Emotion Detection Using EEG Data

Tina Gholami

Electrical and Computer Engineering Department, University of Western Ontario
London, Canada

tgholam2@uwo.ca

Abstract*— In The electroencephalogram (EEG) is a dynamic non-invasive, and relatively inexpensive technique used to monitor the state of the brain. EEG signal analysis has attracted significant research attention because of its number of clinical uses that range from monitoring normal wakefulness or arousal states to complex clinical situations involving seizure or coma. Other applications include detecting any stimulus on the brain and predicting the induced range of emotions. This can be directed towards enhancing the playing experience for the gamers. That is why different research organizations have tried to record EEG data within various tasks. One of these popular datasets is called the DEAP dataset for emotion detection [1]. In recent years, deep learning classification-based approaches have demonstrated modest success due to their ability to model complex systems. However, since the DEAP dataset's targets have continuous values, the papers have adopted different threshold values and methods to better classify emotions. Nevertheless, such diversity makes it harder for the methods to be compared on all target labels. Additionally, accurate continuous values for each target label cannot be calculated. This issue is highlighted in that such delicacy exists for values near the threshold so that with a slight nuance (such as the model error), it will be improperly categorized. Moreover, Tripathi et al [2] showed the accuracy of such a classification model is no greater than 60%-75%. Consequently, the current paper examines a regression-based method to provide better insight into the ultimate predicted emotion based on the analysis of the continuous target labels. Comparing the convolutional neural network (CNN) different parameters with the previous papers, the observed results show an improvement of about 10% on the accuracy depending on the target label and the model. The advantage of a regression-based model is that it is less prone to the threshold error, and it can be more accurately analysed to produce a more reliable emotion.

***Keywords*— EEG signal analysis, DEAP dataset, emotion detection, classification-based approaches, regression-based method, convolutional neural network, threshold error**

I. Introduction

Our emotions play an essential role in our decisions, the satisfaction we feel towards a product and our interactions and cognitive process. As the growing technology is trying to better understand our emotions, there are widespread devices such as Muse or EEG caps that can catch our biological data (such as the brain's EEG signals in the latter case) to predict our inner emotions automatically. Such analysis can be beneficial in both marketing and health care sectors to help people stay mentally healthy and improve customer satisfaction innovatively. There have been successful research breakthroughs on emotion recognition using text, tone, or facial expression by real-time streaming data. However, one of the new directions in emotion recognition is using EEG-based systems as it becomes more affordable. Which has led it to become one of the most potent technologies to give insight into the state of mind. This paper focuses on predicting a subject's emotion from EEG signals through supervised learning models on the DEAP dataset [1]. Towards this end, CNN has been studied in regression tasks, using various parameter tunings and time-series pre-processing methods such as windowing, to come up with the best possible performance. Neural networks are inspired by how the brain performs while learning tasks, introducing the key concepts present in the brain, such as non-linearity and complexity. Furthermore, recent research shows the promising results such machine learning models can bring about in predicting sentiments and detecting hidden patterns within images. The DEAP EEG dataset [1] is a multimodal dataset for the analysis of human affective states. The electroencephalogram (EEG) and peripheral physiological signals of 32 participants were recorded as each watched 40 one-minute-long excerpts of music videos. Participants rated each video in terms of arousal levels, valence, like/dislike, dominance and familiarity. This indicated four different target labels, each having a continuous number in the range of 1 - 9. The device that recorded the EEG signals contained 40 electrodes (aka, channels) while the 40 participants were exposed to the music videos. Aside from that, 22 of the 32 participants frontal face video was also recorded. However, only the EEG data was used in this paper. There are two available versions of the DEAP dataset: the pure noisy version, which requires signal processing techniques to be used in a machine learning task. Alternatively, the other version used in this paper contains a down-sampled (to 128Hz), pre-processed and segmented version of the data in Matlab and pickled python formats. This version of the data is well-suited for data scientists for use in a machine learning task without the hassles of applying signal processing techniques separately. To be more precise, compared to the purse version of the dataset, the second version was down-sampled to 128Hz, EOG artifacts were removed, a band-pass frequency filter from 4.0-45.0Hz was applied, the data was averaged to the common reference, the EEG channels were reordered so that they all follow the Geneva order, the data was segmented into 60-second trials and a 3-second pre-trial baseline removed, the trials were reordered from presentation order to video order More insight into the dataset is presented in Figure 1.

