Modernizing MDD Diagnosis using Deep Learning from EEG Data*

Milena Lebedinsky[1 [0009-0002-0037-6852]], Rocío Leguizamón[1 [0009-0002-7639-4693]], Pytel, P[1 [0000-0002-3920-1710]], Chatterjee, P[1 [0000-0001-6760-4704]], and Pollo-Cattaneo, Maria[1 [0000-0003-41973880]]

GEMIS - GEMISba - Universidad Tecnológica Nacional. Buenos Aires. Argentina.

Abstract. Major depressive disorder (MDD) is a widespread illness significantly impacting individuals' quality of life. Its diagnosis through Electroencephalogram (EEG) has long been studied in mental health research. Recent advancements in deep learning present a promising pathway for enhancing MDD diagnosis through EEGs. This study integrates state-of-the-art deep learning techniques, including ConvNext and Transformers architectures, into MDD prediction models. Results demonstrate ConvNext models' robustness and efficiency, in terms of precision and specificity, while Transformer models exhibit high recall and sensitivity for diagnosing MDD from incomplete studies.

Keywords: Major Depression Disorder · Electroencephalogram · Deep Learning.

1 Introduction

Depression is a common yet serious illness that interferes with daily life, affecting the ability to work, sleep, study, eat, and enjoy life in general. [1]. The potential of Electroencephalogram (EEG) studies in diagnosing Major Depressive Disorder (MDD) has been a subject of considerable interest in the psychiatric and neuroscientific community [2]. These studies offer a non-invasive window into the neural patterns that may be indicative of MDD. Recent advancements in deep learning have opened new avenues for extracting meaningful patterns from EEG data [3] [4]. The rapid progression of deep learning technologies for sequence analysis presents a significant opportunity to enhance the accuracy and reliability of EEG-based diagnostic models. Transformer models have recently emerged as powerful tools for sequence and image analysis. Their ability to capture long-range dependencies and intricate patterns in data sets them apart from conventional convolutional neural networks (CNNs, or ConvNets). ConvNext models were proposed, demonstrating a predictive capability comparable to that of transformer models, with the advantage of having a smaller number of parameters, demonstrating the capability of modernizing the existing ConvNet models [6]. This study aims to integrate these recent advancements, into existing

* Supported by GEMIS, Universidad Tecnológica Nacional, Buenos Aires, Argentina

models, particularly for the classification of EEG studies used as a biomarker for diagnosing MDD.

2 Methods and proposed architectures

In this study, three principal architectures were explored: Convolutional Neural Networks (CNNs), architectures based on ConvNext layers, and Transformers, each subjected to multiple topology variations. Initially, the methodology involved incrementally modifying individual layers within a base model to ascertain the configuration that exhibited the most promising metrics. Then, for the best performant models found, 7-fold cross-validation was used to collect training statistics. This performance evaluation was made using a publicly available dataset [7], containing 181 EEG studies, split into 86 studies for the control group and 95 MDD-diagnosed studies, which correspond to 64 different patients. The dataset includes eyes-closed (EC), eyes-open (EO), and task-based (TASK) EEG recordings. The mean performance metrics were computed using 26 test studies, corresponding to 5 patients, that were isolated to validate each fold with unseen data. This methodology provided results that reflect the models' stability and generalizability across different data partitions.

The data was preprocessed by selecting the common channels across all studies and normalizing the study durations to the shortest study length. The short-time Fourier transform (STFT) was computed, to use the absolute value of the resulting matrices as the network input. Building on the foundation laid by previous research [8], indicating a higher accuracy can be achieved through artificially enlarging the dataset, an additional preprocessing stage was tested on the Transformer models. Ten different time windows were cut from each study as a data augmentation technique. Hence, the Transformer model differentiated itself from the other architectures in its ability to predict the diagnosis from shorter or potentially incomplete studies and was trained with ten times more data. The impact of specific layer configurations on the diagnosis prediction accuracy was evaluated. Variations included the selective application of L2 kernel regularizers, introducing spatial dropout layers to mitigate overfitting, and incorporating different pooling strategies, such as global average pooling and global max pooling, to refine feature extraction.

