Aaron Sossin 06431543
Vivian Zhu 06424921
CS229 Milestone

CNNs Applied to sMRIs of Heavy Cannabis Users

Motivation:

Cannabis is the most consumed illicit drug in the world (1). Its effect on brain mechanisms thus has serious implications - giving rise to a myriad of research in this area. In 2008, Yucel et. al found that heavy cannabis users had bilaterally reduced hippocampal and amygdala grey matter volumes, suggesting cannabis can lead to cell death in brain regions with a high count of endocannabinoid receptors (2,3). These results were further supported by Batistella et. al in 2014 who implicated even more brain regions such as the temporal and orbitofrontal cortices (4). These studies use Magnetic Resonance Imaging (MRI) to create a 3D image of the brain, specifically outlining regions of grey/white matter (2,4). Although cannabis also affects neurons on the cellular level (5), MRIs reveal the global effect on neuron count and shape of the brain. There is thus ample evidence to suggest cannabis use can be detected by MRIs (2,3,4).

Although ML has not been used in the context of cannabis use, it has successfully examined MRIs in similar fields. For example, in 2018, Xi Zhu et al. used MRIs in conjunction with a Random Forest (RF) classifier to predict alcohol dependence (6). Moradi et. al devised an algorithm that predicts the onset of Alzheimer's disease better than clinicians (7) and Salvatore et. al similarly succeeded in classifying brains with Parkinson's (8). Since MRIs can detect cannabis use (2,4) and ML has garnered success in similar tasks (6,7,8), we suggest that applying classic ML techniques to the analysis of 3D MRI scans will provide valuable insight into the brain mechanisms of heavy cannabis use.

Koenders et al. published a dataset in 2016 containing T1-weighted structural Magnetic Resonance Imaging (sMRI) data of patients at a baseline and at a 3-year follow up (9). This sMRI data is augmented by patient scores on the Cannabis Use Disorder Identification Test (CUDIT) and Alcohol Use Disorder Identification Test (AUDIT). In this study, Koenders et. al were primarily focused on correlating CUDIT scores to grey matter volumes in select brain regions (9). They found grey matter volume of the left hippocampus, amygdala and superior temporal gyrus to be negatively correlated with cannabis use (9). We aim to supplement their research with ML algorithms that have the power to extract more complex MRI features, and perhaps increase predictive ability. Moreover, we aim to provide valuable research to the neuroscience community on how to successfully implement ML on brain MRI images.

There are three main experiments we wish to perform: 1) a regression analysis predicting CUDIT score from MRI image, 2) a classification analysis distinguishing controls from patients

with heavy cannabis use, and 3) a segmentation analysis which isolates the most predictive features/voxels from the above analyses.

Methods:

The data was downloaded from the public repository openneuro.org at this link https://openneuro.org/datasets/ds000174/versions/1.0.1. Each subject has an sMRI image at a baseline and at a 3 year follow-up with corresponding CUDIT/AUDIT scores. In all, there are 42 subjects of varying age and sex, 23 of which are not considered heavy cannabis users.

We began to implement neural networks using MONAI which is a PyTorch-based, open-source framework for deep learning, that specifically has functions that support downloading and converting MRI files of the format ".nii.gz". All implementations run on the Google Cloud Platform and in Python. To start, the MRI images underwent basic preprocessing. Using the 'Transforms' implemented by MONAI, the MRI images were resized to a 3D dimension of 96x96x96 and the voxel intensities were normalized.

Central to our intended methodology is to create a grid search that tests several popular loss functions (Entropy Loss, Mean Squared Error, L1 Loss), optimizers (Adam, SGD), and epoch sizes on different CNN models. We will start with the easiest models to implement: Dense CNN architectures specified by the MONAI API. We will then incorporate more sophisticated CNN architectures like the SegResNetVAE which uses an autoencoder to incorporate unsupervised learning, a method we hope will buoy our relatively small dataset. We will also explore pre-trained CNN architectures that have performed well on the ImageNet Classification Challenge, since this has proven to garner success in analyzing MRIs (10). There has also been evidence to suggest using random classifiers that combine several of the aforementioned networks can improve score (7).

Perhaps more important than model selection is the pre-processing pipeline of the MRI images. We plan on using the *nilearn* python module to perform classically successful techniques like skull-stripping, bias-field correction, registration as well as statistical parametric mapping (11). Moreover, the *FreeSurfer* neuroimaging software has shown similarly positive results (12). These modules also output useful statistics that can serve as their own features during training.

In order to perform the segmentation task, we aim to follow MONAI's segmentation tutorial. However, we are not constrained to classical MONAI implementations as we can devise our own PyTorch supported models. Should the segmentation task isolate brain regions/voxels with high endocannabinoid counts and/or regions highlighted by Koenders et. al, this would provide evidence our network is producing results consistent with the literature.

