EE798R Project 1 Report
ERTNet: an interpretable transformer-based framework for EEG emotion recognition

Devansh Ojha

October 22, 2024

1 Implementation Details

I worked on the code provided by the authors at https://github.com/ruix6/ERTNet
I implemented the code in a python notebook uploaded athttps://github.com/DeVaNsHoJhA/ERTNet_IPR/tree/main, which implements the ERTNet model proposed by the paper and four other models for comparison, namely, EEGNet (Lawhern et al., 2018), DeepConvNet, and ShallowConvNet (Schirrmeister et al., 2017) and CNN-BiLSTM(Zhang et al., 2020).

The training was done using the free gpu python notebook instance on kaggle.

1.1 Dataset Details

I worked with the DEAP datasetP (Koelstra et al., 2012), which contains labeled EEG signal data for 32 test subjects. Each subjects dataset contains 32 channels of EEG data, which are used for testing the architecture. Each data file has data for 63s, sampled at 128hz, which is too large for the model. So each time-series data is segmented into non-overlapping 4s segments for training.

The paper also suggested training on other datasets like SEED-V(Liu et al., 2022), but I couldn't ddo it due to computational and data-size contraints.

1.2 Training and Evaluation Method

For **subject dependent** testing, 10-fold training was done on the ERTNet, but for the other four comparision models training with a 1 : 9 test-train split was done due to computational contraints.

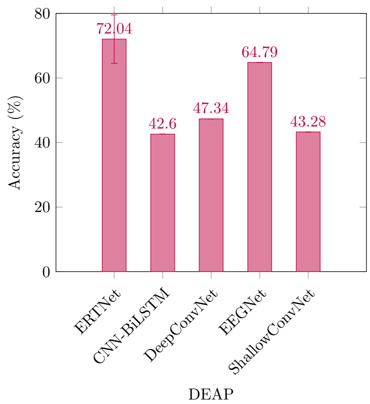
For **subject independent** testing, Leave-One-Subject-Out Cross Validation (LOSOCV) was implemented, i.e. leaving out data of one test subject from training and using it as test data.

For both subject dependent/independent testing, the hyper parameters used were the optimal parameters given in the paper, not acquired by bayesian optimization (but the code is present) as given in the paper, due to computational limitation of the kaggle platform.

2 Results

2.1 Subject Dependent Training Results

Note: that the other four models have not been 10-fold trained due to computational contraints, thus don't have an error bar.



Accuracy of Subject-Dependent DEAP

We can see that these results fairly match the results given in the paper.

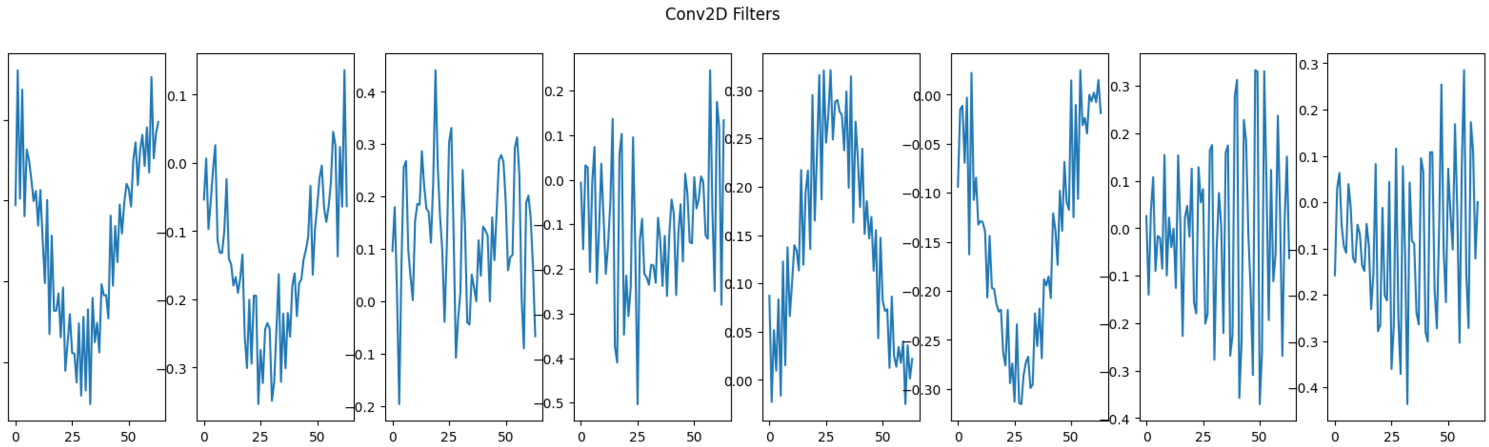
2.2 Subject Independent Training Results