Convolutional neural network tunning is used in this paper to predict the test data in a regression task, and since the data is 3D, CNNs are one of the best options to use. Previous works have shown satisfactory performance in classifying the EEG data Tripathi et al [2], Djemal et al [4]. However, none have tried to use regression. Additionally, in the case of Tripathi et al [2], only the first two targets (valence and arousal) were analysed. The advantage of regression is that the results can be further analysed to perform classification. If needed, the classification thresholds can finally be applied to the regression results to classify them into the desired categories. That way, even if there are huge values in the regression predictions created by the training error, they can be included in the last category of the classification.

Nevertheless, the issue with a classification approach is that it cannot be mapped to continuous values (like the results achieved in regression) for a more accurate analysis. Besides, thresholding the training targets is a rather sensitive task since a slight nuance created by the training error can classify test data into an inappropriate category. Moreover, while performing regression, higher accuracies were achieved. When performed on the pre-processed data after applying windowing, the CNN model showed better performance, achieving accuracies of 63.11%, 82.11%, 74.44%, and 77.89% for valence, arousal, liking, and dominance, respectively. Whereas Tripathi et al [2] presented a deep neural net model with achieved accuracies of only 75.58% and 73.28% for valence and arousal respectively for classification on two classes (high and low); 58% and 54% for valence and arousal, respectively for classification on three classes (high, normal and low). The CNN model presented by Tripathi et al [2] also only performed with 81.41% and 73.36% for 2 class classification and 66.79% and 57.58% for three classes on valence and arousal, respectively. Furthermore, Djemal et al [4] struggled to reach 75% and55% on classification involving two and three classes, respectively. This shows how this paper's CNN model surpassed the previous models using windowing and appropriate vigorous tunings. The rest of the paper is organized as follows: Section II describes the background, and section III reviews related work. Section IV presents the methodology, Section V explains the experiments and discusses the corresponding results. Finally, Section VI concludes the paper.

Array name	Array shape	Array contents
data	40 x 40 x 8064	video/trial x channel x data
labels	40 x 4	video/trial x label (valence, arousal, dominance, liking)

Fig. 1 DEAP dataset characteristics[1]

II. Background

In this section, some background information on the methods used is provided:

A. *Convolutional Neural Networks (CNNs)*

Convolutional neural networks or CNNs are a class of deep neural networks commonly applied to image data. The architecture of these networks consists of an input and an output layer. In between, there are convolutional layers that convolve the input images with specific kernels (filters) to abstract the data into a feature map. There are also dropout layers to help combat overfitting by randomly subsampling the data at the corresponding layers. Then, there are max-pooling layers to abstract the data into smaller matrices even further in order to recognize vital features in each input matrix data. Finally, after repeating the desired number of such layers, the data should be flattened into a 1D vector capable of being fed into a deep neural network to be analysed.

B. *Feed Forward Deep Neural Network (DNN)*

Deep neural networks (DNNs) have recently become the standard tool for solving various computer vision problems. They are heavily resource-intensive, all the more when ensembles of multiple models are involved. DNNs were inspired by information processing and distributed communication nodes in biological systems. They are based on a collection of connected units called artificial neurons (analogous to biological neurons in the brain). Typically, neurons are organized in layers. Different layers may perform different kinds of transformations on their inputs. DNNs can model complex non-linear relationships since their architectures generate compositional models. The extra layers enable the composition of features from lower layers, potentially modelling complex data with fewer units than a similarly performing shallow network.

C. *Recurrent Neural Networks (RNNs):*

A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time-series data. These deep learning algorithms are commonly used for ordinal or temporal problems, such as language translation, natural language processing (NLP), speech recognition and image captioning. They are incorporated into popular applications such as Siri, voice search and Google Translate. Like DNNs and CNNs, recurrent neural networks utilize training data to learn. They are distinguished by their "memory" as they take information from prior inputs to influence the current input and output. While traditional deep neural networks assume that inputs and outputs are independent of each other, the output of recurrent neural networks depends on the sequence's prior elements. While future events would also help determine the output of a given sequence, unidirectional recurrent neural networks cannot account for these events in their predictions.

