CZ-1 – DL Alzheimer's

Software Design Documentation (SDD)

CS 4850 - Sections 02 & 04 - Fall 2025
Sharon Perry
September 14, 2025

 Julia Johnson Developer	 Jordan Rainford Developer

Team Members:

Name	Role	Cell Phone / Alt Email
Jordan Rainford (Team Leader)	Developer	404.277.6285 Jrain.dev@outlook.com
Julia Johnson	Developer	404.956.2715 cjohnson.julia120@gmail.com
Chen Zhao	Project Owner	470.578.6005 czhao4@kennesaw.edu

Table of Contents

1. Introduction and Overview

Alzheimer's disease (AD) is a common disorder that has many challenges for an early diagnosis. Our project aims to improve AD staging prediction by integrating deep learning models with MRI images to better AD staging predictions. We will perform comprehensive preprocessing of MRI images and electronic health record data to develop a deep learning model with improved predictive accuracy for diagnosing AD. Our goal is to gain a deeper understanding of AD progression and contribute to the development of more advanced diagnostic and prognostic tools for the disease, as well as advance clinical outcomes for people suffering from AD. In more narrow language, our main goal is to utilize machine learning techniques to improve predictive accuracy for AD.

2. Design Considerations

2.1. Assumptions and Dependencies

For our software, we are assuming that the user has access to the required ADNI, AIBL, and OASIS datasets. The initial segmentation of data also requires the user to have MATLAB R20203b installed, as other versions of the application are not able to run some of the software involved in pre-processing the data.

We are also assuming that the user has access to SPM and CAT12 software for viewing and processing of the MRI images prior to running the data through our deep learning model. While we do not have precise specification requirements yet, our model will also depend on the end user having a suitably functional computer to run the software on due to deep learning models requiring more powerful hardware, such as better GPUs and CPUs.

2.2. General Constraints

Our goal is to gain a deeper understanding of AD progression and contribute to the development of more advanced diagnostic and prognostic tools for the disease, as well as advance clinical outcomes for people suffering from AD. In more narrow language, our main goal is to utilize machine learning techniques to improve predictive accuracy for AD.

- Program must run locally on researcher's computer(s)
- End user must be able to run Python, MATLAB, SPM, and CAT12 on their computer(s)
- Model must intentionally utilize resources when handling simultaneous tasks
- Program must ensure security of sensitive health. data
- Program must run on Major Operating Systems (Windows, Linux, MacOS)
- Program must output understandable, relevant information
- Model performance must be able to be observed (Ex. Accuracy, data loss, etc.)
- Data repository and distribution requirements
- Model must be made as efficiently as possible to reduce processing time
- Testing and training data must be randomly split into 80% training and 20% testing
- Model must still be functional for increasingly larger datasets

2.3. Development Methods

Our approach to creating our software is to first understand the desired output. Then, we made a flow chart of how the data needs to be preprocessed and discussed which aspects of the data were important versus which could be ignored.

Next, we followed an existing tutorial and used a pre-existing script to help us understand the software we would be using to process the MRI images and create the model. Once the datasets were processed, we were able to begin planning on how to create the deep learning model the data would be used with.

To build the model, we have assigned each of the datasets to one team member. Each team member will work with their dataset to begin building a model that is functional with what they have. Then, each week we will exchange, combine, and edit the drafted models to determine what works well with all the data sets and which aspects of the model are not ideal for all parts.

We will plot and analyze aspects of the model, such as accuracy and data loss, to guide us in making each update to the model more accurate than the previous version. As we work on each version of the model, we will also continuously research and experiment with different kinds of neural network layer to ensure we have the best possible outcome.

3. Architectural Strategies

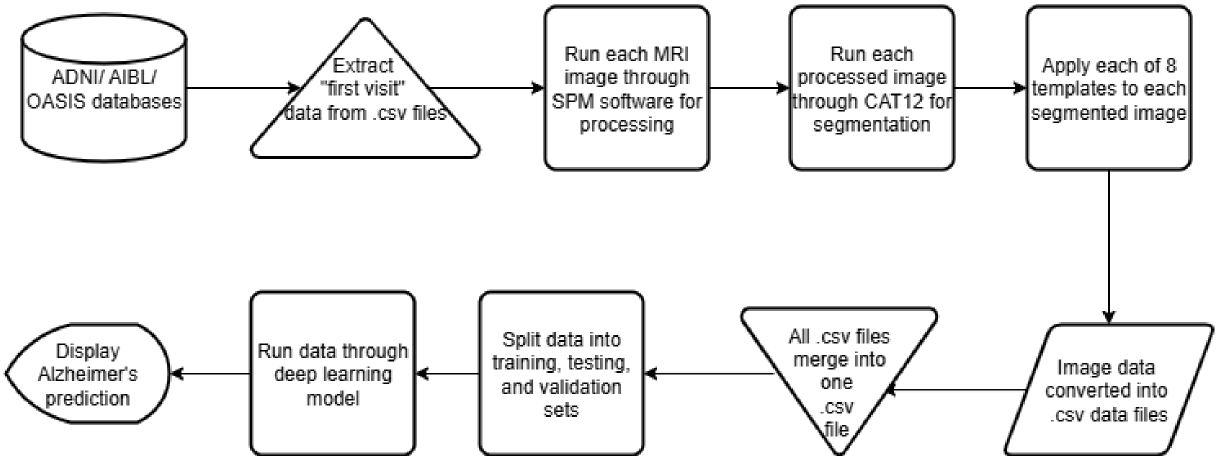
For our program, we are using the ADNI, AIBL, and OASIS datasets combined to create our initial dataset. We have chosen these databases because they contain a vast amount of MRI images taken at various levels of AD progression. Combining multiple databases also ensures the data we are working with is more diversified in random, to help ensure that the model is functional for any MRI image.

We are also utilizing SPM and CAT12 with MATLAB R2023b for preprocessing and segmentation of the original datasets. These programs are important because the versions of SPM and CAT12 we are using can only be run with MATLAB due to its mathematical capabilities. We are using SPM to take each MRI image and make a digital map that can be viewed and analyzed by the computer. We then pass the map made by SPM to CAT12 for segmentation, so that all parts of the brain can be seen and processed. The segmented images can then be passed back to SPM to be normalized based on different templates that focus on different regions of interest. A combined file of all regions of interest is what we are using t craft our final dataset, so it is important that it is built properly with the programs listed.

We will also be using the Python language to create our deep learning model because it is uniquely equipped with specialized libraries focused on neural networks and creating deep learning model. Because most of our data is stored on a Linux lab computer, it will also be easier to migrate our Python scripts across operating systems then it would be for other programming languages.

Currently, our model will only be built around initial MRI scans from patients' first visits to the doctor. To enhance the model in the future, we would want it to train on every MRI image across multiple scans over time so that the model can better predict AD progression as well. We also want to train the model not just on recognizing AD, but also on recognizing other calcification on the brain that could be prerequisites to the disease.

4. System Architecture


ADNI/ AIBL/
OASIS databases
Extract
"first visit"
data from .csv files
Run each MRI
image through
SPM software for
processing
Run each
processed image
through CAT12 for
segmentation
Apply each of 8
templates to each
segmented image
Image data
converted into
.csv data files
All .csv files
merge into
one
.csv
file
Split data into
training, testing,
and validation
sets
Run data through
deep learning
model
Display
Alzheimer's
prediction