Note: that the other four models have not been trained due to computational contraints (but the code is present), thus don't have any data.

$$Accuracy = 33.4 + -(7.4)\%$$

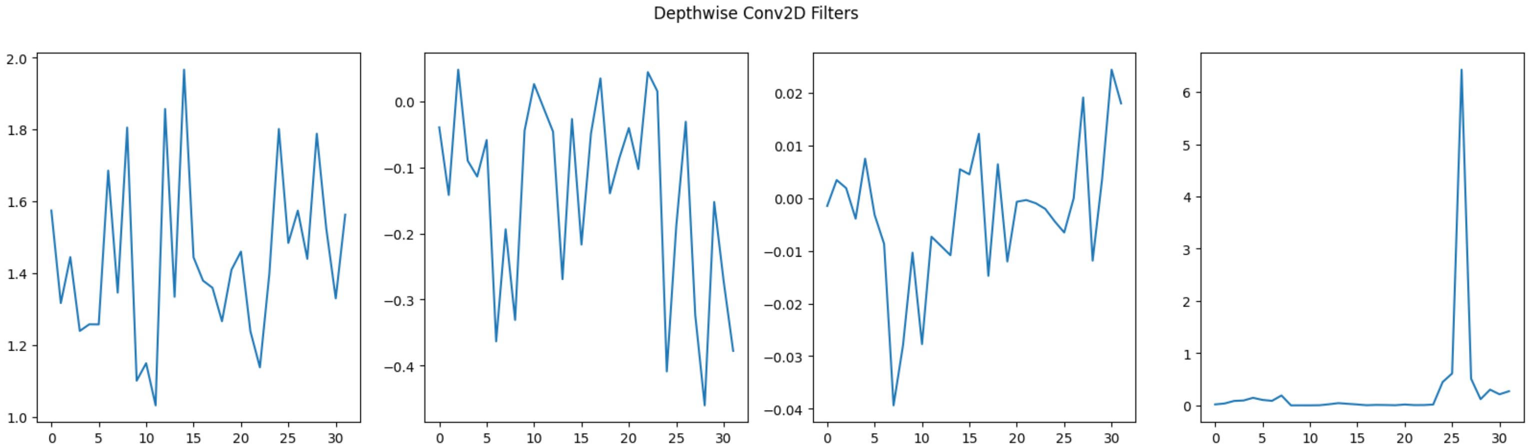
2.3 interpretability of the Model

Visualizations of the different convolution kernels learned by the ERTNet model are made to interpret the different activations for each kernel. I made visualization for 2 layers, the 8 temporal 1D-CNN layer and 4 saptial DepthWise Convolution layer.



8 Temporal CNN kernels

The 8 temporal convolution kernels show what artifacts in the EEG signals that the corresponding kernels activates to



4 spatial CNN kernels

The 4 spatial convolution kernels show which relation between the 32 channels of the data activates this layer

3 Acknowledgement and references

I would like to acknowledge prof. Sandhan and Ms. Swathi Pratapa for helping me in this project.

3.1 References

- Liu R, Chao Y, Ma X, Sha X, Sun L, Li S and Chang S (2024) ERTNet: an interpretable transformer-based framework for EEG emotion recognition. Front. Neurosci. 18:1320645. doi: 10.3389/fnins.2024.1320645