Table 1). It could be explained by the low number of samples to test and the misclassification error, in this case, is also low. CNNs also achieve the highest precision, recall and F1 values for LVHA_2 in case of 9 and 12 classes (Figure 7). However, metrics for LVHA_0 are close to the other classes, meaning that CNN is capable to handle the classification of the underrepresented samples.

The heavily biased dataset for 9 and 12 classes affected the accuracy of both classifiers. We tried to mitigate the effect of the non-uniform distribution of instances in each class on learning by training both classifiers on the dataset oversampled with repetitions. There was no significant improvement in accuracy, most probably due to overfitting the model, therefore we abandoned the idea of oversampling. We did not try undersampling, since it will result in tiny datasets which might not be informative for learning emotion representations.

Overall, both classifier demonstrate the same average accuracy fluctuating around 0.85 for datatest of 4 classes with evenely distributed samples and 0.7 on the datasets containing heavily underrepresented classes (for 9 and 12 classes). The SVM demonstrated slightly better performance than CNN, it trains faster than CNN and easy to implement. 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

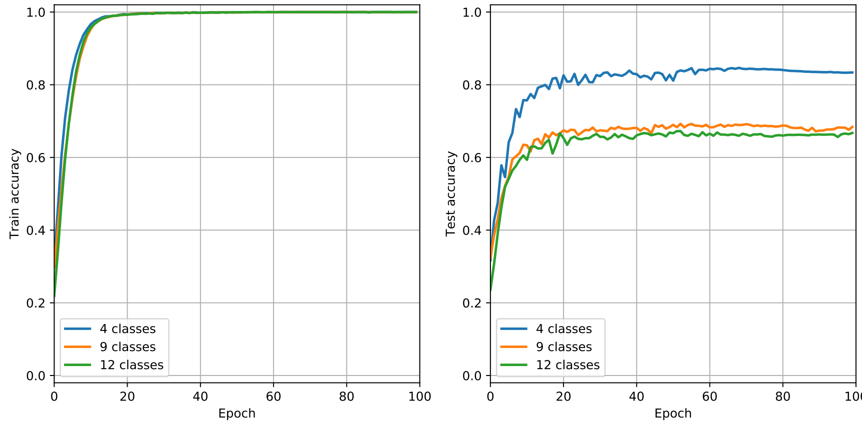
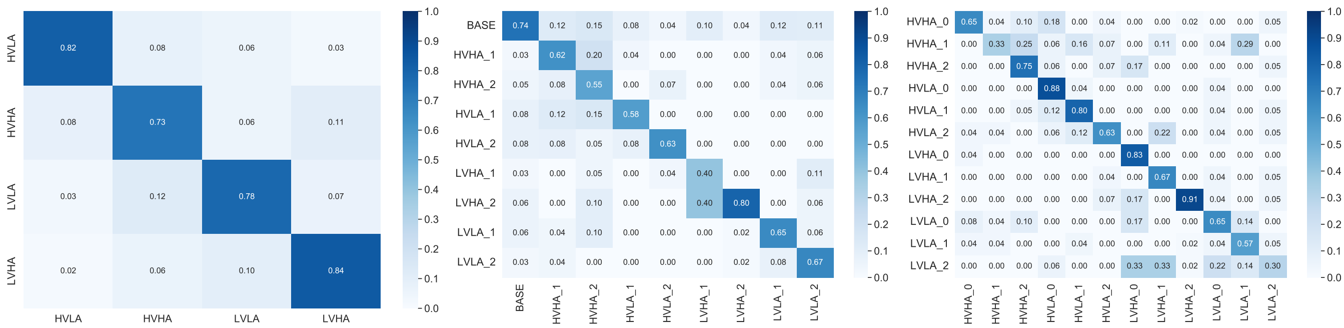
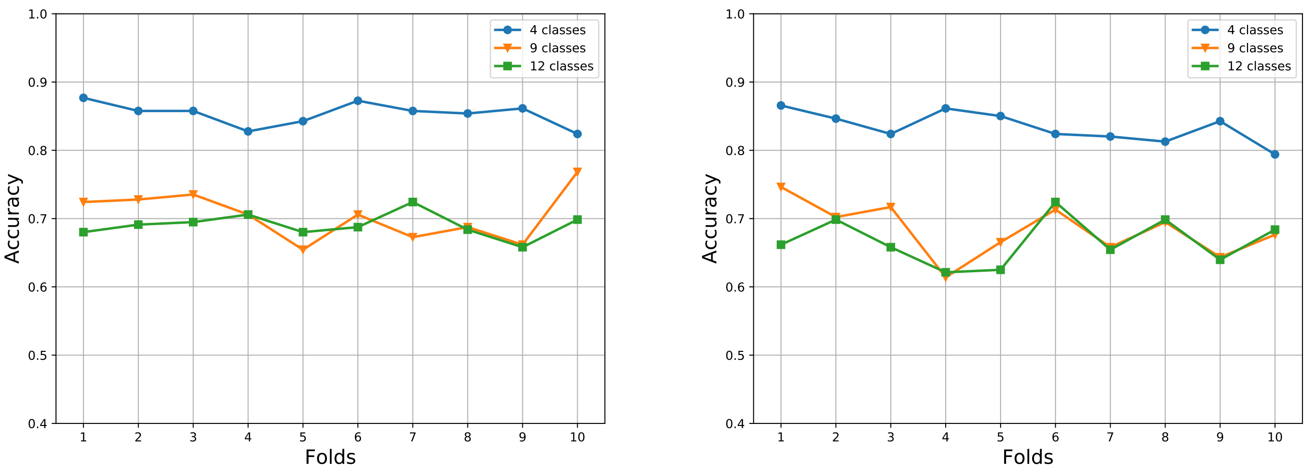