III. RELATED WORKS

Before releasing the DEAP dataset, which helped the growing community of HCI researchers in emotion recognition, most of the studies were based on facial expression videos or speech recognition to get insight into emotion. However, one of the downsides of such data is that it requires massive memory storage to be processed and analysed. Additionally, no related data was available for EEG brain signals. But recently, there has been more research done to produce more EEG data, such as DEAP. One other popular EEG dataset is SEED [5]. DEAP uses Russells valence-arousal scale, widely used in research on affect, to quantitatively describe emotions. In this scale, each emotional state can be placed on a 2D plane with arousal and valence as the horizontal and vertical axes. Arousal can range from inactive (e.g., uninterested, bored) to active (e.g., alert, excited). Whereas valence ranges from unpleasant (e.g., sad, stressed) to pleasant (e.g., happy, elated). DEAP has the highest number of participants in publicly available databases for analysis of spontaneous emotions from physiological signals. In addition, it is the only database that uses music videos as emotional stimuli. Djemal et al [4] used ResnetV2 in a classification task, where an average accuracy of about 57% and 62% was achieved. Tripathi et al [2] only checked the valence and arousal target labels. Each users' data was classified into two (high and low) and three (high, normal, low) classes for valence and arousal using DNN and CNN models. Their CNN model had 81.41% and 73.36% accuracy for two-class classification and 66.79% and 57.58% for three-classes on valence and arousal respectively as shown in Table 1. Candra et al [7] investigates how the window size affects the classification of DEAPs, EEG data using wavelet entropy, and SVMs.They conclude that an overly broad window can lead to information overload, which causes the features to be mixed up with other information. Similarly, the information about emotion might not be adequately extracted if the time window is too short. They then use the popular discrete wavelet transform (DWT) coefficient for extracting time-frequency domain features in EEG signals. Their investigation revealed that arousal could be classified up to 65.33% accuracy using a window length of 310 seconds, while valence can be classified up to 65.13% accuracy using a window length of 312 seconds.

IV. METHODOLOGY

This section explains data pre-processing in preparation for neural network models. The main step is the sliding window technique. Also, no missing value was detected in the data. Next, neural network models will be studied.

A. Sliding Window Technique:

As seen in Figure 1, each of the forty people watched forty randomly chosen music videos, while 8064 samples of EEG data were recorded through forty electrodes. Therefore, for each music video, there are samples of size (40, 8064), meaning 40 rows and 8064 time-based features. However, each sample should be transposed to (8064, 40) to have the time dependency in the rows and the 40 electrode samples in the features. A sliding window can then be applied to the data, with each window sliding along the rows (aka the time) with a specific stride and window size. Using the windowing operator, we can convert a time series problem into a machine learning problem. This method allows us to use all the additional tools and techniques to train and optimize models. This step is also considered to lie in the scopes of feature engineering, as by applying windowing, not only is the model far better able to find patterns within the data, but also it becomes easier to train the model as windows in this case function similarly to batches. Hence, the data is broken down into chunks (aka, sub-samples), and the targets should accordingly increase as well in order to maintain the same size as the features. That is why corresponding targets for each music video were repeated for each music video's sub-samples.

B. Splitting the Data into Train and Test Sets:

If a naive split (such as choosing the first thirty-two music videos as the train set and the remaining eight as the test set) is performed on the data, the scale of the validation loss gets much larger than that of the training loss because the test data is significantly different from the training data (about 200 to even 500 times). To tackle this issue, each music video should contain both train and test data simultaneously. That is why two methods have been employed for train-test-split. One is to choose the first 80% of the sub-samples in each music video as the training set and the remaining 20% as the test set, and then repeat. This process is repeated for all forty music videos for each person. However, this method did not show promising results when the model was being trained. The issue with the large validation loss scale was still present. Hence, a second method was employed, which first mixes up all the sub-samples from all 40 music videos together before shuffling them. Next it randomly chooses 80% of them as the training set and the other 20% as the test set. Implementing the second technique reduced the difference within the train-test scales. Introducing more randomness while splitting the data is actually beneficial in the results. This technique was applied on both CNN and DNN models. However, in the case of the RNN model, the sliding window technique was not used as the LSTM layer only accepts 2D data, whereas the windowing technique would result in 3D input vectors.

