NeuroVision: Multimodal Deep Learning Prediction of Higher-Order Cognition

Woody Hulse, Micah Lessnick, Vandana Ramesh
Brown University, Providence, RI 02912

May 11, 2023

1 Introduction

In recent years, there has been rapid research into applications of deep neural networks to brain biometric data. Most work has been done on the classification of neurological disorder by evaluating abnormalities in either MRI or EEG data, but there has been little focus, in the conventional sense or with novel deep learning methods, on the prediction of higher-order behavioral cognition. Our goal is to predict over an array of behavioral metrics (CVLT, LPS, RWT, TAP, TMT, WST, DSM-V) using an array of 3D Deep Convolutional Neural Network (DCNN) approaches paired with an EEGNet-based 2D convolutional model to interpret both MRI and EEG data. We hypothesize that a deep learning model will be able to detect and interpret structures and activities in the brain that indicate certain behavioral characteristics.

2 Methodology

Our approach to this kind of multimodal model was to bifurcate our final model into its constituent EEG and MRI model parts; then, we develop those parts optimally before combining latent outputs for those models into our combined model. This way, we can experiment with lightweight EEG models more easily, however we assume here that optimally performing constituent models will create an optimally performing joint model.

2.1 Data and Preprocessing

We collect data from the Max Planck Institute in Lepzig, Germany from 228 healthy participants. Our data contains paired MRI, EEG, and behavioral data across an array of 27 behavioral metrics.

We first standardize the length of each EEG readout by truncating each end of potentially faulty data, and filter any patients that don't have the necessary 60 electrical readouts that our models require. We also pass a low-pass filter (50 Hz) to remove excess noise from the EEG data.

With MRI data, we again filter participant data which doesn't have the necessary resolution requirements. Because we use transfer learning for MRI data, we take two additional approaches to preprocessing 3D MRI data into 2D data interpretable by a 2D transfer learning model. Our intuitive approach was to split the data along each axis (XYZ) and pass each set of slices through VGG, or a VGG-like feature extraction model. This was repetitive (tripling the raw amount of data) and computationally inefficient (requires VGG passed over each slice of each axial direction). We settled on Axial-Coronal-Sagittal (ACS) convolution, which instead leverages transfer learning color channels to process the local pattern for each convolution of all three axial directions (Yang et al., 2021).

With behavioral data, we extract all relevant tests, then take representative statistics from those tests, then standardize those scores across the population of samples. Post preprocessing, we were left with 118 valid samples of complete, paired patient data.

2.2 EEG Model

There have already been several studies that have developed great EEG-specific DCNN models, the most popular of which being EEGNet (Lawhern et al., 2018) (Guleva et al., 2020)

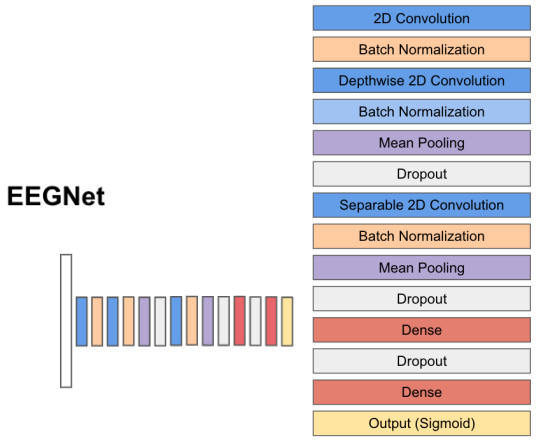

Figure 1: EEG Model Architecture

Unlike previous applications of EEGNet, primarily for diagnoses, we are not classifying, so we modify the head of the architecture with a light dense model head. We standardize earlier so we can apply sigmoid to this final layer and ensure that all model predictions will always lie within the reasonable domain of our model. The drawback to this, of course, is that our model then relies on knowledge of the distribution of the general population–in our case, we disregard this and substitute it with our data distribution.

We also try custom CNN architectures and using LSTM units. To the former, EEGNet is already the state-of-the-art, so we decided to use that for the best chance of success. Intuitively, it makes sense to use recurrency for sequential data like EEG. Although this has been done, it is difficult to structure a model that is half recurrent and half non-recurrent, since we must maintain some sort of joint latent space that can operate over a time series. Further, it makes more sense to use convolution on periodic data, since ideally the size of the kernel is able to segment each period and thus learn interpretations of individual cycles.

