HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 3 (2023)

# 3-D Image Based Deep Learning for Dementia Diagnosis

George Liu[1] and Guillermo Goldsztein[#]

[1]Johns Creek High School
[#]Advisor

## ABSTRACT

Dementia is a neurodegenerative disorder that greatly affects memory, thinking, and reasoning, impacting millions of people worldwide. Although dementia diagnosis is challenging and time-consuming, recent studies have shown promising results in using deep learning for dementia diagnosis by analyzing MRI scans. However, these studies are limited by access to data and the depth of analysis. In this study, we developed a deep learning model that utilizes the T1-weighted MRI scans from the Open Access Series of Imaging Studies (OASIS-3) dataset, which contains data from 1378 patients with varying degrees of cognitive decline. The model classified MRI scans into two categories, negative or positive diagnosis of dementia, based on the patients' clinical dementia rating (CDR). A 3-D convolutional neural network (CNN) was constructed with the TensorFlow framework and achieved an accuracy of 76.25%. This study demonstrates the potential that deep learning models have in the future of dementia diagnosis.

## Introduction

Dementia is the general term that describes a host of neurodegenerative disorders that greatly affect memory, thinking, and reasoning. It currently impacts more than 55 million people worldwide, with almost 10 million new cases every year (World Health Organization, 2023). Many forms of dementia, like Alzheimer's disease and Vascular Dementia, are progressive. The severity of these conditions varies greatly, from mild symptoms in the earlier stages to serious complications in later stages. Although there is currently no cure for dementia (NHS, 2021), starting treatment in the early stages can slow the onset of dementia and greatly improve a patient's quality of life. Therefore, early diagnosis is of utmost importance.

Diagnosing dementia can be extremely challenging, requiring that a clinician recognizes certain patterns and signs. No single test can be used; clinicians must run numerous different tests (cognitive tests, neurological evaluations, psychological evaluations, lab tests) (Mayo Clinic, 2022), making for an extremely time consuming and costly process. Recently, clinicians have also been able to use brain imaging for diagnosis, magnetic resonance imaging (MRI) and computed tomography (CT) being the most popular. Although both are widely used, MRI scans are better at identifying certain conditions like brain atrophy and brain damage (Stanford Medicine, 2017). However, despite the plethora of available methods, diagnosis still consumes an incredible amount of time and resources. The many intricacies of brain activity require the focused collaboration of experts to decipher. Machine learning presents an opportunity to make improvements to this process and unlocks an alternate path forward.

Machine learning allows for the classification of images based on a trained model. Deep learning simulates the behavior of a human brain, allowing for multiple layers of processing that can make more complex conclusions. This more intricate processing method provides the level of detail necessary for use in the medical field. Several studies using deep learning for dementia diagnosis have already been conducted (Buvari & Pettersson, 2020; Falahati et al., 2014; Gill et al., 2020; Yagis et al., 2020). These studies have shown promising results in detecting cognitive impairment with the classification of MRI scans. However, these studies are limited by access to data and depth of analysis. In general, access to MRI data is limited; the available data sets are relatively small which could have a

negative impact on the training of the learning model. 3-D images are also difficult to work with, resulting in researchers (Buvari & Pettersson, 2020) opting for simpler 2-D analysis. However, this could limit the learning model's ability to identify more obscure details and patterns. The objective of this study is to develop a learning model capable of accurately diagnosing Dementia that fully utilizes the small amount of data available.

## Method

### Dataset

The Open Access Series of Imaging Studies (OASIS) project is made for the purpose of making neuroimaging data freely available to the scientific community. The data that OASIS provides spans a broad demographic containing thousands of subjects. The Oasis project has 5 compiled datasets: OASIS-1, OASIS-2, OASIS-3, OASIS-3_TAU, and OASIS-4. The OASIS-3 data set (LaMontagne et al., 2019) was chosen for this project.

The OASIS-3 dataset is a Longitudinal Multimodal Neuroimaging, Clinical, and Cognitive dataset that covers age related deterioration and Alzheimer's disease. The Dataset was retrospectively compiled from data through the Washington University in St. Louis Knight ADRC, spanning the course of 30 years. The data set contains the information of 1378 patients, consisting of 755 cognitively normal subjects and 622 subjects showing various degrees of cognitive decline. The patients span the ages of 42-95 years of age. The data is anonymous and the dates of which the patients where scanned has been normalized to reflect the days from entry into the study. The data in the OASIS-3 dataset consist of both MRI and PET data. The MRI data consists of data from 2842 MR sessions which include T1-weighted, T2-weighted, FLAIR, ASL, SWI, time of flight, resting-state BOLD, and DTI.