C. **Scaling the Data:**

Next, the data was scaled using z-standardization which uses mean and standard deviation of the data samples. This is a crucial step since if a feature in the dataset is significant in scale compared to others, it can dominate others. As a result, predictions of the neural network will not be accurate. Also, the front propagation of neural networks involves the dot product of weights with input features. Hence, if the values are too large, the calculation of the output will take a lot of computation time and memory. The same case occurs with backpropagation. Consequently, the model converges slowly if the inputs are not normalized. At the same time, the model was run using (i) no scaling on the targets, (ii) a min-max scaler, and (iii) z-standardization. However, the best and most stable results were achieved with the non-scaled targets. So, method (i) was in general applied in this paper.

D. Reshaping the Data for the Model:

Since each windowed sub-sample is 3D and a CNN was applied to the data, then each input that is going to be fed into the model has to have the proper shape. That is why the data should be reshaped into an array of three dimensions instead.

E. Convolutional Neural Networks (CNN):

Two convolutional layers were used (each having 128 and 64 neurons with kernels of size 3*3), intermingled with two Max Pooling layers of pool size 2*2. Next, the output was flattened, and the data was fed into a neural network of three layers (each having 128, 64, and 32 neurons) following by two dropout layers 0f 0.5 after each of the first two dense layers. The output layer was considered to have (i) four (for multiple-label regression) and (ii) one (for single-label regression) neurons, respectively. The mean absolute error (MAE) was used as the loss function as this paper reclassified the problem into a regression problem. Different optimizers Adam, SGD, and RMSprop, were used along with various flat learning rates of 0.001 and 0.0001. Batches of size 64 were employed. Moreover, since the focus in this paper was to find a good regression model for DEAP EEG recordings, only the data for the first subject was used as the benchmark. To have a more general result, the model can be applied to any number of participants involved and calculate the average error. Finally, to evaluate the model performance, MSE and MAE metrics were used. The regression accuracy was calculated based on the MAE metric.

V. RESULTS

In this section, the results of different model configurations are presented. The CNN model is broken down into a multi-label versus a single-label format. The number of epochs proved to make a considerable difference in the results. Initially, 50 epochs were tested using CPU, but since that did not provide sufficient speed to train the data, 200 epochs were used with GPU instead. However, as seen in Figure 2, any number of epochs ranging from 200 to about 320 should be acceptable to use

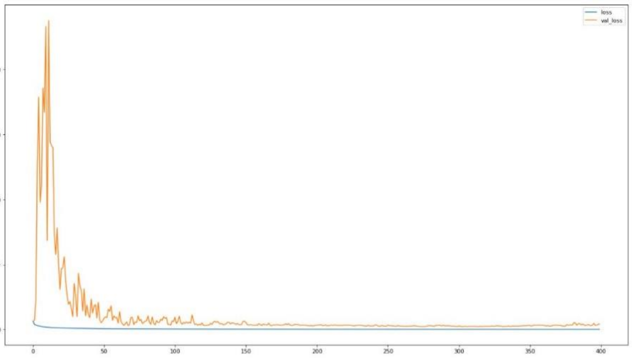

Fig. 2 Optimal number of epochs

The CNN model was broken down into two formats: multiple-label regression (predicting all four targets simultaneously) and single-label regression (predicting one single target at a time). Different optimizers Adam, SGD, and RMSprop, and learning rates of size 0.001 and 0.0001 were considered. The results are presented in tables 2-7. As can be seen in the tables, for most of the optimizers, especially those with better performance, such as Adam(0.001) and Adam(0.0001), the single-label regression has produced better results with lower error (both MAE and MSE) and higher accuracy compared to the multiple-label regression. Moreover, among the top two optimizers, Adam(0.001) and Adam(0.0001), Adam(0.001) is doing an even better job with accuracy 72.44%, 82.22%, 85.55%, and 89.89% for valence, arousal, dominance, and liking respectively. Compared to the previous studies' results, this model surpasses the previous works in table 1 in regard to both valence and arousal, except for the valence target presented by Tripathi et al [2], which uses a CNN classification model with two output classes. This valence accuracy was about 81.40%, whereas the current paper has achieved 72.44%. Moreover, on the arousal target, Tripathi et al [2] achieved an accuracy of 73.36%, whereas the current paper has an arousal accuracy of 82.22%. Furthermore, none of the previous works in table 1 tried to predict the other two targets, dominance and liking, whereas this paper has shown an extensive analysis of them through tables 2-7.

