

LSTM vs Plot-based CNN for EEG Emotion Detection Tasks

Jared Kelnhofer, Marcus Blaisdell, Mona Ghandi

Washington State University
Wearable Computing REU Program
Pullman, WA, USA

jared.kelnhofer@wsu.edu, marcus.blaisdell@wsu.edu, mona.ghandi@wsu.edu

Abstract—Emotion detection using machine learning and data gathered from an electroencephalogram (EEG) holds the potential for architecture and creating smart adaptive spaces which can respond to the user’s current emotional state detected from the Neurophysiological data in real-time. This technology can help people with mental and physical disabilities to have a greater role in shaping their environment and live more independent lives.

In this paper, two different machine learning approaches, the Long Short Term memory network, (LSTM) and Convolutional Neural Network (CNN) are compared in order to assess their potential to satisfy this goal of emotion detection. The LSTM network was trained on eight-channel time-series data which had undergone a Fast Fourier Transform, and the CNN was trained on the un-transformed data in the form of a unique plot-image based approach.

Keywords—component; CNN; LSTM; EEG; Fast-Fourier Transform; Short-time Fourier Transform

I. INTRODUCTION (*USE CASE*)

Emotion detection with the use of EEG devices holds great potential for the improvement of the lives of people unable to clearly communicate their emotions, such as people who have certain forms of Autism or PTSD. Environments equipped with emotion detecting capabilities and adaptive systems could offer a better experience and adapt to the emotional state of the user. Here, the space plays an informative role to communicate the person’s emotions with the caregivers or parents and help them to prevent aggressive behaviors before they happen.

A. EEG Classification tasks (*Similar approaches*)

The use of EEG signals to detect emotions has been accomplished with relative success in the past through the use of Recurrent Neural Networks [2]. These networks have been shown to be able to achieve good performance when classifying emotions once the input data have been thoroughly pre-processed with techniques such as Short-time Fourier Transforms (STFT) and Zero-time Windowing.

An interesting approach towards time-series classification tasks involves the use of images rather than time-series data as the training medium. G. Zu et al. [1] were able to accomplish a 74% accuracy by implementing

an image-based approach to the problem of EEG-fed motor-imagery recognition. Instead of feeding the raw EEG data into the network the researchers used STFT and built time-frequency spectrum images to use as inputs. This practice of using images as features for EEG classification tasks is not common and only around 20% of recent studies have employed methods similar to this [3]. In 2018 spectrogram inputs were used to train a CNN to detect motor impairment neural disorders with promising results [4], however the researchers made se of FFT and other data filtering techniques when constructing their input spectrograms. In this paper we observe a lightweight plot-based approach to EEG classification of raw data in a spectrogram input format and compare it to a LSTM network in terms of classification accuracy.

II. METHODS (NEUROLOGICAL DATA COLLECTION: OPEN BCI EEG)

The data for both networks was collected by showing participants videos and prompting them to self-assess their emotional state at random intervals by pressing a computer key. The eight-channel OpenBCI EEG headset provides open-source software, allowing customization and measurement off the brain’s electrical activity. At their discretion participants in the lab were fitted with EEG headwear. They were also given a wristband. They then watched a video of their choosing until a particular emotion was evoked. Since people vary in their preferences and a same video might not evoke a particular emotion in different people the choice of the video was left up to the user. Data were collected over two years.

The EEG recorded eight channels of constant time-domain data sampling at 256 times per second, and paired time intervals with the self-reported 3motioanal labels. This dataset was then put through a Fast Fourier Transform for the LSTM which achieved an accuracy of about 93%.

A. Data Processing

We processed the dataset differently when we fed it into the CNN. Instead of feeding the raw or transformed numerical data into the network, plot images were constructed in order to make the task a “visual” image-recognition task. To construct an image, each of the eight channels was split by a sampling size of 128 measurements, or one half-second of real-world voltage reading data capture. By choosing a sampling size of one

