A Tutorial on Neural Imaging Data sets from ADNI: Current Practices and Recommended Enhancements

Haymanot Gebre-Amlak*,   and   Narisa Nan Chu[†]

Department of Computer Science and Electrical Engineering
University of Missouri-Kansas City, Kansas City, MO, USA*
CWLab International, Thousand Oaks, CA, USA[†]
Email: *{hhgc77}@umkc.edu
[†]{narisa.chu}@ieee.org

Abstract—**Using the image data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) website can be overwhelming for many researchers. The sheer volume and size of image data (3 terabytes) can be intimidating in addition to the methods and techniques involved which are inherently complicated and are rapidly evolving. Alternatives approaches are not available and many researchers have depended on ADNI data to perform evaluations for their fields of interest. The learning curve required to use ADNI is not trivial. This tutorial will provide supplement for the complexity of the subject matter and non-trivial user interface in times of insufficient software extensions/technological support. Moreover, research involving the sought-after AI models, has shown performance un-reproducible, with a shortage of inter-disciplinary subject experts. This tutorial aims to bridge the gap between the researchers and the data collection. It also recommends areas of enhancements to make the ADNI data-set more user-friendly.**

I. Introduction

The Signal Processing in the Context of Neurotechnologies (SPCN) on cognitive sciences, sensory physiology, ophthalmological and neurological clinical investigation conference in 2020 calls out the ADNI datasets for the study of the Alzheimer's Disease (AD) [1] focused researches. This tutorial can aid researchers to tackle ADNI at the onset, and furthermore to examine brain aging process incorporating state-of-the-art modeling tools, e.g., neural networking and deep learning [2]. The Laboratory of Neuro Imaging (LONI) Image and Data Archive (IDA) is a secure system for the archival of neuroimaging data that protects subject confidentiality by restricting access to unauthorized users. IDA provides an interactive and secure environment for storing, searching, updating, sorting, accessing, tracking and manipulating neuroimaging and relevant clinical data [3]. One of the prominent studies which is part of IDA is Alzheimer's Disease Neuroimaging Initiative (ADNI) [4]. ADNI is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection, tracking, and drug development to treat Alzheimer's Disease (AD). ADNI data can be downloaded from the website in the form of images as well as the study information for those images. The whole data can be divided into 4 groups: ADNI1 (2004-10), ADNIGO (2009-11), ADNl2 (2011-16) and ADNl3 (2017-22).

The remainder of the tutorial is organized as follows. Section II provides information for accessing the data with required permissions. Section III describes navigating the webpage to find the data of interest and also presents data analysis and data cleaning process. Section IV shares the data storage options and memory requirements. Software tools to build the model are listed in Section V. We conclude the paper with topic syllabus in Section VI.

II. Access Request

The access request requires the researcher to fill out a form on the website [4] and the approval process takes approximately 5 business days. After you have been approved to access ADNI data, you can login to search and download a data from the repository.

III. Data

The study data tab which is part of the download tab provides an option to download information about the subjects, diagnosis, clinical study, psychological tests, and others, in comma separated value (csv) format. This is useful while analyzing images and looking up for more details regarding those images.

IV. Storage

The image download section enables the user to search images such as MRI, (Positron Emission Tomography) PET, DTI etc. using the search criteria. The results are generated in the Advanced Search tab with the option to download images one at a time or all together. Advanced downloader allows the user to compress and download many images together in a zipped file format. The images downloaded are in NIfTI, MINC, ANALYZE, DCM and other formats.

The downloading of all the images can take a long time (e.g., in one instance downloading a file of size 40 GB took more than a week at a university because of website security issues) with a total size of images reaching 3 terabytes.

Moreover, it is not useful to download all the images together since they are in different formats, with incomparable quality and modality. Therefore, it is much better to shortlist and download a subset of images as per the research requirements. Also, limited number of images would put less burden on processing capabilities to speed up the computations.