TABLE 1
RESULTS OF THE RELATED WORKS

Metric / Labels	Valence	Arousal	Dominance	Liking
Tripathi et al [2] CNN Accuracy with 2 classes	81.40%	73.36%	None	None
Tripathi et al [2] CNN Accuracy with 3 classes	66.79%	57.58%	None	None
Chung, Youb, and Yoon[8] with 2 classes	66.60%	66.40%	None	None
Chung, Youb, and Yoon [8] with 3classes	53.40%	53.00%	None	None

TABLE 2
RESULTS OF THE REGRESSION CNN MODEL, ADAM(0.001)

Metric/Labels	Valence	Arousal	Dominance	Liking
MAE multi-label	2.38	1.69	1.64	1
MSE Multi-Label	9.49	5.7	5.22	1.76
Accuracy	73.55%	81.22%	81.77%	88.88%
MAE Single-Label	2.48	1.6	1.3	0.91
MSE Single-Label	9.57	5.18	3.48	1.5
Accuracy	72.44%	82.22%	85.55%	89.89%

TABLE 3
RESULTS OF THE REGRESSION CNN MODEL, ADAM(0.0001)

Metric / Labels	Valence	Arousal	Dominance	Liking
MAE Multi-Label	3.79	2.55	2.41	1.42
MSE Multi-Label	23.28	11.41	10.56	3.73
Accuracy	57.88%	71.66%	73.22%	84.22%
MAE Single-Label	3.91	1.97	3.2	1.9
MSE Single-Label	30.48	7.12	14.03	11.43
Accuracy	56.55%	78.11%	64.44%	78.88%

TABLE 4
RESULTS OF THE REGRESSION CNN MODEL, SGD(0.001)

Metric / Labels	Valence	Arousal	Dominance	Liking
MAE Multi-Label	16.74	10.05	13.04	6.52
MSE Multi-Label	1418.07	440.5	820.72	299.76
Accuracy	too off	too off	too off	27.55%
MAE Single-Label	6.89	6.25	4.94	3.25
MSE Single-Label	82.96	56.4	62.66	18.98
Accuracy	23.44%	30.55%	45.11%	63.88%

TABLE 5
RESULTS OF THE REGRESSION CNN MODEL, SGD(0.0001)

Metric / Labels	Valence	Arousal	Dominance	Liking
MAE Multi-Label	19.05	23.5	24.84	24.71
MSE Multi-Label	635.26	920.01	1037.01	1031.76
Accuracy	too off	too off	too off	too off
MAE Single-Label	66.34	17.09	28.75	3.92
MSE Single-Label	7340.73	513.04	1384.77	50.99
Accuracy	too off	too off	too off	56.44%

TABLE 6
RESULTS OF THE REGRESSION CNN MODEL, RMSPROP(0.001)

Metric / Labels	Valence	Arousal	Dominance	Liking
MAE Multi-Label	4.26	2.64	4.51	1.89
MSE Multi-Label	33.7	13.06	49.67	10.15
Accuracy	52.66%	70.66%	49.88%	79.00%
MAE Single-Label	14.59	2.68	2.01	7.6
MSE Single-Label	393.09	11.56	7.8	115.4
Accuracy	too off	70.22%	77.66%	15.55%

TABLE 7
RESULTS OF THE REGRESSION CNN MODEL, RMSPROP(0.0001)

Metric / Labels	Valence	Arousal	Dominance	Liking
MAE Multi-Label	4.91	4.09	7.14	2.39
MSE Multi-Label	43.43	26.5	103.04	8.37
Accuracy	45.44%	54.55%	20.66%	73.44%
MAE Single-Label	4.81	14.31	14.91	7.1
MSE Single-Label	99.54	354.79	366.02	103.1
Accuracy	46.55%	too off	too off	21.11%

As can be seen in the tables, table 2 presents the best results so far. Figures 4-7 Show the test targets versus the predictions for Table 1 results, single-label prediction. As seen in the figures, the predictions (in red) are well able to track the test data (in blue). Figure 3 presents the validation loss against the training loss for the valence target in table 1. The other targets' loss plots were similar to figure 3. That is why only the plot of the valence target is included in this paper.