2.3 MRI Model

As briefly mentioned, the MRI model was our intended point of innovation. Conventional approaches include only convoluting over dominant layers of MRI scans, but these have consistently failed to predict behavior (Blue, 2020) (Elliott et al., 2020). Therefore, we attempt several novel approaches.

We developed a 3D DCNN for processing all MRI data concurrently by modifying VGG19 for 3D input, but this was too memory and processor hungry to train with conventional computing resources. Our subsequent approaches involve slicing the MRI input data. We first pass all slices through VGG-like models, but this also too was too resource intensive to train in a reasonable amount of time. Further, this introduces redundancies in the data (requiring that each pixel be interpreted three times) and disallows model 3D spatial awareness like we set out for. We settled on the ACS approach to concurrently pass in axial, coronal, and sagittal local layers. We also tried several architecture structures within our transfer learning attempts, namely VGG and Inceptionv3 (Google). Because the structures that theoretically dictate behavior may be of variable size, and due to outright performance improvements for both our task and feature recognition in general, we settled on using Inception modules to construct our MRI model.

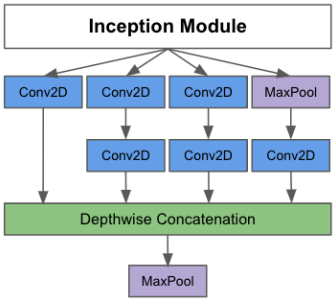

Figure 2: Inception Module

Inception modules concurrently pass information through 4 convolutional steps, each of which contains a different feature size. After concatenating, Inception-like models can implicitly learn the “most informative” size.

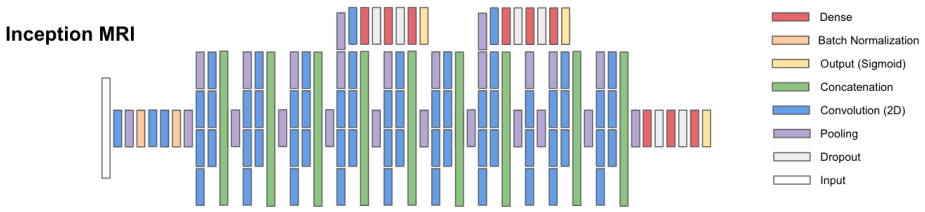

Figure 3: MRI Model Architecture

2.4 NeuroVision

The NeuroVision architecture is essentially the union of the MRI and EEG model architectures. We fuse the latent space prior to the head of each model with a consistent join size (20 neurons). We also include auxiliary outputs, which allow our model to ensure that meaningful information is being extracted within earlier layers of the network, thus preventing vanishing gradients.

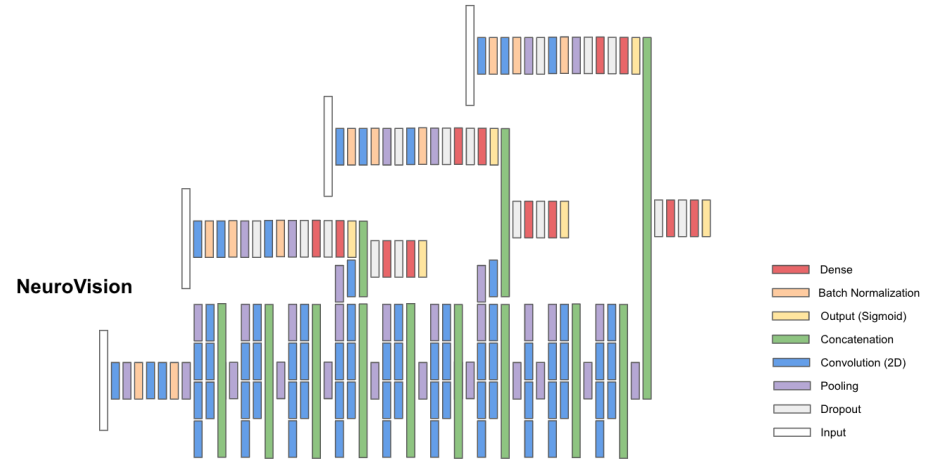

Figure 3: NeuroVision Architecture

3 Results

Because there is no comparative network that evaluates over continuous behavioral metrics, we elect to use control models to benchmark model performance. These include a "center model" (guesses center of the rescaled data), "mean model" (guesses the data mean), "median model" (guesses the data median), "guess model" (guesses a random datapoint from the training dataset), and a simple neural network. Our NeuroVision model, including the constituent models, failed to outperform control models by a statistically significant amount.

