Title
Weakly Supervised Deep Feature Learning to Predict Aβ-positivity from Structural MR Brain Images

Author
Manuel, Cyrus

Publication Date
2017

Peer reviewed|Thesis/dissertation

Weakly Supervised Deep Feature Learning to Predict Aβ-positivity from Structural MR Brain Images

by

Cyrus Manuel

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements

To my committee, thank you for the welcoming support and making it a fun and memorable year. Duygu thanks for the inspiration to explore graphs and allowing to me to dive into deep learning. To Michaël, thank you for helping me digest your code. To my loving wife, my sunflower in the cornfields, thanks for embarking on this journey with me and making me not take myself too seriously.

Abstract

Introduction – Many novel treatments for Alzheimer's Disease (AD) are aimed to target Aβ, one of the pathological hallmarks of AD, but are hampered by potential non responders due to lack of target Aβ pathology in their brains. Specifically, about ~25-40% of those clinically diagnosed with AD or mild cognitive impairment (MCI) would not have significant Aβ pathology. In this study we used a deep learning framework to predict Aβ pathology positivity from baseline clinical assessments and structural MRI data routinely acquired from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Graph convolutional networks (GCNs) were trained on graphs derived from MR imaging and their performances was assessed to see their predictive value based on ground-truth Aβ-positivity estimates from AV45-PET scans. Hidden layers of the first convolutional layer were visualized to observe pertinent networks the model has learned to be important towards predicting Aβ-positivity.

Methods – Baseline MRI and AV45-PET along with demographic and other AD-related predictors (age, gender, education, genotype, baseline cognitive performance) from 771 participants (248 healthy controls (HC), 381 MCI, 101 AD) from the ADNI database were used for evaluation. An undirected graph model reconstructed from diffusion MRI served as inputs for the GCNs. Anatomical brain parcellations with atrophy estimates from structural MRI constitute the vertices; tractography based connectivity estimates defined the edges of the graph model. Separate GCNs models were trained for each clinical diagnostic group. The best performing model architecture were used to assess model performance. Predictive value was compared with models trained on atrophy data and models with atrophy and AD-related predictors. A 10-fold cross validation on independent training and test sets were performed to assess the model performance in terms of classification accuracy, sensitivity, specificity, positive and negative predictive values.

Results – GCNs were able to learn from atrophy descriptors and network connectivity derived from MRI and predict Aβ-positivity. Atrophy was a significant predictor of Aβ-positivity in the AD model, but at a lesser degree in HC and MCI models. The inclusion of other AD-related predictors showed: a significant improvement in test accuracy to 68±4%, sensitivity to 84±7%, specificity to 52±13%, negative predictive value to 77±5%, and positive predictive value to 64±4% in MCI models; and a significant improvement in test accuracy to 69±2% and specificity to 97±4% in HC models. In one MCI network model, filters of the first hidden layer suggest greater contribution of atrophy in left superior parietal, right inferior temporal, right entorhinal, and left postcentral regions.

Conclusion – Patterns of regional brain atrophy within large-scale brain networks might offer predictive value to whether or not a subject will test positive for an AV45-PET exam. Predictions are more accurate with the addition of well-established AD-related predictors, however more features may be necessary to increase the predictive ability in HC and MCI subjects. This assessment offers a practical adjunct to deciding the next course of action for the patient.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease and the most common type of dementia. Synaptic dysfunction can be attributed to the accumulation of β-amyloid and tau proteins that spread in a disease-specific topographic pattern[1,2]. Aβ deposits first appear in the neocortex and spread to the outer cortical regions primarily affecting the frontoparietal regions and then temporal regions of the brain[2]. In contrast, neurofibrillary tangles or tau aggregates develop in locus coeruleus and spread to broad areas of the neocortex[2]. Evidence largely from the autopsy studies suggest that spread of AD-related tau pathology is facilitated in presence of significant brain Aβ pathology. The increasing cost both in the healthcare and familial setting is alarming[3]. Therefore it is crucial to develop predictive models to help better understand the dynamics between formation and spread of AD pathologies to lower costs and provide effective treatment.

According to two independent phase 3 drug trials, about 27% of those who meet the clinical criteria for mild-AD are Aβ-negative[4,5]. The presence of Aβ-negative subjects in trial cohorts represents a potential confound that may introduce variability and dilute a treatment signal in analyses where a slowing of clinical progression is hypothesized. Therefore, measure of brain Aβ as an inclusion criterion has been crucial in recent clinical trials of putative therapeutics for AD. Both positron emission tomography (PET) imaging, using Aβ-specific radiotracers such as [^{11}C]-PIB or [^{18}F]-florbetapir, and measurement of Aβ proteins from cerebral spinal fluid (CSF) samples, are widely used in the research setting to quantify brain Aβ plaque load. However, PET and CSF methods can be challenging to implement in global clinical trial sites outside the western hemisphere, for reasons including patient acceptance, cost, and availability.

Could Aβ pathology status of clinical trial participants be predicted from measurements widely available and already included in clinical trial protocols? Our aim in this work to answer this question by developing a predictive model based on participant's demographics, baseline cognitive and clinical assessments, and state of brain atrophy quantified from structural magnetic resonance imaging (MRI), which is already included in global clinical trial protocols for radiological monitoring.

Recent analytical developments have demonstrated that Aβ status can be predicted to a high accuracy in both early and amnestic mild cognitive impairment (MCI) subjects from Alzheimer's Disease Neuroimaging Initiative (ADNI) using a macroscopic pattern of brain structural deformation obtained from structural-MRI data[6,7]. Furthermore, recent developments in convolutional neural networks (CNNs) have started revolutionizing computational medicine[8,9]. In particular, CNN computed over graph structures could potentially have high impact in neurodegenerative disease models by taking advantage of brain's intrinsic connectivity while developing predictive models[10]. In this work, we will further extend these two approaches, in particular, by representing the brain as a graph and by learning the best Aβ status classifier via a weakly supervised neural network. Nodes of the graph represents different anatomical regions of the brain and their structural connectivity is derived from diffusion tensor imaging (DTI). Each node can contain features describing shape variations like atrophy from structural MRI which have been shown to be a good biomarker for predicting the Aβ-positivity[6,11]. Additional AD-related predictors like age, genetic background, cognitive assessments are also included as these are the strongest known risk factors for AD[7].

An implementation of this model is proposed to predict Aβ-positivity from MR imaging along with demographic, genetic, and cognitive assessment information. The associated AV45-PET labels serve as the ground truth and is used for learning and optimizing the model. This model will be used to test the hypothesis that a disease stage specific anatomical variation pattern detected in structural brain MRI is predictive of Alzheimer's related brain Aβ pathology. The purpose of this study is twofold: 1) to accurately predict Aβ-positive without AV45-PET imaging and 2) identify the important features and/or patterns of regional brain atrophy that are predictive of Aβ-positivity.

Chapter 2

Methods

2.1 Study Data

Data used in this study were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org. Subjects of this study were ADNI participants who underwent AV45-PET imaging and had structural MRI acquired within 60 days of PET scan. According to the clinical assessment done closest in time to neuroimaging visit, the study cohort was composed of 248 healthy control (HC) elderly individuals, 381 individuals with MCI, and 101 individuals with AD clinical diagnosis. The diagnostic criteria for HC and MCI participants in ADNI were previously described[12]. The available ADNI data was combined and organized from multiple files into one main file using custom Python scripts and widely available large-scale data processing modules (iPython, NumPy, Pandas, SciPy) installed with Anaconda

(Continuum analytics, http://www.continuum.io/anaconda). Baseline PET-AV45 exams from all unique participants were compiled first and then the patient's associated data was added.

2.1.1 MRI Acquisition

Structural MRIs were acquired at ADNI sites equipped with 3T MRI scanners using 3D MP-RAGE or IR-SPGR T1-weighted sequences with sagittal slices and voxel size of $1.1\times1.1\times1.2mm^3$, as described online (http://adni.loni.usc.edu/methods/documents/mri-protocols). A designated center quality controlled the MP-RAGE/IR-SPGR images and corrected for system-specific image artifacts such as geometry distortion, B1 non-uniformity, and intensity inhomogeneity[13].

2.1.2 Regional Atrophy Information

Baseline T1-weighted brain MRI within 60 days of their first PET-AV45 scan date was used for evaluation. An algorithm (FreeSurfer, https://surfer.nmr.mgh.harvard.edu/) was used to automatically parcellate of the brain into 86 distinct anatomical regions (Appendix A lists regions in detail). Each FreeSurfer parcellation was visually checked for anatomical accuracy. Only subjects that passed the overall FreeSurfer quality check were included in the model training and testing. The parcellated volumes were divided by the subject's total intracranial volume to account for gross differences in head size. The values were then standardized (z-score) with respect to an average of healthy controls (HC) who are cognitively intact on neuropsychological testing, lack Apolipoprotein E (APOE ε4) copies, and are Aβ-negative on the PET-AV45 exam.

2.1.3 Graph Representation

The white matter connectome was derived by diagnostic group and constructed by taking the mean of diffusion MRI data from a subset of 263 ADNI subjects. A sample size of 86 HC subjects (37 men,

49 women, 72.3±5.5 years), 128 MCI subjects (79 men, 49 women, 72.5±7.2 years), and 49 AD subjects (29 men, 20 women, 74.9±8.5 years) were used to create the respective connectivity information in the form of tractograms. Raw diffusion weighted MRIs were corrected for image artifacts including eddy current, motion, and echo planar imaging distortions using FSL toolbox[14-16]. A single diffusion tensor was modeled at each voxel in the brain from the corrected diffusion weighted MRIs using CAMINO toolbox[17]. Afterward, deterministic simple whole white matter streamlining has been applied on the DTI using CAMINO software[17]. The tissue masks from FreeSurfer processing was rigidly registered to the first frame of the diffusion weighted MRI and used in the white matter tractography. Subject-specific FreeSurfer anantomical parcellations mapped in the DTI subject space is used to calculate the ROI-ROI connectivity matrix.