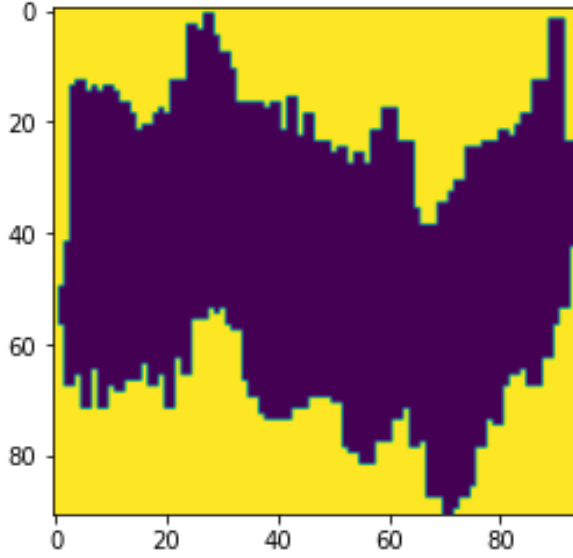


Fig. 1: 128-length channel plot

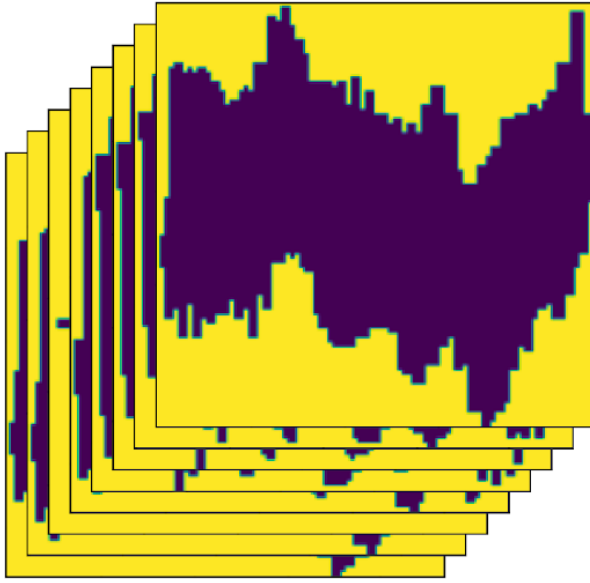


Fig. 2: All 8 channels making up a spectrogram

half-second, ignoring samples which had overlapping emotional labels, time-domain plot images (Fig. 1) were built for each channel and layered as the channels of one input spectrogram (Fig. 2). All of these spectrograms were then fed into the Convolutional Neural Network. By far the most commonly occurring class in the dataset, “NOEMOTION” posed a challenge when first beginning to train the network on the spectrograms (Fig. 3). The model started out with an accuracy nearly at 60%, but when the NOEMOTION class was removed, accuracy dropped to less than 12%. In order to train the model to diversify its predictions, the final output weights were balanced proportionally to the underrepresentation of each class. This instantly boosted the accuracy on the dataset

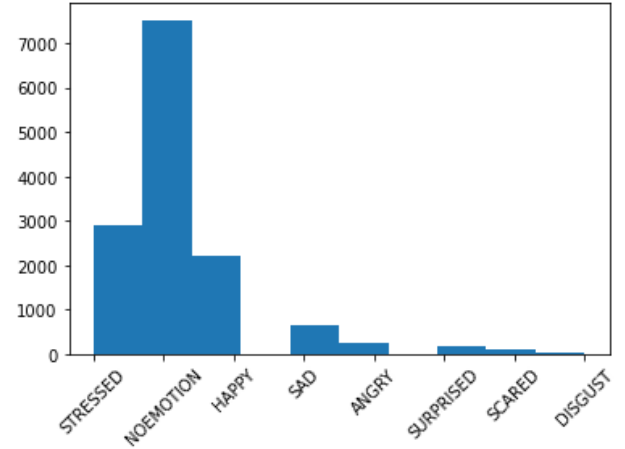


Fig. 3: Class distributions

without the NOEMOTION class to 41% but lowered the training and test accuracies of the network on the full dataset to around 31%.

In these early stages of development and after hyperparameter tuning via grid search, the plot-fed CNN method has achieved a 56% accuracy when classifying 8 emotions as opposed to the LSTM’s 93% accuracy (Table I).

III. RESULTS AND DISCUSSION

The most current accuracy of the trained CNN is ~56%. The LSTM trained on the same data currently outperforms the CNN. There are still improvements to be made to the data transformation pipeline and changing the way that these input spectrograms are generated in order to improve the model’s efficiency. With a quickly reachable 56% test accuracy, this method may indeed prove useful in the future, considering the lack of signal filtering.