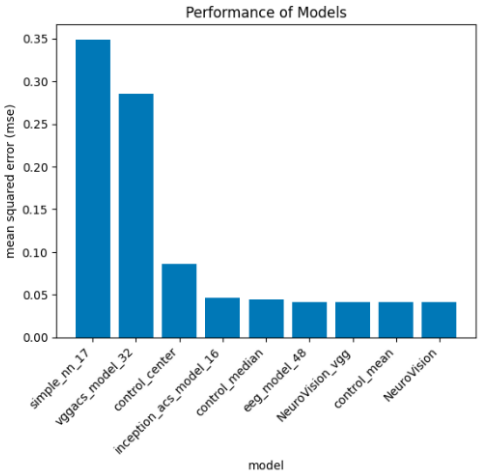

Figure 3: General performance of NeuroVision against controls

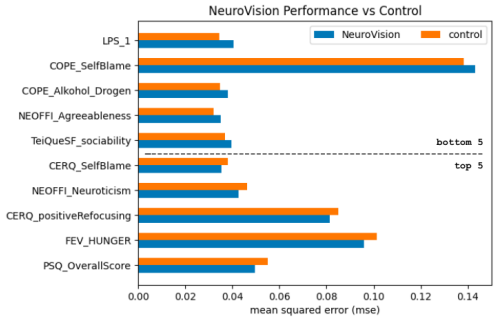

Figure 4: Best five and worst five categories of performance of NeuroVision against controls

Top 5 Performance Categories of NeuroVision

	diff in MSE	MSE	p-value
PSQ_OverallScore	-5.479236e-03	0.049449	0.612672
FEV_HUNGER	-5.321663e-03	0.096000	0.687184
CERQ_positiveRefocusing	-3.721070e-03	0.081348	0.744951
NEOFFI_Neuroticism	-3.644320e-03	0.042516	0.740512
CERQ_SelfBlame	-2.714836e-03	0.035480	0.670292

Figure 5: Sample of statistical significance of NeuroVision performance against control

There could be several reasons for not achieving improvement over control.

Because we haven't yet identified a precise mapping between behavior and structures and activities in the brains scientifically, it could be the case that it is impossible to create such a mapping altogether with available brain biometric data.

Further, we only have 30 validation samples, and so discrepancies found between our model and control performance are difficult to decisively say are significant. It could also be that more data for our model, like has been shown in feature recognition tasks in other settings, may demand significantly more data to be able to properly identify complex relationships in the data. Another issue may be the sample, which is entirely from Lepzig, Germany, which may be responsible for some reduction in variance in behavioral outcomes in the data.

Another limitation is our data representation. We chose to use MRI data, which in our case has resolution of $256 \times 256 \times 256$. It is entirely possible that the structures that we would need to locate to determine a mapping to personality are much finer than this available resolution. We also make no attempt to bake in inductive bias into our dataset regarding known areas of the brain responsible for higher-order thinking, namely the prefrontal cortex, by excluding peripheral areas. Lastly, our use of EEG data may be misleading for our model, as other measures of activity such as MEG are known to be more representative of deeper brain activity.

4 Challenges

How do we create a model on 3D MRI imaging data?

This is a problem that we decide to attack with two approaches. Our initial approach was to create a 3D implementation of existing cutting-edge 2D DCNN architecutres, but as we mentioned this hasn't worked for training reasons. We also try a replica of cutting-edge 2D convolutional architectures for feature matching (namely, VGG's models). A second approach though is a transfer learning one, where we ACS (axial-coronal-sagittal) preprocess data with some downsampling to compute optimal feature maps from already successful architectures along three dimensions.

How do we evaluate performance of multimodal models across a broad range of metrics?

We ideally evaluate the capacity of each component of our model, alongside our conjoined model, to predict across a full range of behavioral tests. We accomplish this by partitioning our model and input data, standardizing, and setting an objective function as a minimization of Euclidean distance between each (MRI, EEG, combined) model's prediction and the percentile vector for each patient. It is important to note, then, that what constitutes as a "close proximity" is not well defined–therefore, we create four "control" models (which guess based on the mean, median, center, and simple neural network predictions) in order to determine the true efficacy of our model.

How do we handle the task of training such large models?

Our MRI model prototypes are very large (10-50 million trainable parameters). While these models are somewhat trainable on local machine resources, we leverage cloud resources like Center for Computation and Visualization's Oscar and Google Cloud Platform to greatly improve the speed at which we're able to train models.