DTI and graph theory methods provide a means to probe the organization of whole-brain white matter networks. DTI elucidates the brain's structural organization using information from the diffusion signal to calculate the most likely direction of water directional diffusion[18]. Probabilistic tractography (as described above) was used to estimate the connectivity matrix, or equivalently, a graph, representing the brain's anatomic network. The network model consists of nodes corresponding to brain regions and edges reflecting connections between nodes[19].

The resulting matrices in the current study were of size 86 x 86, per the 86 cortical and subcortical structures from the FreeSurfer (Desikan-Killarney atlas) gray matter parcellation. To control for inter-subject variance in total fiber count, the number of connections between each tract was divided by the total number of tracts based on each subject's tractography data. The subsequent undirected connectomes were not thresholded and kept as weighted. These networks were represented as a weighted adjacency matrix, $A \in \mathbb{R}^{86 \times 86}$.

2.1.4 PET Acquisition

The radiochemical synthesis of AV45 was overseen and regulated by Avid Radiopharmaceuticals and distributed to the qualifying ADNI sites. PET imaging was performed at each ADNI site according to standardized protocols. The AV45 protocol entailed the injection of 10 mCi of tracer followed by an uptake phase of 50 min. At 50 minutes subjects were positioned in the scanner and 4×5 min frames of emission data collected. PET/CT scans preceded these acquisitions with a CT scan for attenuation correction; PET-only scanners performed a transmission scan following the emission scan. All AV45-PET scans underwent a rigorous quality control protocol and were processed to produce final images with standard orientation, voxel size, and 8 mm^3 resolution[20].

2.1.5 Global Aβ Burden Analysis

Each AV45-PET scan was analyzed in native space using participant's structural MRI acquired closest to the time of the [^{18}F]-florbetapir PET scan as described previously[21]. Based on FreeSurfer cortical parcellation, the averaged cortical AV45 SUVR in lateral and medial frontal, anterior and posterior cingulate, lateral parietal, and lateral temporal cortical gray matter regions normalized with respect to average uptake from a composite reference region (including the whole cerebellum, pons/brainstem, and eroded subcortical white matter regions) was used as an index of global cortical AV45 burden. Furthermore, subjects were characterized as AV45-positive or AV45-negative based on a threshold value of 0.79 as published previously[22].

2.1.6 AD predictors

An additional 9 AD-related predictors was also included for analysis. Demographic information included age, gender, and years of formal education. Genetic factors included APOE ε4 and its measure was described as the number of copies: 0,1, or 2. Baseline clinical assessments include

Alzheimer's Disease Assessment Scale (ADAS), Executive Functioning (EF), Mini-Mental State Examination (MMSE), and clinical dementia rating (CDR).

2.1.7 Data Export

Regional atrophy data, AD-related predictors, and AV45-PET labels were segregated into diagnostic groups: HC, MCI, and AD. The distribution of Aβ-positive and Aβ-negative was determined within each diagnostic group and this distribution was maintained when creating a training, validation, or test set. Random sampling and reordering of the data was also performed prior to exporting these lists. Data within each diagnostic group were divided into 50% training, 20% validation, and 30% test sets for hyperparameter search and model optimization. Similarly, a training and test set for model performance evaluation was 60% and 40% respectively.

2.2 Graph Convolutional Network

A graph classifier via a convolutional neural network previously demonstrated by Defferrard et al was implemented and built upon[10]. A graph convolutional network (GCNs) is a generalization of classical CNNs from low-dimensional regular grids, where neuroimaging data is represented, to high-dimensional irregular domains, such as brain connectomes, represented by graphs. To properly use this method, the number of nodes of the input and succeeding hidden layers need to be divisible by a factor of 2. This allows for the pooling operation to group nodes of the previous layer to the next layer evenly. Because the graph was initially setup with 86 nodes to be fed as input, ten additional empty nodes were added to act as placeholders and allow for proper pooling. These placeholder nodes do not significantly alter the original information passing through the network. The main neural network components consist of two convolutional layers and a fully connected layer (Figure 1). Graphs that flow into each convolutional layer gets filtered, then

coarsened to less nodes (pooling operation), and then thresholded using a rectified linear unit function to produce an activation map. A fully connected or dense layer was append to the 2nd pooling layer which is then fed into a logistic regression layer just prior to prediction: Aβ-negative or Aβ-positive. This framework was written in Python and used the Python API for TensorFlow (Version 1) to build and run predictions.

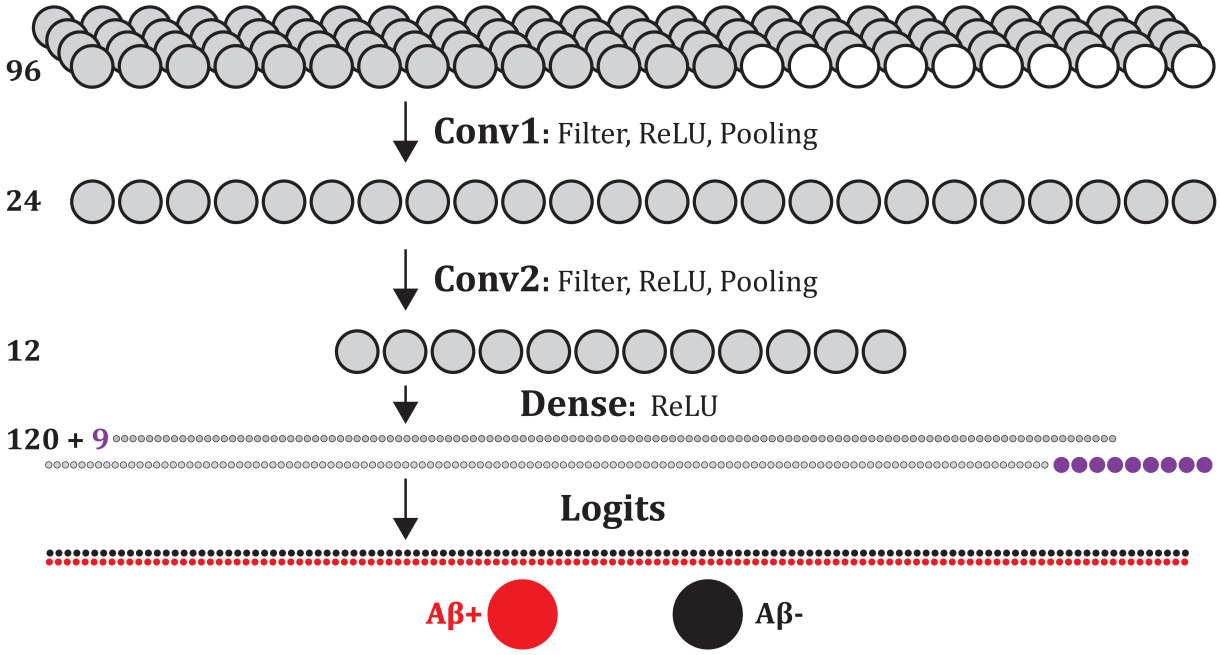


Figure 1. Graph Convolutional Network. Grey nodes are actual data flowing through the network and the extra nodes with empty values are white. The numbers on the left indicate the number of nodes. The purple nodes show the AD-related predictors appended just prior to the logistic regression layer if applicable.

Graphs are abstract data structures and to define and perform convolutions in the nodal domain is cumbersome (Figure 2). Instead, graphs passing through the network are analyzed in its spectral counterpart and the development of convolutional filters used for learning and optimizing the predictive model was accomplished through localized fast spectral filters in the fourier domain[10].

Models in this work used Chebyshev polynomials to filter the graph signals. The extent of local spectral filters was defined by the degree of a polynomial or number of coefficients in the polynomial matrix; the more polynomial coefficients, the larger the filter.

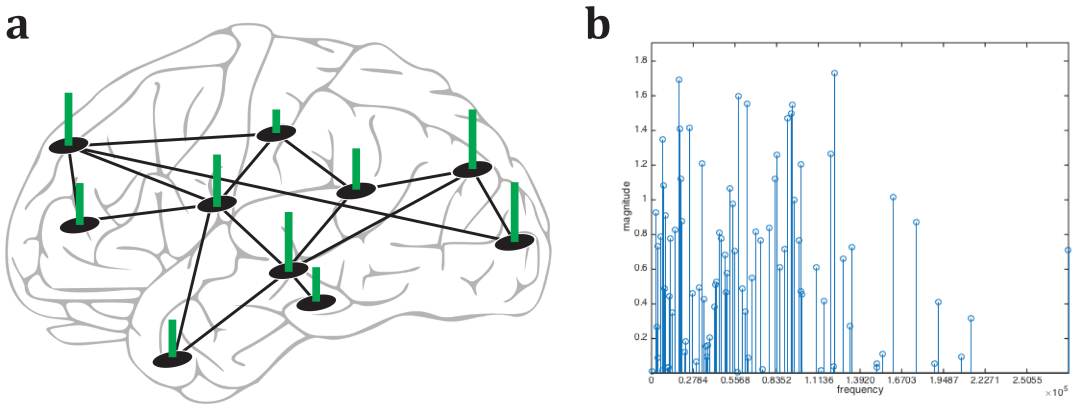


Figure 2. Graph Signal in the Nodal and Spectral Domain. (a) Nodal (b) Spectral

2.3 Hyperparameter Search and Model Optimization

Hyperparameters are values that need to be set prior to training the model, but the most optimal settings are not known *a priori*. A random search was noted to be as effective as a grid search and thus was used in this work[23]. The ranges of the hyperparameter search is listed in Table 1. An initial group of 200 iterations was performed with a broad range and coarse values. The prediction accuracy of the training and validation set was evaluated to identify are more narrow hyperparameter range. A random search was then performed again with another 500 iterations using a narrower range with finer values.