3 Results

The individual training of different architectures revealed that the architecture of a Convolutional model should include a normalization layer before and after the first convolutional, followed by a second convolutional layer. The convolutional block training was the most stable when followed by one MaxPooling2D layer, and finally, a Flatten, one Dense, and one Droput layer, before the single-neuron predictive Dense layer.

The ConvNext model architectures used four ConvNext blocks, as the ones described in the foundational paper [6]. Each ConvNext block was preceded by a SpartialDropout2D layer and for every two of these blocks, a MaxPooling2D layer was inserted.

The same ConvNext (untrained) layers were used as an encoder for the Transformer architecture, followed by the Transformer block, with no positional encoder. After the transformer block, a Dropout layer and a GlobalAveragePooling1D layer followed, before the final Dense layer used for prediction.

Table 1. Trainable parameters of each model.

Transformer	Convolutional	ConvNext
89,137,877	59,850,537	898,025

The three described architectures' number of trainable parameters are shown in the table 1.

Table 2. Average training validation metrics for each model for each fold (the best epochs metric, in terms of training validation loss and training validation accuracy, were stored).

Metrics	Convolutional	ConvNext	Transformer
Best validation loss	0.5810	0.2716	0.3730
Accuracy at best validation loss	0.9193	0.7256	0.7411
Best validation loss epoch	21.2857	12.4286	11.4286
Best validation accuracy	0.9477	0.9088	0.8358
Loss at Best Validation Accuracy	0.6655	0.5146	0.4844
Best validation accuracy epoch	13.0000	10.2857	11.0000

Table 3. Average test metrics for each model across the 7 folds

Metrics	Convolutional	ConvNext	Transformer
Testing loss	0.6115	0.1760	0.4225
Testing accuracy	0.9490	0.9541	0.7804
Presicion	0.9643	1.0000	0.7725
Recall	0.88571	0.8393	0.9409
F1 Score	0.9058	0.8956	0.8332
Sensitivity	0.8571	0.8393	0.9409
Specificity	0.9857	1.0000	0.6032
Gmean	0.9187	0.9070	0.6272

Training validation statistics were captured to exhibit each model's behavior, and are displayed in table 2. Then, metrics calculated with the test data were calculated, and are shown in table 3. The average confusion matrix for each model is displayed in Figure 1, also calculated from the testing set.

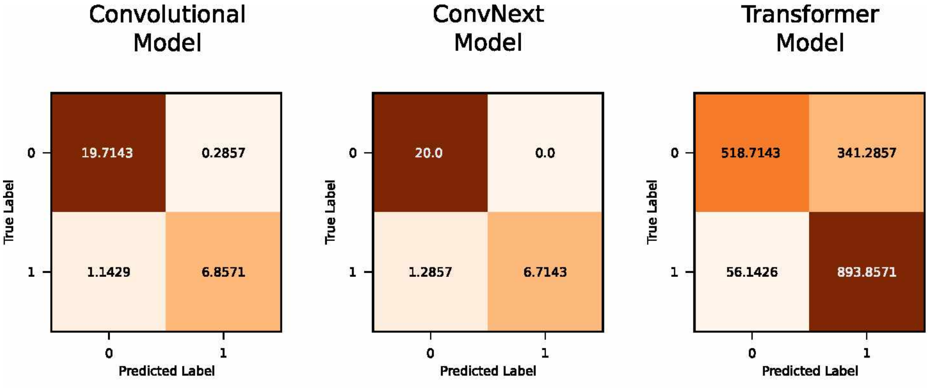

Fig. 1. Average confusion matrix for each model across the seven folds

4 Discussions

The convolutional model reached the highest average F1-score and Gmean, making it the most balanced model in terms of precision and recall, as well as specificity against recall. However, the final architecture resulted in a complex model with a high number of parameters. Presenting the highest loss could be particularly disadvantageous for the specific case of predicting a mood disorder, as it leads to less reliability.