### Retrieving Data

The MRI scans from the OASIS-3 Data set were first selected by scan type. Of the scan types available in the data set, the T1-weighted scan type was selected as it is the best suited for the identification of atrophy, a clear indicator of neurodegenerative dementia. This is illustrated in Figure 1. The strength of the scans was normalized by selecting only 3.0T MRI scans.

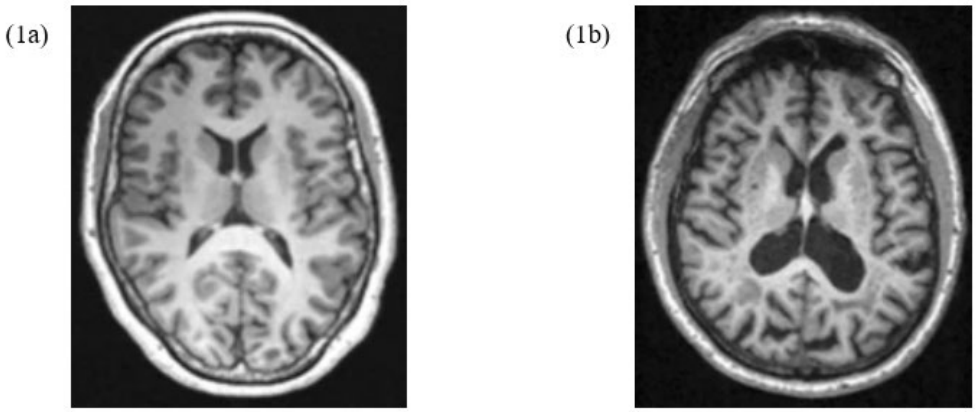


**Figure 1.** Atrophy in T1w Brain Scans of Demented Subjects (Chandra, A., Dervenoulas, G., Politis, M., 2018). Figure 1a shows the T1w Scan of a Normal Brain while Figure 1b shows the T1w scan of a brain with dementia. There are clear signs of regional atrophy in figure 1b.

The data was then sorted by CDR. CDR, Clinical Dementia Rating, is a global rating method used to determine the severity of a patient's cognitive impairment. The CDR is based on a scale of 0–3: no dementia (CDR = 0), questionable dementia (CDR = 0.5), MCI (CDR = 1), moderate cognitive impairment (CDR = 2), and severe cognitive impairment (CDR = 3) (Khan, 2016). In this study, patients with CDR = 0 are considered a negative diagnosis and patients with a CDR > 0 are considered a positive diagnosis. These two diagnoses are used for binary classification.

## Preprocessing

The 1882 T1-weighted MRI sessions consisted of 2684 distinct scans. These scans consisted of 2022 negative diagnoses and 662 positive diagnoses. To create a balanced data set, 662 of the 2022 negative diagnoses were randomly sampled, resulting in a final data set of 1324 images.

The images first underwent a skull stripping brain extraction using the HD-BET CPU (Isensee et al., 2019) implementation. This is a necessary preliminary step that is meant to isolate brain tissue from non-brain tissue. Then, the images are rotated by 90 degrees to ensure that their orientations are fixed. Finally, voxel intensities of the images are normalized with a histogram-based normalization technique (Sun et al., 2015).

After this process, the processed images are partitioned into two parts: 80% in the training set and 20% in the validation set.

## Model Selection

A 3-D Convolutional Neural Network (CNN) is the chosen model for this project. The model (Zunair et al., 2020) starts with four modules consisting of a 3D Convolutional Layer, a Max Pooling Layer, and a Batch Normalization Layer. These modules are followed by a global average pooling layer. This leads to a fully connected layer (512 neurons, ReLU activation function), a dropout layer (0.4 dropout rate), and a final fully connected layer with the sigmoid activation function. Loss is measured by the binary cross-entropy error.