Batch size	Filter Number	Filter Size	Dense Layer	Learning Rate	L2 Regularization
2–10	5–50	5–50	64–256	1e-1 – 1e-6	1e-1 – 1e-6

Table 1. Searched Hyperparameters

Training the model occurred through a series of steps. Each step involved a batch of samples that flowed through the network to make a prediction that was compared to the AV45-PET ground truth label. The error of the model prediction was described with a loss function. Gradients with respect to all the weights in the network were computed by propagating the prediction error backwards through the neural network. The weights used for prediction in this training step were then updated accordingly with the aim to minimize the prediction error for the next batch of samples. An optimizer was used to determine how much to change the weights during the training process. The learning rate described the amount to adjust the weights of the network. Two optimizers were experimented with: gradient descent and ADAM. In gradient decent, all weights in the network were adjusted with a global learning rate and the learning rate attenuated as specified by the number of decay steps. In contrast, the ADAM optimizer adjusted the learning rate on a per weight basis such that each weight will be adjusted by different amounts. A L2 regularization method was used during training to prevent over fitting the training data.

Training may not progress through all the specified training steps and was programmed to cease early when a criteria was met. Training ceased when the value of validation loss of the current step was larger than the mean of the last 10 evaluated steps. Similarly, when training without a validation set, training ceased when the training loss of the current step was larger than the mean of the last 10 evaluated steps.

2.4 Model Assessment

Selection of the best performing models was based on outcome of the test and validation set. Each model was trained on 10 independent and randomly sampled training sets. After training the models performance was evaluated on 10 independent and randomly sampled test sets. The mean and standard deviation of the sensitivity, specificity, and predictive value was calculated. In each

diagnostic group, a t-test was performed to compare any significant differences ($p<0.05$) between model types: I) regional atrophy values, II) atrophy with AD-related predictors added to the fully connected layer (see Figure 1), and III) atrophy and AD-related predictors as input. Model Types I and II had a 2D input of [number of samples x 96] and Model Type III had a 3D data input [number of samples x 96 x 10]. In the case of model type III, each AD-related predictors had in its own dimensions with the same values repeated in every node of the graph.

The trained MCI models of each type and instance were used to predict Aβ-positivity on a cohort of subjects with their 3 month MRI follow up exam. The mean and standard deviation of the sensitivity, specificity, and predictive value was also calculated.

2.5 Hidden Layer Outputs

The best performing models with high prediction accuracy in the test set was used to study the outputs of the learned filters from the first hidden layer. A batch of input data from the test set was fed into these trained networks to obtain the output of each filter. The log percent between the initial and output value was calculated and exported to a comma separated value (CSV) file format. This CSV file was then imported in Matlab (R2014b, MathWorks, Natick, MA) and BrainNet was used to plot and visualize the filter response with respect to each brain region[24].

Chapter 3

Results

3.1 Model performances

Of the total 1093 unique participants with baseline AV45-PET scans, 771 had a baseline MRI scans within 60 days and was the total allotment of subjects used in this study. The best performing neural net architecture was different among diagnostic groups and model types within each diagnostic group. An example of GCN architecture used for evaluation is reported in Table 2. Preliminary results identified that the ADAM optimizer performed better than gradient descent, usually producing ~5% greater accuracy on the validation and test set, and was the optimizer used for evaluation.

Diagnostic Group	Healthy	MCI	Alzheimer's Disease
Batch size	6	4	5
Number of Filters [1st Layer, 2nd Layer]	[11, 11]	[10, 5]	[11, 5]
Filter Size [1st layer, 2nd layer]	[14, 12]	[4, 4]	[14, 14]
Fully Connected Layer	64	128	64
Optimizer	ADAM	ADAM	ADAM
Learn rate	0.01	0.001	0.01
L2 Regularization	0.1	0.001	0.001

Table 2. Example of Top Performing GCN Architecture and Settings

The results of the model evaluations are reported in Table 3. Prediction accuracy, sensitivity, specificity and predictive value evaluated on the respective test sets generally increased with the addition of AD-related predictors either fed into the dense layer or at the beginning of the neural network.

The Normal III model with all patient descriptors fed as an input to the GCN had the highest prediction accuracy that was significant compared to using atrophy by itself (Normal I). This model also had the lowest sensitivity and highest specificity, both were significantly different compared to the Normal I model. There was no significant difference between positive predictive value (PPV) and negative predictive value (NPV) across all three healthy control models.

It made no difference where the AD-related predictors was fed in MCI models and both MCI II and MCI III models produced similar results. Both these models however was significantly better than using atrophy descriptors alone (MCI I), except for PPV of the MCI III model. Each of these MCI models showed a higher sensitivity than specificity and showed a slightly higher NPV than PPV.

All three AD models performed just as well on prediction accuracy, sensitivity, and PPV. The AD models in particular did now show any major difference with the addition of the AD-related predictors. No significant difference was observed between AD II and AD III models. On average, the AD II model was able to identify more Aβ-negative than the other AD models.

Model	Test Set (n)	Accuracy	Sensitivity	Specificity	PPV	NPV
Normal I	149	61.3 ± 2.8	27.0 ± 7.1	76.2 ± 5.1	32.9 ± 4.9	70.6 ± 1.6
Normal II	149	63.2 ± 9.0	43.0 ± 34.8	72.0 ± 27.8	52.9 ± 23.0	77.5 ± 8.5
Normal III	149	69.2 ± 2.2	6.0 ± 6.8	96.6 ± 4.3	31.2 ± 34.3	70.3 ± 1.1
MCI I	152	56.6 ± 2.9	65.6 ± 8	47.5 ± 6.9	55.6 ± 2.4	58.3 ± 4.5
MCI II	152	69.4 ± 2.5	80.0 ± 9.0	58.7 ± 12.9	66.7 ± 4.6	75.9 ± 5.4
MCI III	152	67.6 ± 3.5	83.6 ± 6.7	51.5 ± 13.0	63.8 ± 4.1	76.9 ± 4.5
AD I	40	88.8 ± 1.8	98.3 ± 2.0	2.5 ± 8.0	90.0 ± 0.8	2.5 ± 15.8
AD II	40	90.5 ± 1.1	100 ± 0	5 ± 10.5	90.5 ± 1.0	5 ± 42.2
AD III	40	89.3 ± 1.2	99.1 ± 1.4	0	89.9 ± 0.1	0

Table 3. Model Results. Mean and standard deviation values reported.

An improvement in prediction accuracy, sensitivity, specificity, and predictive value was noted in all three MCI models evaluated on MCI subjects with their 3 month MRI T1 scan. These results are listed in Table 4 along with models' baseline MRI results for comparison. The amount of 3 month follow up MRI exams used for evaluation that met the filtering criteria specified above was 33.

Model	**Test Set (n)**	**Accuracy**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
MCI I	152	56.6 ± 2.9	65.6 ± 8	47.5 ± 6.9	55.6 ± 2.4	58.3 ± 4.5
MCI I *3 month*	33	63.9 ± 3.6	74.1 ± 5.0	53.1 ± 8.5	62.9 ± 3.8	65.9 ± 3.9
MCI II	152	69.4 ± 2.5	80.0 ± 9.0	58.7 ± 12.9	66.7 ± 4.6	75.9 ± 5.4
MCI II *3 month*	33	76.1 ± 3.7	76.5 ± 5.9	75.6 ± 12.0	78.2 ± 8.2	75.4 ± 2.7
MCI III	152	67.6 ± 3.5	83.6 ± 6.7	51.5 ± 13.0	63.8 ± 4.1	76.9 ± 4.5
MCI III *3 month*	33	76.9 ± 6.5	77.6 ± 7.2	78.1 ± 17.0	81.1 ± 10.7	76.8 ± 4.7

Table 4. Comparison of MCI Model Results on Baseline and 3 Month MRI Data

3.2 Learned Hidden Features

This section describes learned filters from trained models of each diagnostic group and their response to regional atrophy data. In these illustrations, the size of the nodes depict the magnitude and the color indicates the direction of the change with respect to the input values. All examples are outputs from the first convolutional layer.

Learned filters from the first hidden layer tend to increase regional atrophy values and activate more networks (more red nodes) as graph models progress from healthy to Alzheimer's disease. An example of brain networks largely affected by a filter from a HC, MCI, and AD model is shown in Figure 3. The nodes in these plots represent the top 50% of the maximum absolute value observed in the graph. For example, healthy control nodes with values greater than 0.0865 (50% of 0.173) was plotted. In all three of these cases the left postcentral and right entorhinal was reduced.

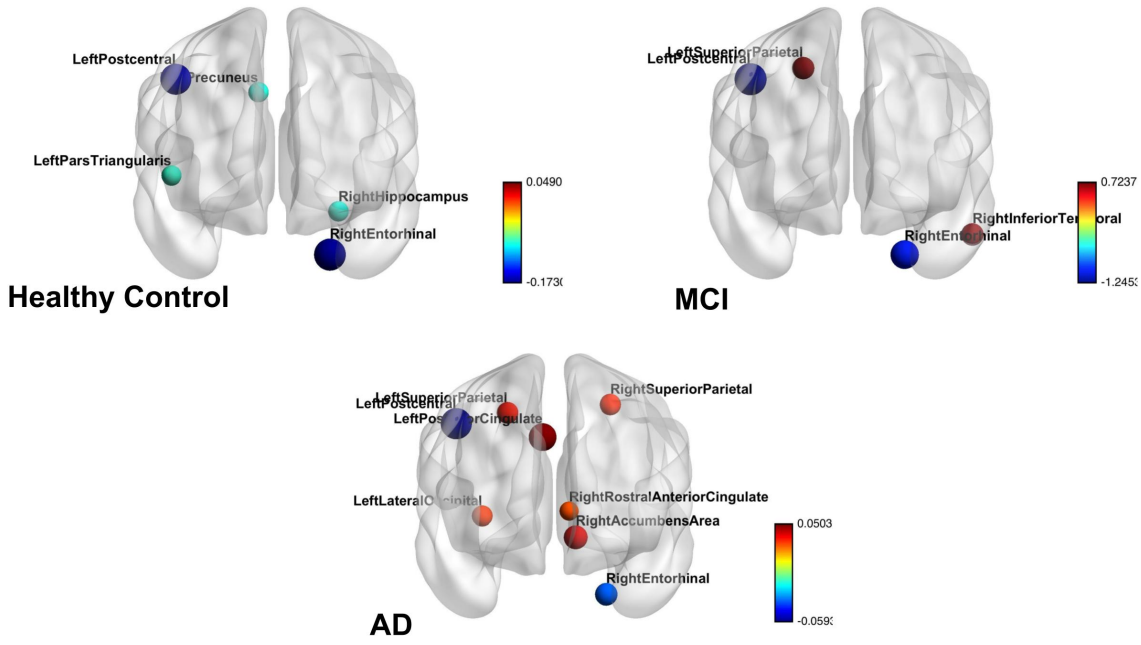


Figure 3. Example of Highlighted Brain Networks from Learned Filters

3.2.1 Healthy Control Filter Outputs

Outputs of five different filters in response to regional atrophy data is shown in Figure 4 to Figure 8. In most of these filters, the largest magnitude in change is negative with less enhanced regions than reduced regions.