FIGURE 6: Training and validation accuracy of CNN averaged over 10 folds for 100 epochs.)



a.) b.) c.)

FIGURE 7: Average of normalized confusion matrices over 10 folds for CNN classifier in case of a.) 4 classes, b.) 9 classes, c.) 12 classes.



a.) b.)

FIGURE 8: Accuracy over 10 folds for a) SVM b) CNN classifier.

SVM. Furthermore, using PCA for dimensionality reduction in the case of SVM helps to remove noise and thus, avoid overfitting. Despite the CNN being more robust in classifying the heavily underrepresented samples, the SVM classifier could be a more preferred choice for emotion recognition from EEG recordings in the case of biased datasets.

VII. CONCLUSION

In this work, we compared the performance of two classifiers – SVM and CNN for emotion recognition from the EEG data. CNN requires little or no preprocessing of the data and is capable of inferring the hidden dependencies in the data. Both classifiers performed well on the dataset with equally distributed samples across four classes (no bias). However, SVM demonstrated better stability in achieving the same accuracies in the case of 9 and 12 subcategories, while CNN demonstrated more robustness for classifying the heavily biased data. Our results are comparable with the existing research that report an accuracy of 80% on four categories. In comparison to several works, that report accuracies ranging between 25% and 70% we were able to achieve better performance. Some works also report averages accuracies around 60% for classifying with CNN and there is also a tendency in the dropping of classification accuracy with the increasing number of categories.

For future work, it is important to reconsider the preprocessing stage and extract the most representative features which can be further purified by reducing the dimensionality with PCA. Also, the recognition accuracy might be increased by considering the multiclass classification with SVM as a binary problem, i.e. applying a one-against-one scheme. In this case, the SVM will be recognizing only one class of emotions and then the recognition results will be combined. Also, in this work we didn't take into account the spatial positioning of the electrodes, however, it might provide additional significant features for emotion recognition. The neural activity responsible for eliciting emotions from different emotional categories is captured by different electrodes and thus, analysis of the relationship between spatial arrangement of electrodes and emotional categories might shed more light on the processes behind emotional response and result in more accurate recognition.

VIII. ACKNOWLEDGEMENT

We are extremely grateful for the demos, in-class exercises, and home assignments developed by Prof. Sundeep Rangan for ECE-GY 6123 "Introduction to Machine Learning" course from where we reused most of the code for this project. Special thanks to Dr. Wanjoo Park for sharing the EEG data and Haneen Alsuradi for suggestions about how to preprocess the EEG data. It won't be possible to conduct this analysis without Dr. Park's and Dr. Eid's help and their kindly sharing the details of the experiment and the data.