How do you prevent exploding or vanishing gradients in huge (90M) parameter models?

We take advantage of a structure pioneered by models like Google's Inception by optimizing over additional auxiliary outputs as well as the final loss gradient calculation, such that we ensure that in intermediate layers of the network there is no information loss.

5 Reflection

Q: How do you feel your project ultimately turned out? How did you do relative to your base/target/stretch goals?

A: We of course did not meet our goals (any acceptable level of statistical significance) but that isn't to say the project was a failure. We used the most comprehensive and advanced methods possible to solve an unsolved problem, and we believe that attempted enough architectures and configurations to show that, at least with data and resources currently available, this problem isn't yet solvable. It certainly would've been cool to create a model that can "solve personality," but we didn't necessarily expect to be able to do this.

Q: Did your model work out the way you expected it to?

A: Our original idea going into the project is that we'd have a joint LSTM and 3D convolutional model. Of course, this isn't how it turned out, and the process of figuring out what works best was probably the most rewarding part of the process.

Q: How did your approach change over time? What kind of pivots did you make, if any? Would you have done differently if you could do your project over again? What do you think you can further improve on if you had more time?

A: Through literature review and reasoning we first changed the structure of the EEG part of the model to an EEGNet modification. We had several pivots regarding the MRI architecture, which were mostly experimental before settling on Inception modules. There are several things that we would like to go back and try instead though, including finding higher-resolution data and blocking off areas that we know aren't responsible for behavior, as mentioned earlier. We only had access to cloud computing resources for about a week and a half, so gaining access earlier would be more of a priority if we could do this over–with more time, we think we could've come up with a more sophisticated and potentially better model.

Q: What are your biggest takeaways from this project/what did you learn?

A: As frustrating as it can be to spend weeks developing a model and not be able to achieve statistically significant results, it was rewarding in some respect. Character building, probably. We came in really hoping to develop a model that could do something that nobody else has been able to do before, with which an explainable version could legitimately impact neuroscientific theory. But we couldn't achieve this, and that's probably okay. Again, a combination of better planning and swifter execution may have given us a better chance at this, but as it stands, we are satisfied with the learning experience alone.

6 References

Burgos, Ninon; Colliot, Olivier. Machine learning for classification and prediction of brain diseases: recent advances and upcoming challenges. Current Opinion in Neurology 33(4):p 439-450, August 2020. — DOI: 10.1097/WCO.0000000000000838

Wutao Yin, Longhai Li, Fang-Xiang Wu. Deep learning for brain disorder diagnosis based on fMRI images, Neurocomputing, Volume 469, 2022, Pages 332-345. ISSN 0925-2312 — DOI: https://doi.org/10.1016/j.neucom.2020.05.113.

Zhang, Li, et al. "A Survey on Deep Learning for Neuroimaging-Based Brain Disorder Analysis." Frontiers, Frontiers, 2 July 2020, https://www.frontiersin.org/articles/10.3389/fnins.2020.00779/full.

V. Guleva, A. Calcagno, P. Reali and A. M. Bianchi, "Personality traits classification from EEG signals using EEGNet," 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON), Palermo, Italy, 2022, pp. 590-594, doi: 10.1109/MELECON53508.2022.9843118.

V. Shah and S. Modi, "Comparative Analysis of Psychometric Prediction System," 2021 Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, 2021, pp. 1-5, doi: 10.1109/STCR51658.2021.9588950.

Vernon J Lawhern et al 2018 EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. J. Neural Eng. 15 056013, DOI 10.1088/1741-2552/aace8c

Yang, Jiancheng, et al. "Reinventing 2D Convolutions for 3D Images." IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 8, Aug. 2021, pp. 3009–18. arXiv.org, https://doi.org/10.1109/JBHI.2021.3049452.

Blue, Charles, and APS Staff Writer. "Predicting Behavior by Scanning the Brain: Does Task-FMRI Really Resonate?" APS Observer, vol. 33, June 2020. www.psychologicalscience.org, https://www.psychologicalscience.org/observer/predicting-behavior-by-scanning-the-brain-does-task-fmri-really-resonate.

Elliott, M.L. and Knodt, A.R. et al. (2020) What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. Psychological Science, https://doi/10.1177/0956797620916786

Emily B. Falk, et al. (2010) Predicting Persuasion-Induced Behavior Change from the Brain. Journal of Neuroscience 23 June 2010, 30 (25) 8421-8424; DOI: https://doi.org/10.1523/JNEUROSCI.0063-10.2010