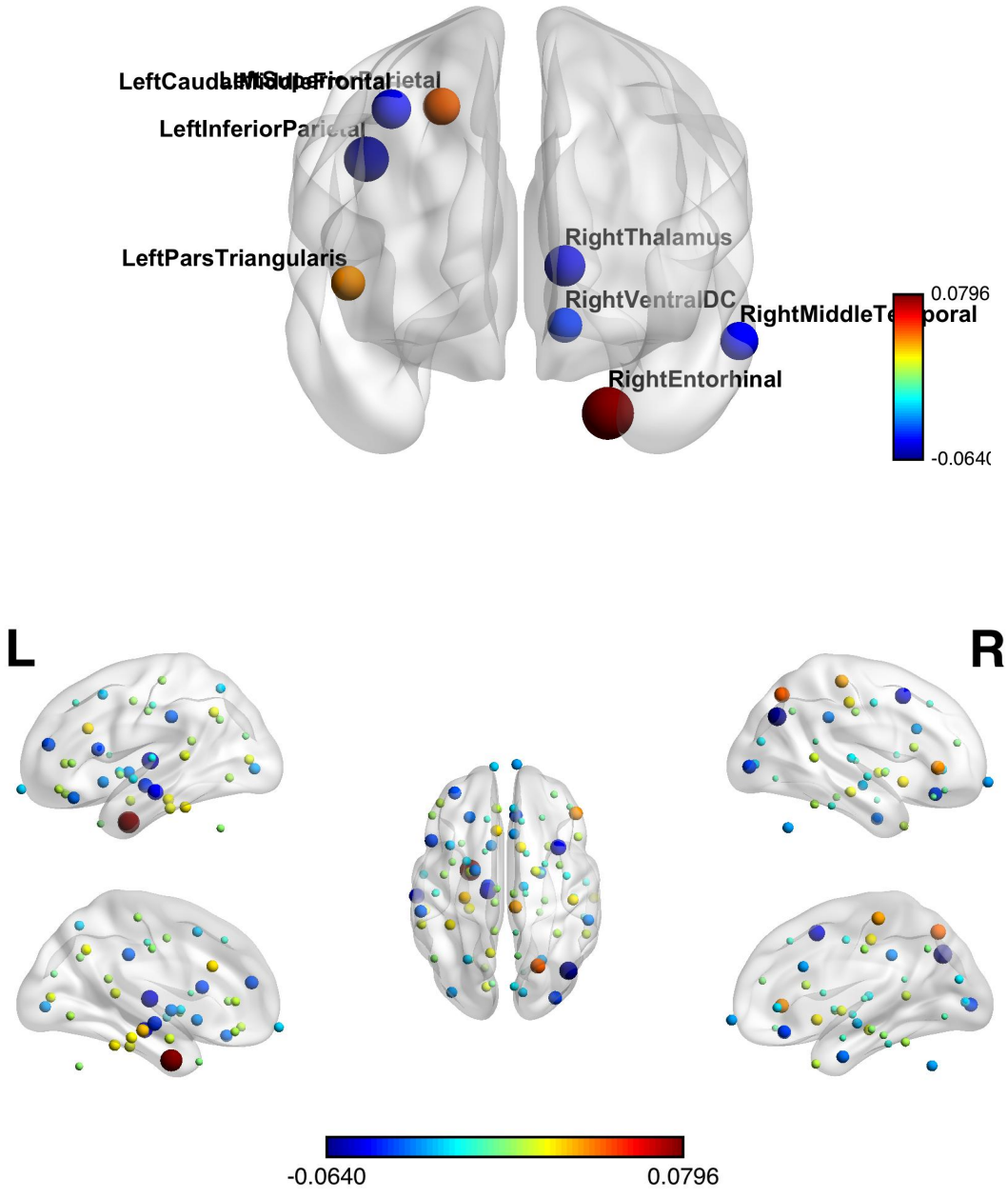


Figure 4. Convolutional Filter #1 Output from a Healthy Control Model

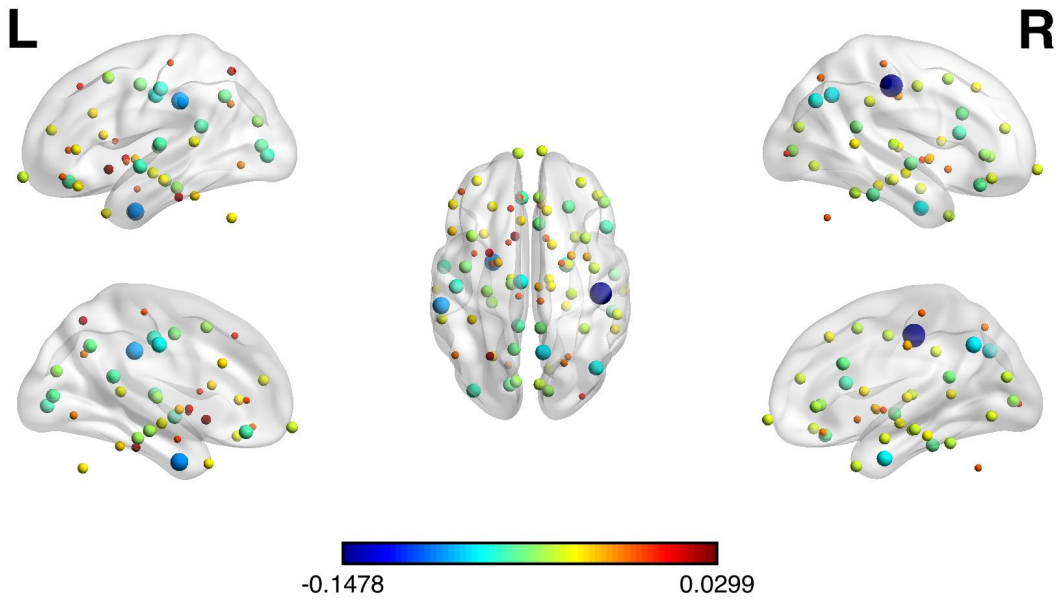


Figure 5. Convolutional Filter #2 Output from a Healthy Control Model

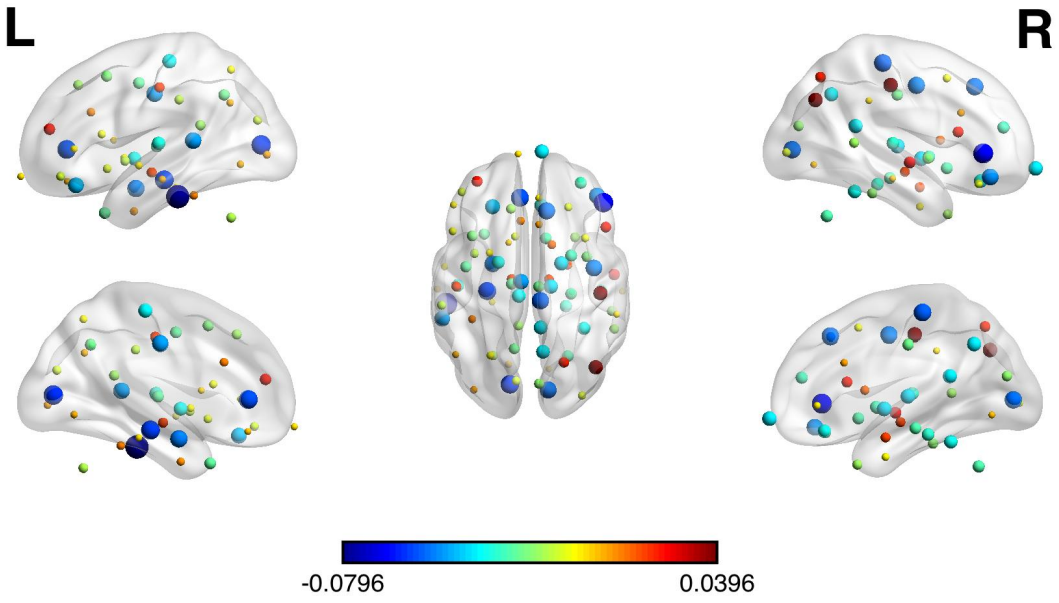


Figure 6. Convolutional Filter #3 Output from a Healthy Control Model

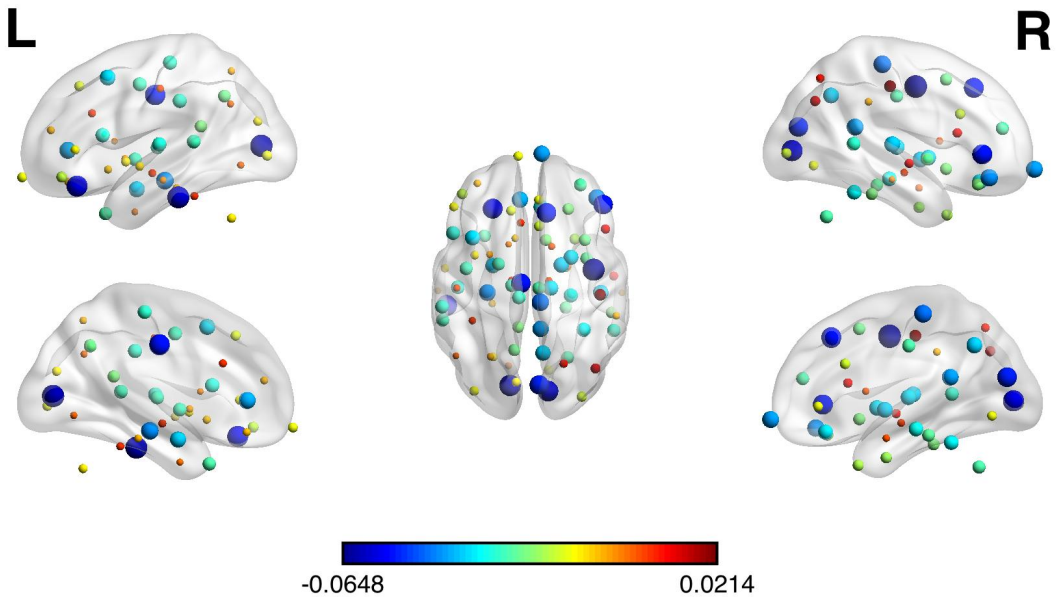


Figure 7. Convolutional Filter #4 Output from a Healthy Control Model

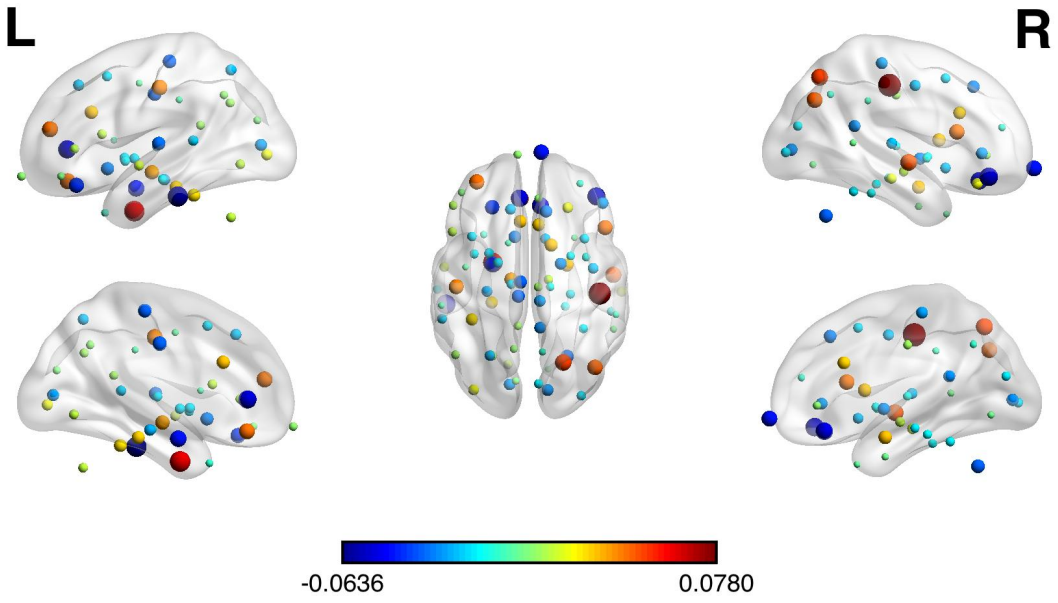


Figure 8. Convolutional Filter #5 Output from a Healthy Control Model

3.2.2 Mild Cognitive Impairment Filter Outputs

Examples of five filters from a MCI model is shown in Figures 9-13. Output of these filters show a mixture of activation and reduction.