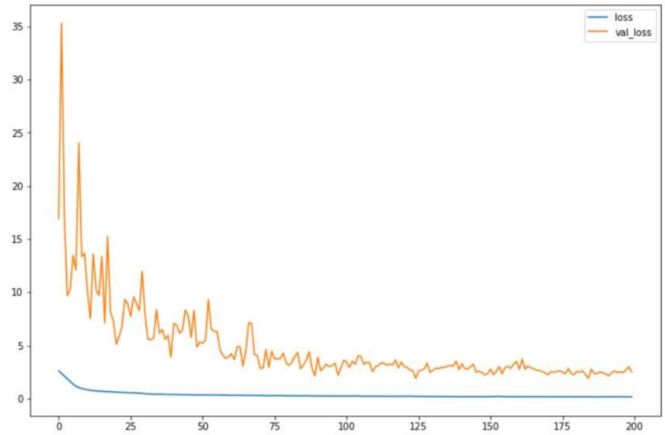

Fig. 3 Training loss (in blue) versus the validation loss (in orange), for the valence target, single-label regression, CNN model, Adam(0.001)

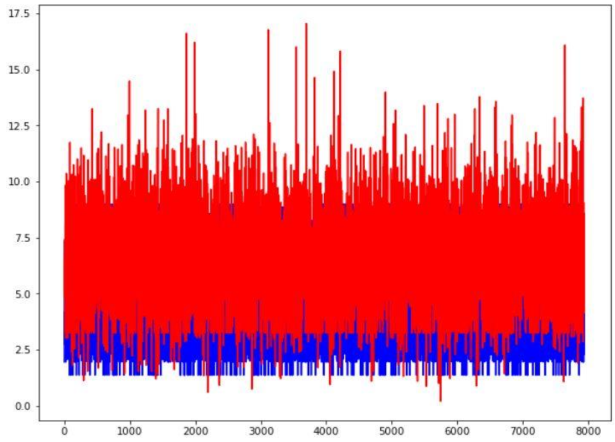

Fig. 4 Test targets (in blue) versus the predictions (in red) for the valence target, single-label regression, CNN model, Adam(0.001)

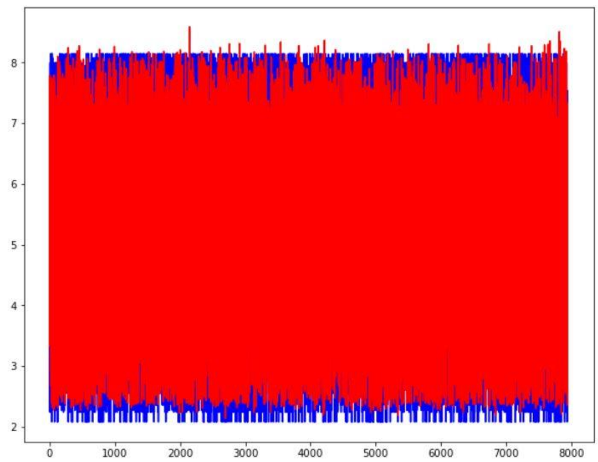

Fig. 5 Test targets (in blue) versus the predictions (in red) for the arousal target, single-label regression, CNN model, Adam(0.001)

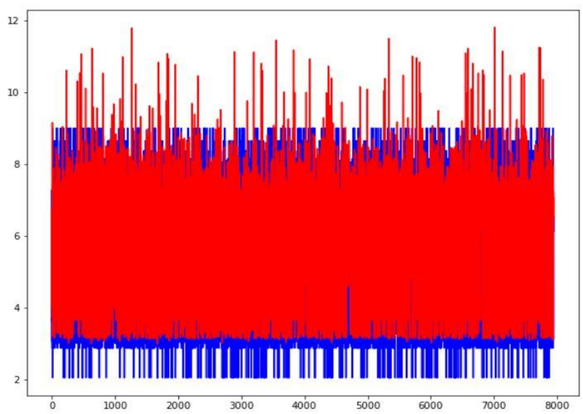

Fig. 6 Test targets (in blue) versus the predictions (in red) for the dominance target, single-label regression, CNN model, Adam(0.001)