IV. FUTURE WORK

In the future, this method will be refined in order to increase the accuracy of the networks’ predictions, potentially enabling its use in the field of emotionally responsive architecture. There are a number of ways to work towards improving the accuracy for this network in

TABLE I. ACCURACY SCORES

	Original Dataset	Balanced Weights	Grid Search	LSTM
Train Accuracy	~0.57	0.3115	0.5479	0.9256
Test Accuracy	~0.57	0.3115	0.5633	0.934
Cross Val Accuracy	N/A	N/A	0.5597	0.938
NOEMOTION Removed Acc.	0.1037	0.4079	0.3495	N/A

the near future One such improvement could be the changing of the sample size for building images. This number could be treated as a hyperparameter and assessed to find the optimal value for consistent emotion detection. Another tweak may be to use more information-dense input images for the CNN training. The images currently consist of numpy ndarrays composed of numpy uiny8 data types (Fig. 4) in order to save drive space. This results in pixel values which are either 0 or 1. These spectrogram pixels could be converted into another data type which can hold more information (Fig. 5) in order to preserve visual spectrogram properties that have been sacrificed for the sake of speed and efficiency during dataset generation.

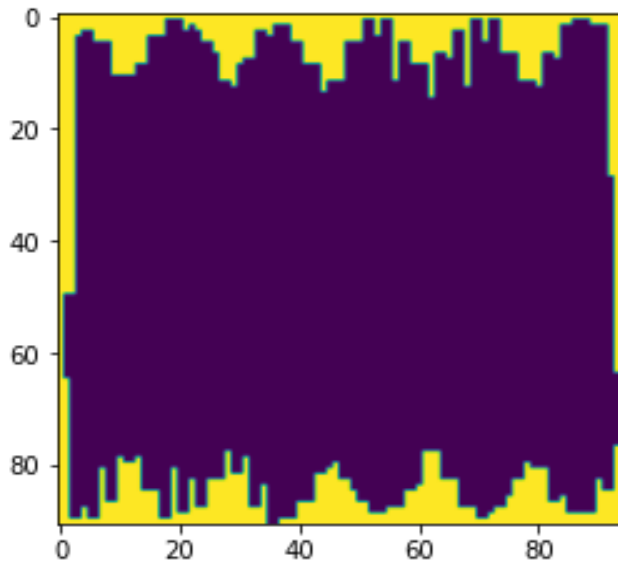


Fig. 4: Binary pixel values may lose information

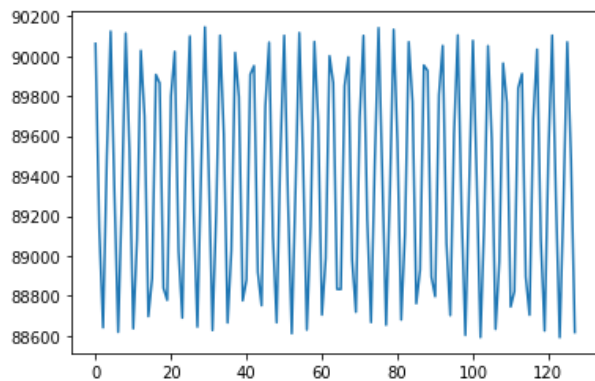


Fig. 5: More detail possible in input channels at the cost of greater storage requirements for dataset

REFERENCES

- [1] G. Xu et al., "A Deep Transfer Convolutional Neural Network Framework for EEG Signal Classification," in *IEEE Access*, vol. 7, pp. 112767-112776, 2019, doi: mn 10.1109/ACCESS.2019.2930958.
- [2] Gannouni, S., Aledaily, A., Belwafi, K. et al., "Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification," in *Sci Rep* 11, 7071 (2021). <https://doi.org/10.1038/s41598-021-86345-5>
- [3] Alexander Craik, Yongtian He, Jose L. Contreras-Vidal. 2019. Deep Learning for electroencephelogram (EEG) classificaiton tasks: a review. *IOPScience Journal of Neural Engineering* 16, 3 (April 2019). Retrieved from <https://iopscience.iop.org>
- [4] Vrbancic, G., & Podgorelec, V. (2018). Automatic Classification of Motor Impairment Neural Disorders from EEG Signals Using Deep Convolutional Neural Networks. *Elektronika IR Elektrotehnika*, 24(4), 3-7. H <https://doi.org/10.5755/j01.eie.24.4.2146>

ACKNOWLEDGMENT

This research was supported through the National Science Foundation Research Experiences for Undergraduates Wearable Computing Program at Morphogenesis Lab at Washington State University