The ConvNext architecture was the one that, while sacrificing some sensitivity and recall, achieved perfect specificity and precision, making no false negative predictions within the testing examples. The loss and accuracy converged at early epochs, showing a great learning capacity in this model which, moreover, is significantly lighter in parameters than the others.

Although the performance of the transformer model yielded lower scores in most of the metrics, it showed the highest recall and sensitivity, parameters that may be of interest when considering a tool for diagnosis assistance. On the other hand, it should be highlighted that the results are not as directly comparable with those of the other models, since both, training and prediction, were made from short snippets taken from the studies in this case. Being this a work in progress, it is worthwhile to continue studying this predictive capacity from incomplete studies.

The relevance of each metric will depend on the model's clinical use and applicability to specific diagnostic scenarios. However, these preliminary results can highlight the ability to obtain better outcomes by tuning the architecture of traditional ConvNets. The results obtained are, to some extent, proof of the success of the ConvNext architecture's underlying hypothesis [6]. In this case, the ConvNext model reached better results with the test data than the best epoch validation results when comparing the averages, meaning it's more robust given the amount of data used for training and testing.

Our future work proposals aim to achieve more robust results. Training the same models with more numerous EEG depression samples will ensure that the performance variations found between the different architectures are consistent. This will also answer to what degree these techniques are subject to overfitting. To make a more robust comparison between the presented architectures, the same preprocessing and data augmentation techniques should be used in the Convolutional and ConvNext models. Also, it's worthwhile to compare the Transformer prediction ability from complete studies. Finally, a hyperparameter exploration must be conducted for each network.

For the transformer models in particular, a pre-trained encoder-decoder that predicts how a study should follow could be used to replace the encoder part. Including the STFT transformation of the data as a layer, so that the parameters of this calculation can be explored as hyperparameters, and the addition of a positional encoder are included in our future work proposals.

References

1. Pan American Health Organization, https://www.paho.org/en/topics/depression. Last accessed 3 March 2024
2. de Aguiar Neto FS, Rosa JLG.: Depression biomarkers using non-invasive EEG: A review. Neurosci Biobehav Rev. 2019 Oct;105:83-93. https://doi.org/10.1016/j.neubiorev.2019.07.021 Epub 2019 Aug 7. PMID: 31400570.
3. Loh, Huiwen, Ooi, Chui, Aydemir, Emrah, Tuncer, Turker, Dogan, Sengul, Acharya, U Rajendra. (2021). Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals. Expert Systems. 39. e12773. 10.1111/exsy.12773.
4. Zhang, J., Xu, B., Yin, H. Depression screening using hybrid neural network. Multimed Tools Appl 82, 26955–26970 (2023). https://doi.org/10.1007/s11042-023-14860-w
5. A. Rafiei, R. Zahedifar, C. Sitaula and F. Marzbanrad, "Automated Detection of Major Depressive Disorder With EEG Signals: A Time Series Classification Using Deep Learning," in IEEE Access, vol. 10, pp. 73804-73817, 2022, doi: 10.1109/ACCESS.2022.3190502,
6. Zhuang Liu and Hanzi Mao and Chao-Yuan Wu and Christoph Feichtenhofer and Trevor Darrell and Saining Xie: A ConvNet for the 2020s. 2022, eprint: 2201.03545, archivePrefix=arXiv, primaryClass=cs.CV
7. Mumtaz, Wajid (2016). MDD Patients and Healthy Controls EEG Data (New). figshare. Dataset. https://doi.org/10.6084/m9.figshare.4244171.v2
8. George O, Smith R, Madiraju P, Yahyasoltani N, Ahamed SI. Data augmentation strategies for EEG-based motor imagery decoding. Heliyon. 2022 Aug 17;8(8):e10240. doi: 10.1016/j.heliyon.2022.e10240. PMID: 36060998; PMCID: PMC9433610.