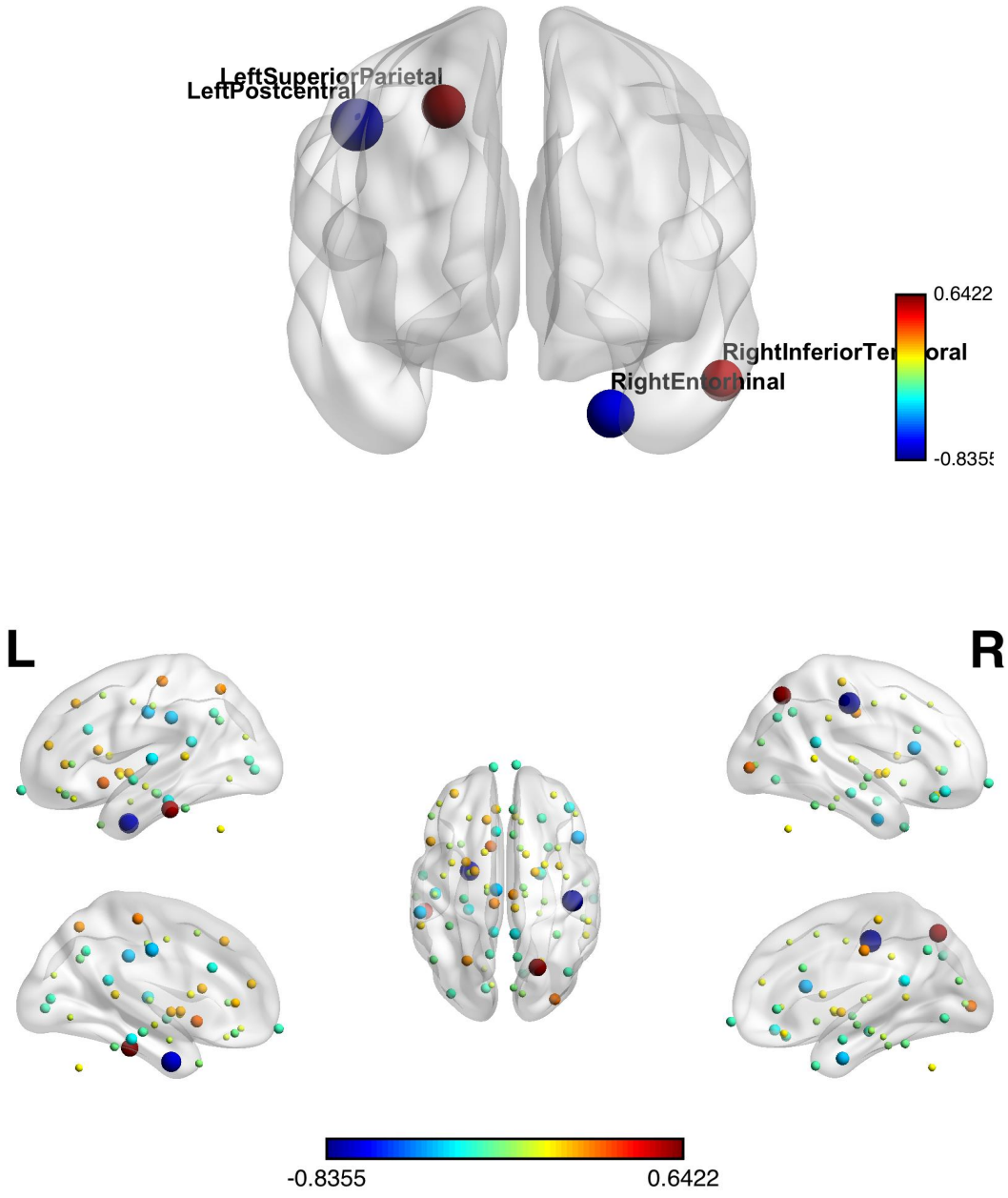


Figure 9. Convolutional Filter #1 Output from a Mild Cognitive Impairment Model

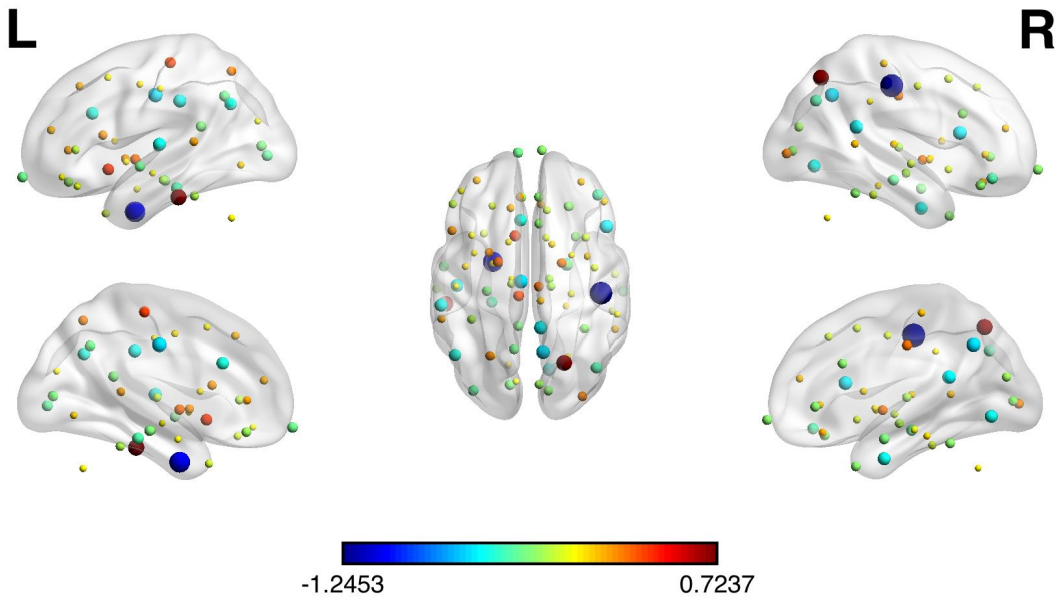


Figure 10. Convolutional Filter #2 Output from a Mild Cognitive Impairment Model

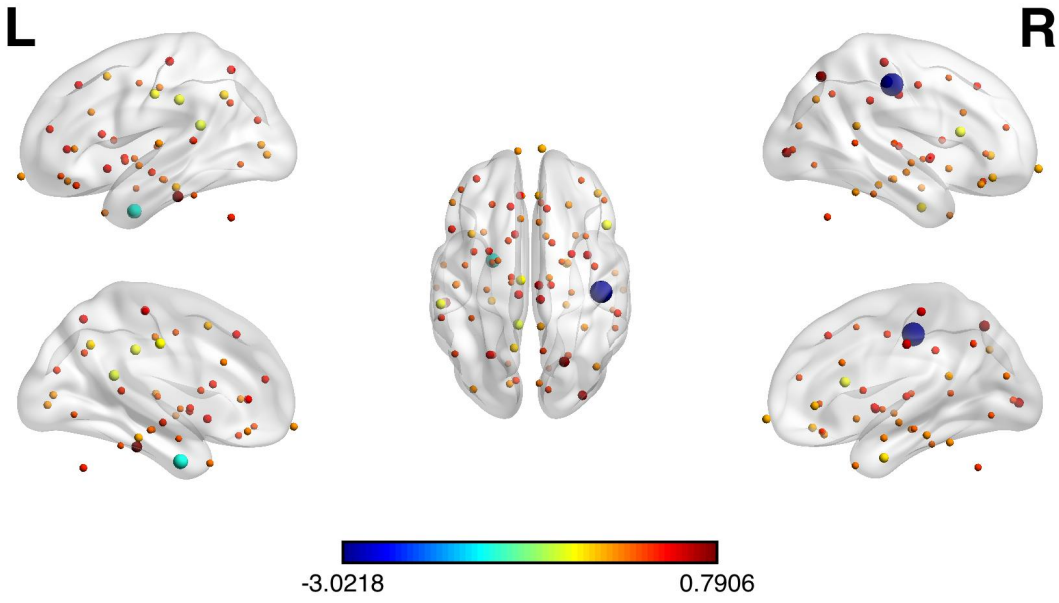


Figure 11. Convolutional Filter #3 Output from a Mild Cognitive Impairment Model

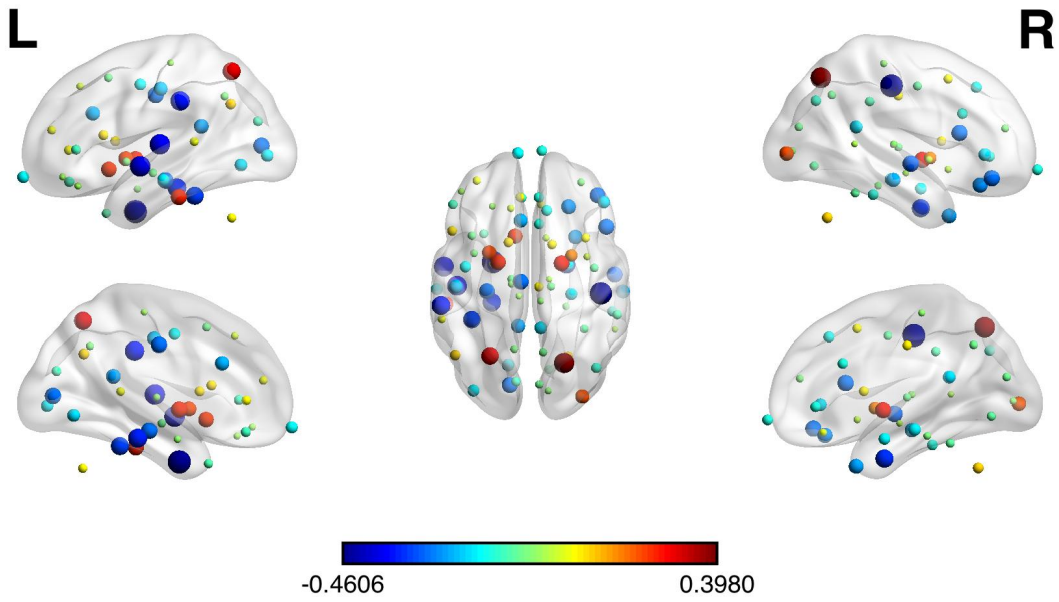


Figure 12. Convolutional Filter #4 Output from a Mild Cognitive Impairment Model

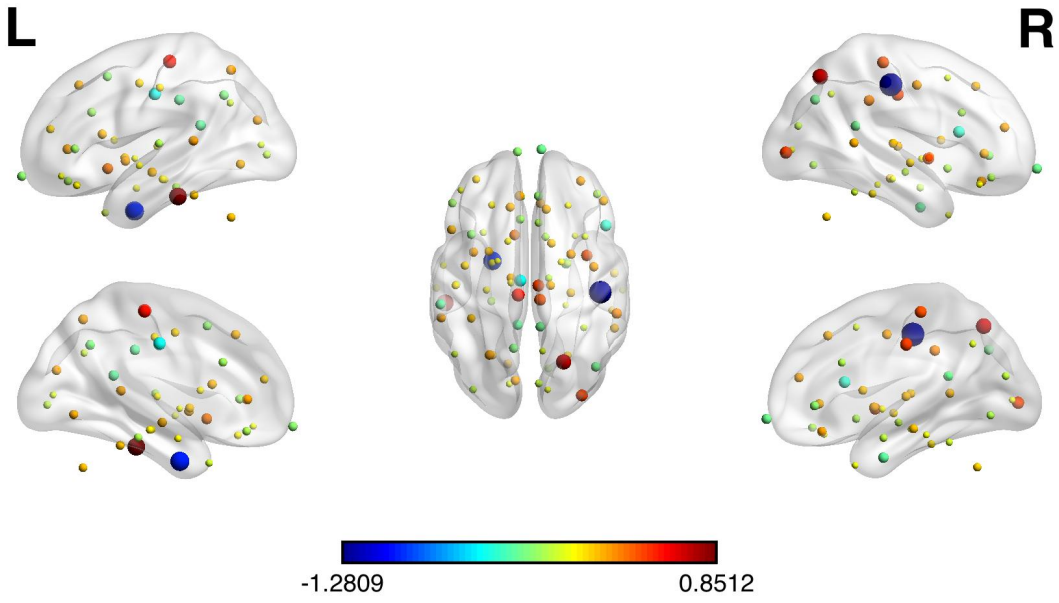


Figure 13. Convolutional Filter #5 Output from a Mild Cognitive Impairment Model

3.2.3 Alzheimer's Disease Filter Outputs

Outputs from five different filters of an AD model are shown in Figures 14 to Figure 18. Two of these filters show similar filtering patterns (Figure 14 and Figure 15). Atrophy in the cortical regions are more apparent these examples, some filters show more regions than others.