V. AVAILABLE TOOLS

There are many software tools for analyzing the image data on ADNI website. Some of the prominent ones are as follows: (i) FreeSurfer is an open source software package for the analysis and visualization of structural and functional neuroimaging data from cross-sectional or longitudinal studies [5]. It runs on Mac OS and Linux [6] (ii) Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL), Analysis Group [5]. Oxford University has a library of analysis tools for fMRI, MRI and Diffusion Tensor Imaging (DTI) brain imaging data. It runs on Apple and PCs (both Linux, and Windows via a Virtual Machine) [7], (iii) Adult brain atlases and image registration tools can be used from [8] for non-commercial research. (iv) Structures can be segmented from 3D images and information can be extracted using the software tools ITK-SNAP and ANTs from the Penn Image Computing and Science Lab [9]. (v) Feature extraction and implementation of 3D separable convolutions using TensorFlow and Keras libraries can be done through [10] and hippocampal segmentation using Convolution Neural Network (CNN) [11] can be done through github [12]. Figure 1 shows a typical CNN network. (vi) Supercomputing environment can also be
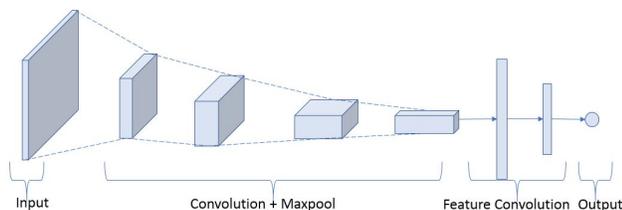

Fig. 1. Typical Convolution Neural Network

used by automatically parallelizing data-independent programs in a given analysis 8 through Pipeline Software provided by LION which requires an application to become a member and collaborator.

VI. SYLLABUS

A. Part I

ADNI Data sets, study data and image collections will be covered in first part of the tutorial.

B. Part II

Supporting tools including methods software application, models [13], case studies [14] and suggestions will be addressed in the second part of the tutorial.

ACKNOWLEDGEMENT

REFERENCES

[1] S. R. Sabat, *The experience of Alzheimer's disease : life through a tangled veil*. Oxford ;: Blackwell.

[2] D. P. Subha, P. K. Joseph, R. Acharya U, and C. M. Lim, "Eeg signal analysis: A survey," *Journal of Medical Systems*, 2010.

[3] "usc-data archive," https://resource.loni.usc.edu/resources/data-archive//, 2020, accessed: 2020-08.

[4] "Alzheimer's Disease Neuroimaging Initiative," http://adni.loni.usc.edu/about/, 2020, accessed: 2020-08.

[5] "FreeSurfer," https://surfer.nmr.mgh.harvard.edu/, 2020, accessed: 2020-08.

[6] "What is Linux?" https://www.redhat.com/en/topics/linux/what-is-linux, 2020, accessed: 2020-08.

[7] "fmrib," https://fsl.fmrib.ox.ac.uk/fsl/fslwiki, 2020, accessed: 2020-08.

[8] "brain-development," http://brain-development.org/, 2020, accessed: 2020-08.

[9] "upenn," http://picsl.upenn.edu/software/, 2020, accessed: 2020-08.

[10] "github," https://github.com/simeon-spasov/MCI, 2020, accessed: 2020-08.

[11] "Convolutional Neural Networks," https://cs231n.github.io/convolutional-networks/, 2020, accessed: 2020-08.

[12] "github," https://github.com/bthyreau/hippodeep, 2020, accessed: 2020-08.

[13] H. Lan, , A. W. Toga, and F. Sepehrband, "Sc-gan: 3d self-attention conditional gan with spectral normalization for multi-modal neuroimaging synthesis," *bioRxiv*, 2020. [Online]. Available: https://www.biorxiv.org/content/early/2020/06/11/2020.06.09.143297

[14] M. Grassi, N. Rouleaux, D. Caldirola, D. Loewenstein, K. Schruers, G. Perna, M. Dumontier, and A. D. N. I. , "A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to alzheimer's disease using socio-demographic characteristics, clinical information, and neuropsychological measures," *Frontiers in Neurology*, vol. 10, p. 756, 2019. [Online]. Available: https://www.frontiersin.org/article/10.3389/fneur.2019.00756