Capstone Project: ADNI Data

Brian Collica, Ben Searchinger, Ryan Roggenkemper, James Koo

2/12/2021

Data

Access and Acquisition

The main source of our data will be DICOM format files of PET and MRI brain scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. For the purposes of this assignment, the data is coming from the Australian Imaging, Biomarkers, and Lifestyle (AIBL) database while we wait for final access approval from ADNI. Both databses are provided courtesy of the Laboratory of Neuro Imaging at the University of Southern California. It is unclear at this moment whether or not AIBL data will play a part in the final analysis, although this is a possibility we are exploring.

Exploratory Data Analysis

Downloading and Reading DICOM files

Much of the data is in DICOM format, a standardized file and metadata formatting system used in medical imaging. Other formats are also available such as NiFTI and ANALYZE, although we plan to restrict ourselves to the DICOM format.

There are many existing open-source software libraries in R and Python for reading and processing DICOM files. One such R package is `radtools` which is available on the Neuroconductor repository and also on GitHub. We have been able to successfully download PET brain scan images from AIBL and process them into R using the `radtools` package.

Below is an example work flow where we read in the entire 3D image from a directory containing 90 separate DICOM files, one for each image slice. Using the functions in `radtools`, we are able to inspect the dimensions of the data along with the metadata attributes. We can also transform the slices into a three dimensional array and inspect the results. `radtools` also facilitates viewing the actual image slices.

```
# Install Package
source("https://neuroconductor.org/neurocLite.R")
neuro_install('radtools', release = "stable", release_repo = "github")

library(radtools)

# Path to image directory
img_path <- "AIBL/10/summed.img__RSRCH_RAMLA3D-SUV/2006-10-17_13_53_08.0/I153055/"

# Read in 90 slices of PET image
PET <- read_dicom(img_path)

# Inspect image dimensions and number of slices
img_dimensions(PET)
num_slices(PET)
head(header_fields(PET))
```

```
## [1] 128 128  90
## [1] 90
## [1] "AccessionNumber"            "AcquisitionDate"
## [3] "AcquisitionTime"            "ActualFrameDuration"
## [5] "AttenuationCorrectionMethod" "BitsAllocated"
```

```
img_array <- img_data_to_mat(PET)
dim(img_array)
img_array[65:70, 65:70, 65]
```

```
## [1] 128 128  90
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]  535  569  625  670  668  612
## [2,]  517  540  594  635  646  616
## [3,]  486  494  495  547  634  613
## [4,]  459  435  432  531  630  609
## [5,]  423  404  408  506  571  554
## [6,]  359  373  380  434  488  494
```

This single 6 x 6 matrix corresponds to the region outlined below in the image representing slice 65.
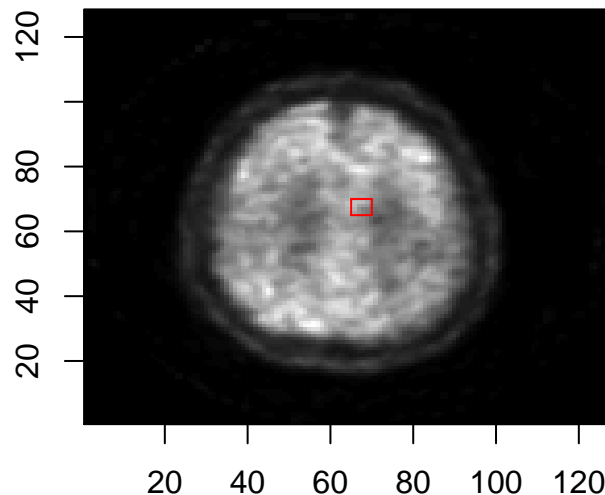
Project Outline and Analysis Plan

State of Current Research

Many papers have been published addressing the challenges associated with Alzheimer's research and using neuroimaging data. There is an overall lack of consensus regarding standardization of metrics and also which data and attributes are of importance. Additionally, there is much heterogeneity in the analysis approach employed by researchers with everything ranging from multiple linear regression models to deep learning with 3D convolutional neural networks.

Question of Interest

The specific question of interest being addressed in this project is. . .

Analysis Approach

We plan to approach this question by. . .

Project Roadmap

The project can be broken up into the following main steps:

Step Number	Due Date	Description
1	2/16	Preliminary Plan, EDA, and Project Description
2	2/23	Define the exact data to be used for analysis and have it downloaded and processed
3		
4		
5		

While each step has a specified target completion date, this is flexible depending on the unknown complications we will encounter.