REFERENCES

[1] 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), pp. 180–185, 2015.
[2] P. Lang, "International affective picture system (iaps) : Technical manual and affective ratings," 1995.
[3] M. M. Bradley and P. J. Lang, International Affective Picture System, pp. 1–4. Cham: Springer International Publishing, 2017.
[4] J. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, pp. 1161–1178, 12 1980.
[5] N. Dar, M. Akram, S. Khawaja, and A. Pujari, "Cnn and lstm-based emotion charting using physiological signals," Sensors, vol. 20, p. 4551, 08 2020.
[6] Y. Zhao, X. Cao, J. Lin, D. Yu, and X. Cao, "Multimodal emotion recognition model using physiological signals," 2019.
[7] S. Siddharth, T.-P. Jung, and T. J. Sejnowski, "Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing," 2019.
[8] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using eeg," in Foundations of Augmented Cognition (D. D. Schmorrow and C. M. Fidopiastis, eds.), (Berlin, Heidelberg), pp. 492–501, Springer Berlin Heidelberg, 2013.
[9] Y.-P. Lin, C.-H. Wang, T.-L. Wu, S.-K. Jeng, and J. Chen, "Eeg-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine," pp. 489–492, 04 2009.
[10] 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.
[11] V. Anh, M. Van, B. Ha Bang, and T. Huynh Quyet, "A real-time model based support vector machine for emotion recognition through eeg," pp. 191–196, 11 2012.
[12] R. Horlings, D. Datcu, and L. Rothkrantz, "Emotion recognition using brain activity," p. 6, 01 2008.
[13] Y. Zhu, S. Wang, and Q. Ji, "Emotion recognition from users' eeg signals with the help of stimulus videos," vol. 2014, pp. 1–6, 07 2014.
[14] P. Rani, C. Liu, N. Sarkar, and E. Vanman, "An empirical study of machine learning techniques for affect recognition in human–robot interaction," Pattern Analysis and Applications, vol. 9, 05 2006.
[15] C. C. Aggarwal, Neural Networks and Deep Learning. Cham: Springer, 2018.
[16] L. Wenqian, C. Li, and S. Sun, Deep Convolutional Neural Network for Emotion Recognition Using EEG and Peripheral Physiological Signal, pp. 385–394. 12 2017.
[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," Neural Information Processing Systems, vol. 25, 01 2012.
[18] N. Liu, Y. Fang, L. Li, L. Hou, F. Yang, and Y. Guo, "Multiple feature fusion for automatic emotion recognition using eeg signals," pp. 896–900, 04 2018.
[19] H. Mei and X. Xu, "Eeg-based emotion classification using convolutional neural network," in 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), pp. 130–135, 2017.
[20] 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.
[21] E. Shawky, R. El-Khoribi, M. Shoman, and M. Wahby Shalaby, "Eeg-based emotion recognition using 3d convolutional neural networks," International Journal of Advanced Computer Science and Applications, vol. 9, p. 329, 09 2018.
[22] Y. Wang, Z. Huang, B. McCane, and P. Neo, "Emotionet: A 3-d convolutional neural network for eeg-based emotion recognition," in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–7, 2018.
[23] S. Koelstra, C. Muhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis ;using physiological signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, 2012.

Alg.	Metrics	4 classes				9 classes									12 classes											
		HVLA	HVHA	LVLA	LVHA	BASE	HVHA_1	HVHA_2	HVLA_1	HVLA_2	LVHA_1	LVHA_2	LVLA_1	LVLA_2	HVHA_0	HVHA_1	HVHA_2	HVLA_0	HVLA_1	HVLA_2	LVHA_0	LVHA_1	LVHA_2	LVLA_0	LVLA_1	LVLA_2
CNN, 4 classes	Precision	0.84	0.81	0.82	0.87																					
	Recall	0.87	0.78	0.82	0.87																					
	F1 score	0.85	0.80	0.81	0.87																					
CNN, 9 classes	Precision					0.69	0.54	0.60	0.62	0.68	0.65	0.90	0.60	0.49												
	Recall					0.76	0.54	0.55	0.60	0.64	0.65	0.90	0.54	0.49												
	F1 score					0.72	0.53	0.56	0.61	0.66	0.58	0.90	0.56	0.47												
CNN, 12 classes	Precision														0.57	0.58	0.59	0.62	0.65	0.66	0.64	0.61	0.88	0.70	0.57	0.48
	Recall														0.55	0.53	0.61	0.60	0.64	0.65	0.65	0.62	0.90	0.68	0.57	0.43
	F1 score														0.55	0.53	0.61	0.60	0.64	0.65	0.65	0.62	0.90	0.68	0.57	0.43
SVM, 4 classes	Precision	0.84	0.82	0.86	0.90																					
	Recall	0.92	0.91	0.90	0.81																					
	F1 score	0.86	0.81	0.85	0.90																					
SVM, 9 classes	Precision					0.67	0.54	0.66	0.68	0.72	0.70	0.92	0.62	0.51												
	Recall					0.77	0.51	0.65	0.57	0.69	0.56	0.93	0.57	0.45												
	F1 score					0.71	0.53	0.65	0.61	0.70	0.61	0.92	0.59	0.47												
SVM, 12 classes	Precision														0.64	0.54	0.65	0.60	0.64	0.70	0.81	0.72	0.91	0.72	0.58	0.50
	Recall														0.68	0.55	0.63	0.64	0.59	0.69	0.80	0.61	0.95	0.70	0.58	0.48
	F1 score														0.65	0.54	0.64	0.62	0.60	0.69	0.77	0.63	0.93	0.71	0.58	0.48

TABLE 1: Evaluation metrics averaged over 10 folds.