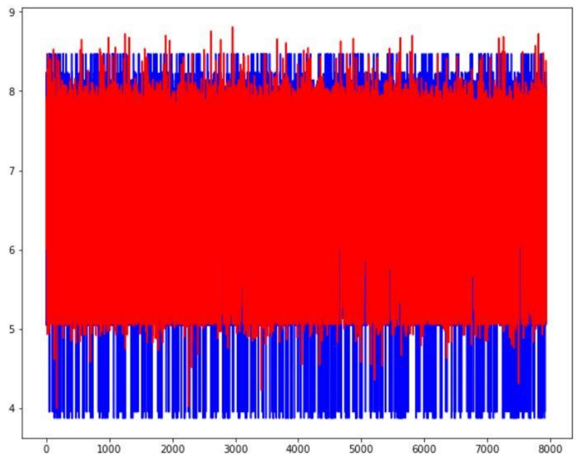

Fig. 7 Test targets (in blue) versus the predictions (in red) for the liking target, single-label regression, CNN model, Adam(0.001)

VI. Discussion on other models

Aside from the proposed CNN model, two other modes DNN and RNN have been considered for their performance. The DNN model had an architecture of three hidden layers with 1000, 500, 250 neurons respectively. Dropout layers were employed as a means of regularization. Adam optimizer with learning rate of 0.001 was employed as well. The MSE loss on the test set was 3.34. The RNN model contained four LSTM layers, each having 50 neurons. Adam(0.001) was used as the optimizer. Four Dropout layers of rate 0.2 were employed here as well. Table 8 is the results for the RNN model. But this is just the early results which were inconsistent because the validation loss was slightly smaller than the training loss when the model was training. Such cases are called an “unknown fit”. To get a better understanding, this case should be studied more extensively. However, Alhagry. Aly Fahmy, A. El-Khoribi [9] presents ideas for applying regression RNN on the EEG data.

TABLE 8
Results of The Regression RNN model, Adam(0.001)

Metric / Labels	Valence	Arousal	Dominance
MAE Single-Label Regression	0.12	0.13	0.12
MSE Single-Label Regression	0.26	0.31	0.24
Accuracy	98.66%	98.55%	98.66%

VII. Conclusion

In this paper, various configurations on emotion detection for the CNN regression model was studied to find a model that can surpass the accuracy in the previous works on the EEG data (more specifically, the DEAP dataset). Adam, SGD, and RMSprop optimizers with flat learning rates 0.001 and 0.0001 as well as single-label versus multiple-label regressions were studied. Tables 2-7 indicates the best model to have Adam optimizer with learning rate 0.001. The model architecture was kept similar for all the tables' results, which is more thoroughly described in parts IV, and V. The previous works mostly employed classification on the DEAP dataset, whereas this paper aims to show that a tuned regression-based approach shows promising results as well. Furthermore, as discussed in section VI, other neural network models such as RNN and DNNs were considered too as suggestions for future works. RNN model, in particular, shows considerable improvements over the proposed CNN model but to come up with a firm result, more studies have to be conducted on the RNN model.

References

[1] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, Ioannis Patras, DEAP: A Database for Emotion Analysis using Physiological Signals, IEEE, Volume 3, Issue 1, March 2012

[2] Samarth Tripathi, Shrinivas Acharya, Ranti Dev Sharma, Sudhanshi Mittal, Samit Bhattacharya, Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset, Twenty-Ninth AAAI Conference on Innovative Applications (IAAI-17), 2017

[3] Davoud Gholamiangonabadi, Nikita Kiselov, Katarina Grolinger, Deep Neural Networks for Human Activity Recognition with Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection, IEEE, Volume 8, July 2020

[4] Ridha Djemal, Ayad G. Bazyed, Kais Belwafi, Sofien Gannouni, Walid Kaaniche, Three-Class EEG-Based Motor Imagery Classification Using Phase-Space Reconstruction Technique, Brain Science (MDPI), 2016

[5] https://bcmi.sjtu.edu.cn/home/seed/

[6] https://www.bitbrain.com/blog/ai-eeg-data-processing

[7] Henry Candra, Mitchell Yuwono, Rifai Chai, Steven W Su, Investigation of Window Size in Classification of EEG-Emotion Signal with Wavelet Entropy and Support Vector Machine, The 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, 2015

[8] Chung, Seong Youb, and Hyun Joong Yoon, Affective classification using Bayesian classifier and supervised learning, 12th International Conference on Control, Automation and Systems (ICCAS), IEEE, 2012

[9] Salma Alhagry, Aly Aly Fahmy, Reda 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, 2017