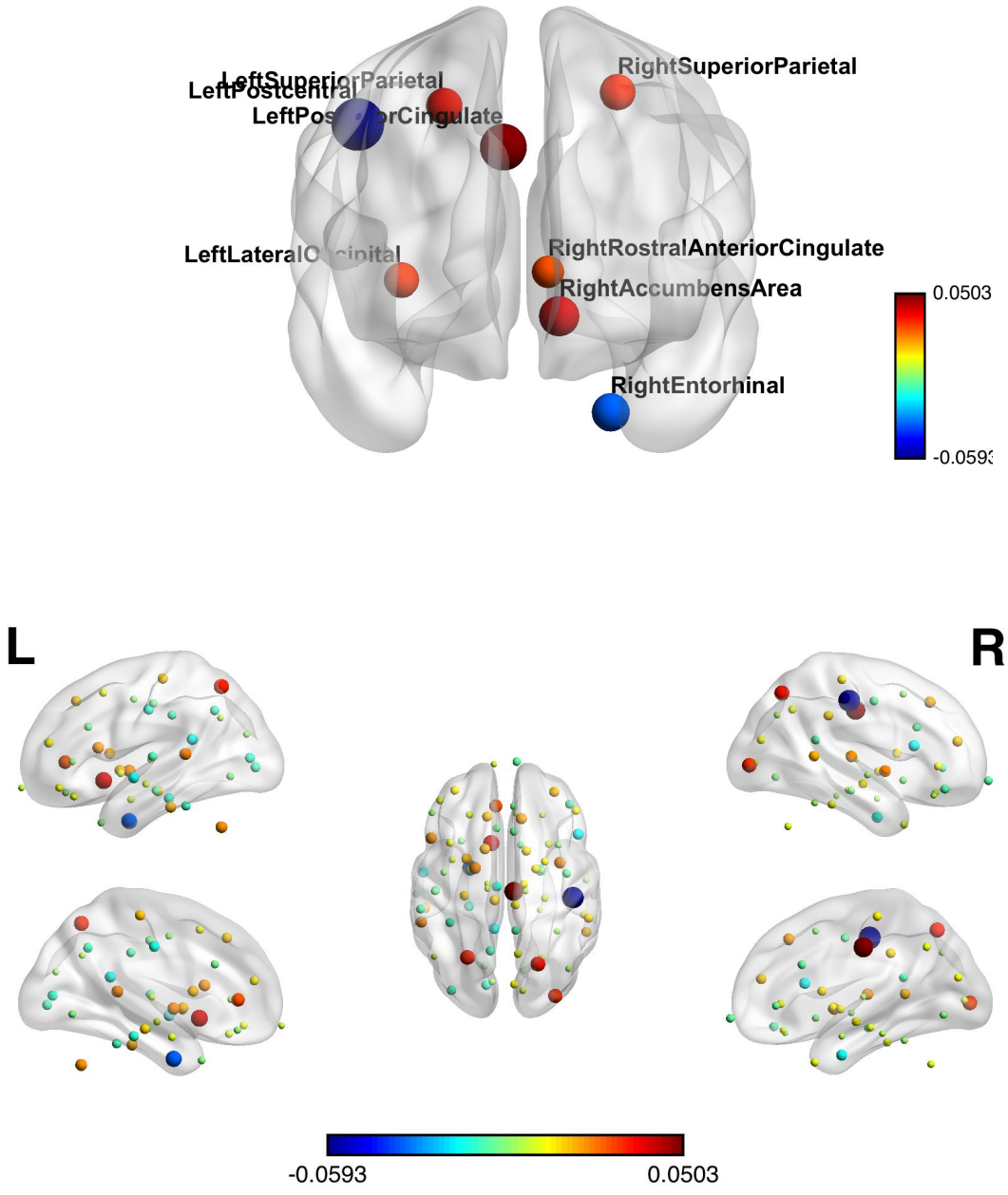


Figure 14. Convolutional Filter #1 Output from an AD Model

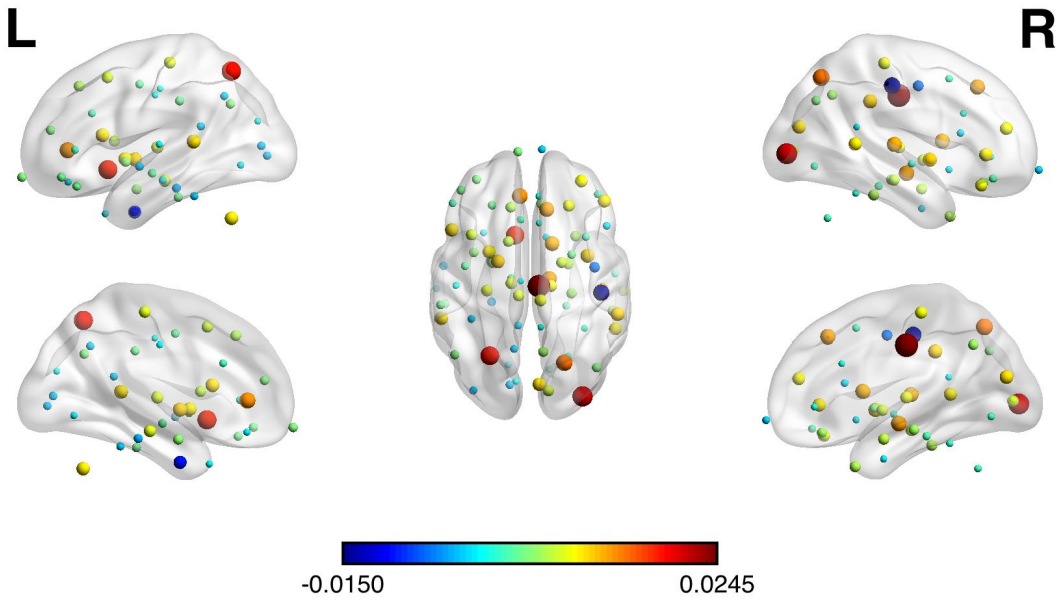


Figure 15. Convolutional Filter #2 Output from an AD Model

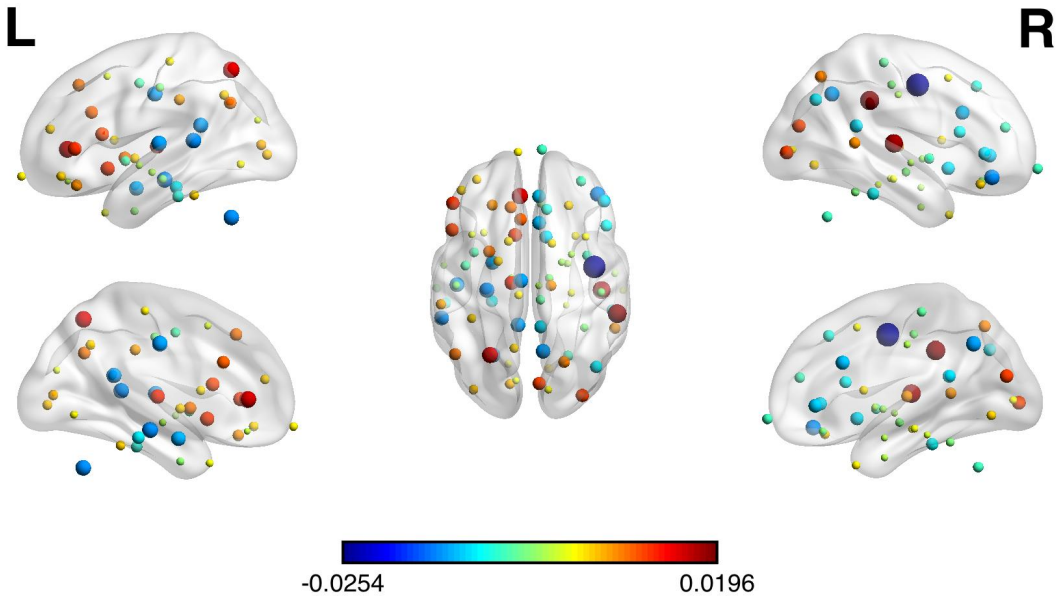


Figure 16. Convolutional Filter #3 Output from an AD Model

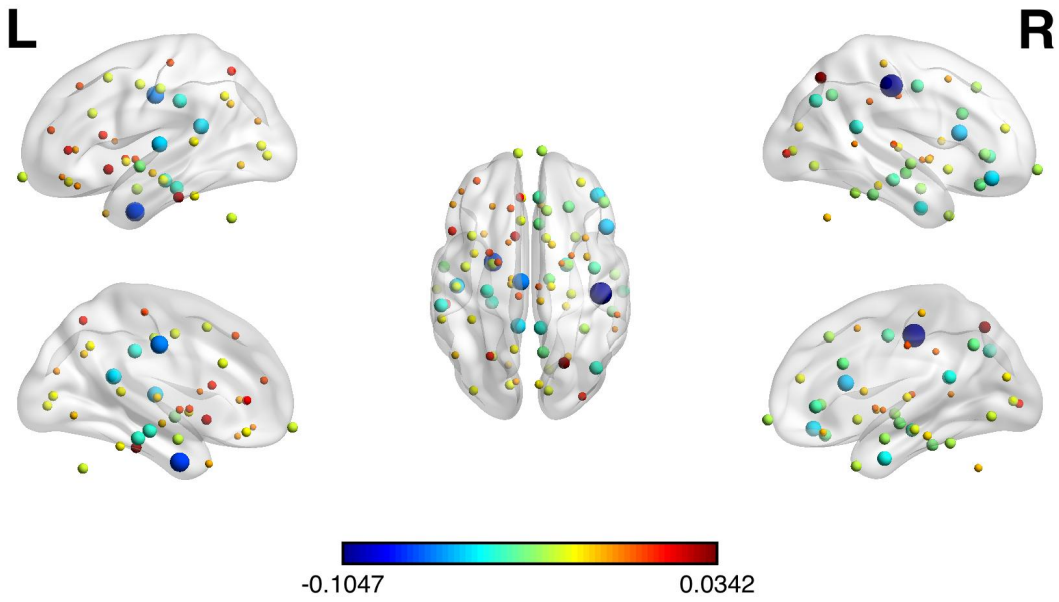


Figure 17. Convolutional Filter #4 Output from an AD Model

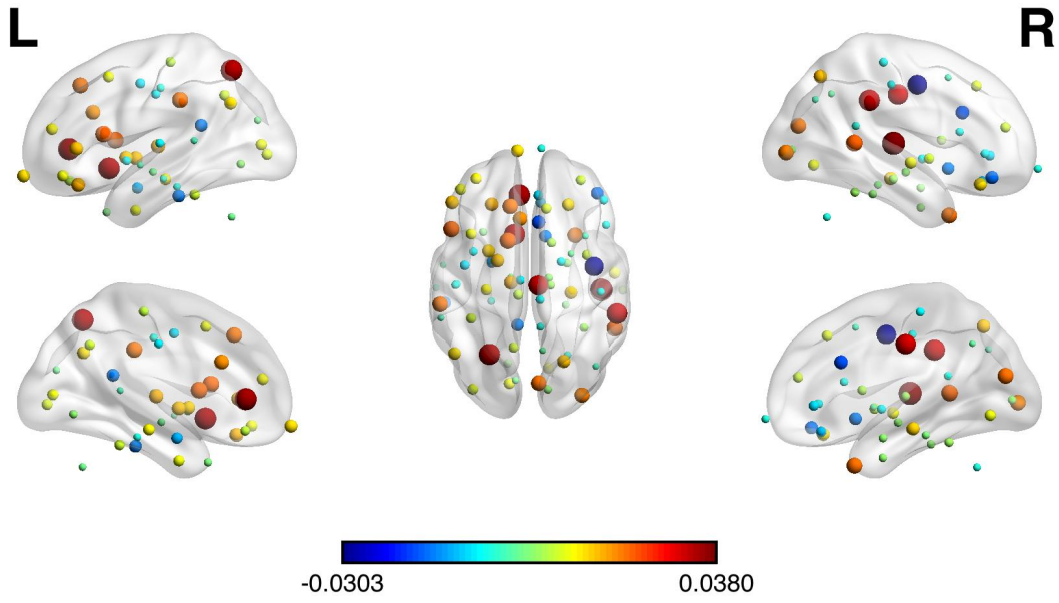


Figure 18. Convolutional Filter #5 Output from an AD Model

Chapter 4

Discussion

4.1 Summary of Findings

The dense image information from structural MRI can be reinterpreted as graph which can be fed into a convolutional neural network to predict whether or not a person would test positive in the AV45-PET exam. This method can be used as a practical adjunct to deciding the next course of action for the patient. With proper validation of the model, the costs involved with using Aβ targeted radiopharmaceuticals and amount of related diagnostic tests like AV45-PET and CSF assay could be minimized. Additionally, areas with limited access to Aβ radiotracers and PET scanners or persons hesitant with getting their CSF probed would benefit from using this predictive modeling method.

Training the model and running predictions with it can be performed with today's off the shelf laptops or desktops without the need of specialized resources like using graphics processor units or cloud services dedicated to these kinds of computations. The network which describes connectivity between the 86 brain regions in this study is sparse and allows for even fewer convolutional computations. By comparison, convolutions in deep learning frameworks with images for classification would normally require the filter to transverse through all or most of the image without the option to skip unnecessary regions. An idea maybe worth exploring would be to combine a graph neural network coupled with its image neural network counterpart for higher predictive power and more insight into understanding the progression of Alzheimer's disease. Activated regions in the graph network could specify more important regions to look at in the image network or visa versa.

The learned features or filters provide clues as to which regions are pertinent for classification and disease staging. The filter output from the first hidden layer directly corresponds with the original input data making it easier to interpret as it is a one to one relationship. Additionally, these filters may identify characteristic networks or groups of regions that could be attributed to a particular disease stage. In contrast with neural networks trained in the task of image classification, the first convolutional filters provide simple but meaningful descriptors useful for classification like the detection of edges. A similar but not direct relationship could also be noted with the first convolutional layer filter of the graph neural network. While analyzing the 2nd hidden layer filters may offer an added value, they are harder to interpret as the describe even more abstraction of the input data that depends on the activations from the first hidden layer. A detailed look into the interpreting the complete hierarchal learning of the hidden layers may offer additional insight to how combined brain regions or networks contribute to classification, but this analysis will be left for future endeavors.