In order to evaluate our models, we plan on using cross validation with k = 5, to measure several metrics like AUC score, ROC curve, accuracy, balanced accuracy, precision, recall, etc…

Preliminary Experiments:

Figure 1 shows some preliminary results on Dense CNN architectures specified by the MONAI interface. In order to create a baseline performance for the classification task, we started by using several DenseNet models, along with the Adam optimizer and Cross Entropy loss function. The best result is the DenseNet169 architecture at 90 epochs with accuracy: 0.61. Experimentation has been limited so far due to time constraints, but we have tested the assumption that our intended methodology using MONAI works, and also that there is enough drug use variability to find predictive features in the MRIs.

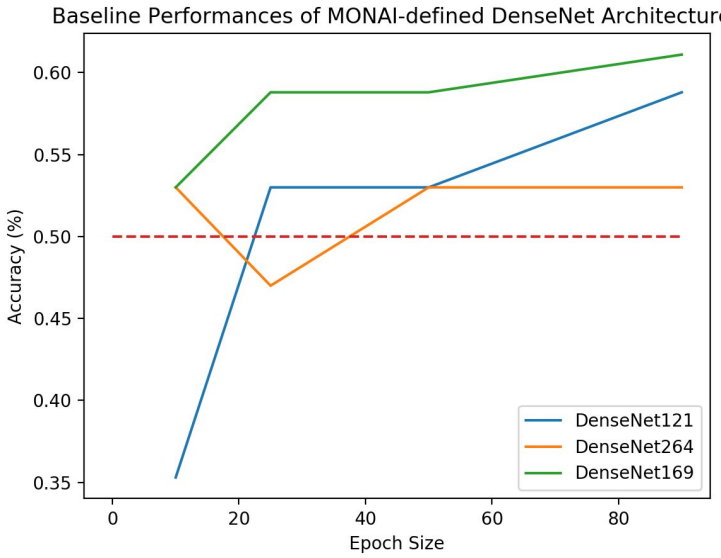

Figure 1: DenseNet models with Adam Optimizer, Cross Entropy Loss Function.

Next Steps:

Based on Figure 1, it appears as though more epoch sizes can improve the model as accuracy hasn't plateaued. So, we will try with 200, and 500 epochs. Then, we will increase the pre-processing of the MRI images using *nilearn* and *FreeSurfer*. Then, we will perform a proper grid search on the dense nets. This will help us optimize parameters for our future explorations. Using these optimal parameters, we will begin to experiment with new model types as defined in the Methods section. Regardless of the efficacy of the regression and classification tasks, we hope that segmenting the MRIs into regions/voxels of high predictive value will shed light into the specific anatomical regions impacted by cannabis use.

Contributions: Aaron and Vivian Equally: Project Idea, Software, Writeup

References:

(1) 1. Degenhardt L, Hall W (2012): Extent of illicit drug use and dependence, and their contribution to the global burden of disease. The Lancet 379:55–70

(2) Yücel, Murat, et al. "Regional brain abnormalities associated with long-term heavy cannabis use." *Archives of general psychiatry* 65.6 (2008): 694-701.

(3) Iversen, Leslie. "Cannabis and the brain." *Brain* 126.6 (2003): 1252-1270.

(4) Battistella, Giovanni, et al. "Long-term effects of cannabis on brain structure." *Neuropsychopharmacology* 39.9 (2014): 2041-2048.

(5) Hall, Wayne, and Donald MacPhee. "Cannabis use and cancer." *Addiction* 97.3 (2002): 243-247.

(6) Zhu, Xi, et al. "Random forest based classification of alcohol dependence patients and healthy controls using resting state MRI." *Neuroscience letters* 676 (2018): 27-33.

(7) Moradi, Elaheh, et al. "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects." *Neuroimage* 104 (2015): 398-412.

(8) Salvatore, Christian, et al. "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy." *Journal of neuroscience methods* 222 (2014): 230-237.

(9) Koenders, Laura, et al. "Grey matter changes associated with heavy cannabis use: a longitudinal sMRI study." *PLoS One* 11.5 (2016): e0152482.

(10) Antropova, Natalia O., Hiroyuki Abe, and Maryellen L. Giger. "Use of clinical MRI maximum intensity projections for improved breast lesion classification with deep convolutional neural networks." *Journal of Medical Imaging* 5.1 (2018): 014503.

(11) Bernstein, Alexander, et al. "MRI brain imagery processing software in data analysis." *Trans. Mass-Data Analysis of Images and Signals* 9.1 (2018): 3-17.

(12) Lin, Weiming, et al. "Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment." *Frontiers in neuroscience* 12 (2018): 777.