The sensitivity, specificity, and predictive value was generally enhanced with the addition of AD related cofactors. The MCI model improved the most with the addition of the AD related cofactors and was significant across all measures. With the MCI group being the heterogeneous among other diagnostic groups, using atrophy descriptors alone may not offer high enough predictive value with this method. Even with the AD related cofactors, the average prediction accuracy was ~70% and may require additional biomarkers to train and improve predictions. The least affected with the addition of AD related factors was noted in the AD model. This result could possibly mean that the use of atrophy descriptors provided enough context to predict Aβ-positive reliably. The AD group contained small portion of Aβ-negative (~10%) which was difficult for the model to identify, however some model instances was able to identify 1 of 3 Aβ-negative subjects in the test set. The healthy control model showed marginal improvement on prediction accuracy with AD-related

predictors. Nonetheless, the healthy control models can better identify Aβ-negative and a somewhat reasonable portion of Aβ-positives. Normal III model in particular showed a very high specificity and NPV, which showed that the model learned more features that characterize a healthy brain.

Most models reported in Table 3 showed a relatively marginal variance in their measures of predictive value, sensitivity, and specificity. One model that deviated from this was the Normal II model where there was a large standard deviation in sensitivity and PPV. This larger variance could be attributed the random sampling of the training and test set where some datasets had more difficult subjects to classify. Moreover, every model begins training with randomly chosen initial values of the networks' parameter weights and this affects the learning trajectory. A model with the same training and validation sets, for example, will have slightly different outcomes if ran again *de novo*. The values being initialized however are from a tight distribution with a mean of 0.0 and within a standard deviation of 0.1.

Trained MCI models are robust enough to predict Aβ-positivity on 3 month MRI data with comparable accuracy to baseline MRI predictions, albeit the evaluation was performed on a smaller test set. The narrow list could be attributed to the variability of when participants actually come in for the 3 month MRI exam and also the strict tolerance of 15 days when aggregating the data. A larger test set size is warranted, but this at least demonstrates the potential flexibility of the learned MCI models.

4.2 Limitations

This model assumes that the density of tracts describing the connectivity between brain regions is consistent among all subjects at a particular disease stage, that each diagnostic group could be described as one network topology. For example, subjects diagnosed with MCI will have the same connection strength between the amygdala and the hypothalamus as others who are diagnosed with MCI. Incorporating the person's unique network topology in training the model to predict Aβ-positivity may be of value especially if diffusion MRI becomes readily available and implemented as a standard routine of care. Another aspect of this modeling method to note is that the learned filters cannot be transferred to other graph structures so any changes such as an addition of another node or node connection would require re-training and optimizing the model. The sample size is limited to participants in the ADNI study and could be larger.

4.3 Future Work

More work can be done to reap the full potential of this numerical modeling method. While this work looked at two hidden layers, the pooling combinations and number hidden layers have not been explored in detail. One thing to explore would be to augment the input and feed more graph signals into the neural network with varying regional information. Regional metabolite concentrations (i.e. N-acetyl-asparate, creatine/phosphocreatine, choline, and myo-inositol) derived from magnetic resonance spectroscopy could also be included with regional atrophy information. Another idea worth exploring would be to incorporate a multi-scaled graph whereby each node would contain another network. Each brain region can hold another network information like the interplay of protein expression, metabolite concentration, and finer structural characteristics.

References

1 Jacquemont, T. *et al.* Amyloidosis and neurodegeneration result in distinct structural connectivity patterns in mild cognitive impairment. *Neurobiol Aging* **55**, 177-189, doi:10.1016/j.neurobiolaging.2017.03.023 (2017).

2 Brettschneider, J., Del Tredici, K., Lee, V. M. & Trojanowski, J. Q. Spreading of pathology in neurodegenerative diseases: a focus on human studies. *Nat Rev Neurosci* **16**, 109-120, doi:10.1038/nrn3887 (2015).

3 Stefanacci, R. G. The costs of Alzheimer's disease and the value of effective therapies. *Am J Manag Care* **17 Suppl 13**, S356-362 (2011).

4 Siemers, E. R. *et al.* Phase 3 solanezumab trials: Secondary outcomes in mild Alzheimer's disease patients. *Alzheimers Dement* **12**, 110-120, doi:10.1016/j.jalz.2015.06.1893 (2016).

5 Salloway, S. *et al.* Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med* **370**, 322-333, doi:10.1056/NEJMoa1304839 (2014).

6 Tosun, D., Joshi, S., Weiner, M. W. & Alzheimer's Disease Neuroimaging, I. Neuroimaging predictors of brain amyloidosis in mild cognitive impairment. *Ann Neurol* **74**, 188-198, doi:10.1002/ana.23921 (2013).

7 Vlassenko, A. G., Benzinger, T. L. & Morris, J. C. PET amyloid-beta imaging in preclinical Alzheimer's disease. *Biochim Biophys Acta* **1822**, 370-379, doi:10.1016/j.bbadis.2011.11.005 (2012).

8 Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118, doi:10.1038/nature21056 (2017).

9 Quellec, G., Charriere, K., Boudi, Y., Cochener, B. & Lamard, M. Deep image mining for diabetic retinopathy screening. *Med Image Anal* **39**, 178-193, doi:10.1016/j.media.2017.04.012 (2017).

10 Defferrard, M., Bresson, X. & Vandergheynst, P. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *ArXiv e-prints* **1606** (2016). <http://adsabs.harvard.edu/abs/2016arXiv160609375D>.

11 Tosun, D., Joshi, S., Weiner, M. W. & the Alzheimer's Disease Neuroimaging, I. Multimodal MRI-based Imputation of the Abeta+ in Early Mild Cognitive Impairment. *Ann Clin Transl Neurol* **1**, 160-170, doi:10.1002/acn3.40 (2014).

12 Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* **74**, 201-209, doi:10.1212/WNL.0b013e3181cb3e25 (2010).

13 Jack, C. R., Jr. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* **27**, 685-691, doi:10.1002/jmri.21049 (2008).

14 Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23 Suppl 1**, S208-219, doi:10.1016/j.neuroimage.2004.07.051 (2004).

15 Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* **45**, S173-186, doi:10.1016/j.neuroimage.2008.10.055 (2009).

16 Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782-790, doi:10.1016/j.neuroimage.2011.09.015 (2012).

17 Cook, M., Fleming, J. J. & Buchanan, N. S. El Camino Hospital: ten years later. *Comput Hosp* **2**, 22-25 (1981).

18 Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34**, 144-155, doi:10.1016/j.neuroimage.2006.09.018 (2007).

19 Micheloyannis, S. Graph-based network analysis in schizophrenia. *World J Psychiatry* **2**, 1-12, doi:10.5498/wjp.v2.i1.1 (2012).

20 Jagust, W. J. Amyloid imaging: coming to a PET scanner near you. *Ann Neurol* **68**, 277-278, doi:10.1002/ana.22144 (2010).

21 Landau, S. M. *et al.* Amyloid-beta imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *J Nucl Med* **54**, 70-77, doi:10.2967/jnumed.112.109009 (2013).

22 Landau, S. M. *et al.* Measurement of longitudinal beta-amyloid change with 18F-florbetapir PET and standardized uptake value ratios. *J Nucl Med* **56**, 567-574, doi:10.2967/jnumed.114.148981 (2015).

23 Bergstra, J. & Bengio, Y. Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research* **13**, 281-305 (2012).

24 Xia, M., Wang, J. & He, Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One* **8**, e68910, doi:10.1371/journal.pone.0068910 (2013).

Appendix A

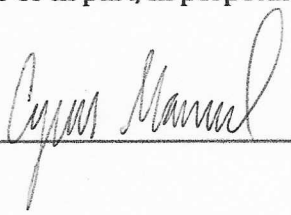
Left Cerebellum Cortex	Left Rostral Middle Frontal
Left Thalamus	Left Superior Frontal
Left Caudate	Left Superior Parietal
Left Putamen	Left Superior Temporal
Left Pallidum	Left Supramarginal
Left Hippocampus	Left Frontal Pole
Left Amygdala	Left Temporal Pole
Left Accumbens Area	Left Transverse Temporal
Left Ventral DC	Left Insula
Right Cerebellum Cortex	Right Bankssts
Right Thalamus	Right Caudal Anterior Cingulate
Right Caudate	Right Caudal Middle Frontal
Right Putamen	Right Cuneus
Right Pallidum	Right Entorhinal
Right Hippocampus	Right Fusiform
Right Amygdala	Right Inferior Parietal
Right Accumbens Area	Right Inferior Temporal
Right Ventral DC	Right Isthmus Cingulate
Left Bankssts	Right Lateral Occipital
Left Caudal Anterior Cingulate	Right Lateral Orbitofrontal
Left Caudal Middle Frontal	Right Lingual
Left Cuneus	Right Medial Orbitofrontal
Left Entorhinal	Right Middle Temporal

Left Fusiform	Right Parahippocampal
Left Inferior Parietal	Right Paracentral
Left Inferior Temporal	Right Pars Opercularis
Left Isthmus Cingulate	Right Pars Orbitalis
Left Lateral Occipital	Right Pars Triangularis
Left Lateral Orbitofrontal	Right Pericalcarine
Left Lingual	Right Postcentral
Left Medial Orbitofrontal	Right Posterior Cingulate
Left Middle Temporal	Right Precentral
Left Parahippocampal	Right Precuneus
Left Paracentral	Right Rostral Anterior Cingulate
Left Pars Opercularis	Right Rostral Middle Frontal
Left Pars Orbitalis	Right Superior Frontal
Left Pars Triangularis	Right Superior Parietal
Left Pericalcarine	Right Superior Temporal
Left Postcentral	Right Supramarginal
Left Posterior Cingulate	Right Frontal Pole
Left Precentral	Right Temporal Pole
Left Precuneus	Right Transverse Temporal
Left Rostral Anterior Cingulate	Right Insula

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  Date 9/11/17