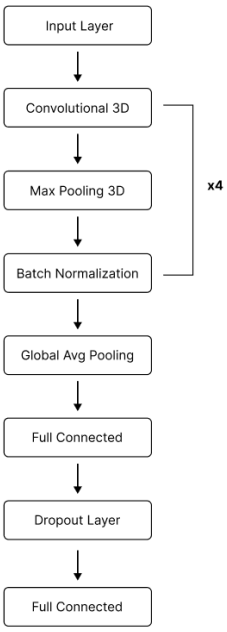


**Figure 2.** CNN Model used with four modules of a 3D Convolutional Layer, a Max Pooling Layer, and a Batch Normalization Layer

### Implementation

The project was implemented in Python 3.11.0. The diagnosis data provided alongside the OASIS-3 Dataset (LaMontagne et al., 2019) was read and processed with the Pandas and NumPy libraries, two of the most used libraries in machine learning. Pandas is a library for data analysis and manipulation and NumPy is a library that provides a host of mathematical functionality. Preprocessing was done using the Nibabel, ANTs, and FSL libraries, all libraries used for medical image processing. The model was implemented with the TensorFlow Library and Keras. TensorFlow is a machine learning library developed by Google, and Keras is a high-level deep learning library that runs on top of it. The model was trained with the dataset of images, using the diagnosis information as labels. A patience level of 10 was used for validation accuracy to terminate the training.

### Hyperparameter Optimization

The hyperparameters used for classification of dementia included training for 100 epochs with early stopping if there was no improvement in validation loss for 10 epochs, dropout of 0.3, batch size of 2, an Adam optimizer with learning rate of 1e-4. We used image rotations with angles of {−15, −10, −5, 5, 10, 15}, randomly selecting angles to augment the images in the data set.

### Execution

The code was executed in the Visual Studio Code IDE on a Windows machine with 4 cores and 16 gigabytes of ram.

## Results

The accuracy and loss (binary cross-entropy loss) on the training and validation sets were tracked during the training. All plots displayed are generated by matplotlib, a library for graphical visualization. The model was trained for 100 epochs. Each epoch consisted of 80 steps and ranged from 1500 seconds to 1800 seconds in length. The accuracy of the training and validation sets are plotted in figure 3a. The loss of the training and validation sets are plotted in figure 3b. After 100 epochs, a training accuracy of 76.25% and a validation accuracy of 70.80% were reached.

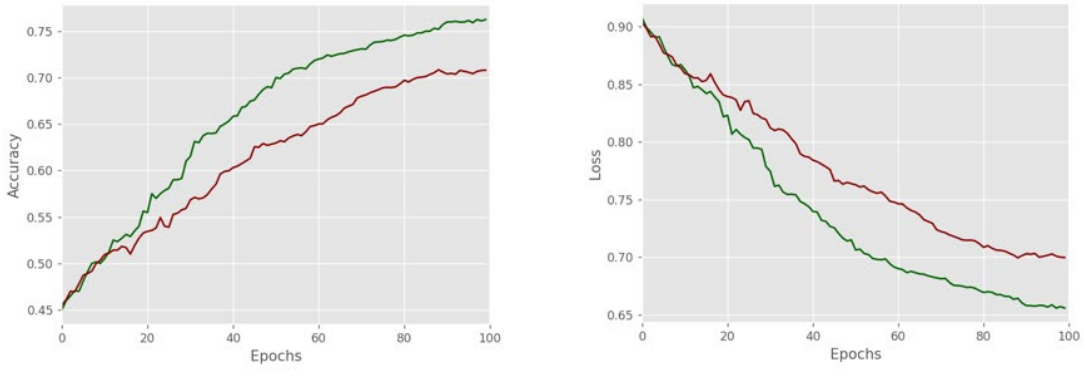


**Figure 3.** Accuracy and Loss Plots from training. Figure 3a shows the accuracy of the model vs epochs while figure 3b shows the loss (binary cross entropy) vs epochs.

# Discussion

## General Discussion

This study has resulted in the development of a complex deep learning model that diagnoses dementia with an accuracy of 76.25%. When we examine the accuracy and loss plots, we can see that accuracy is generally increasing and that loss is generally decreasing every epoch for both training and validation data. This is the expected trend of a machine learning network. We can also see that the curve starts to flatten in later epochs, meaning that the model is getting close to its peak performance and accuracy. By stopping the training after the model seems to have peaked in performance, we reduced the risk of overfitting.

This study places a greater emphasis on complex 3-D analysis than the image classifier that Bovari and Petterson (2020) developed. They chose to analyze a single slice of the image, resulting in an image classifier with an accuracy of 71.43% on the same OASIS-3 Dataset. This notable difference in accuracy suggests that increasing the complexity of analysis is a crucial part of increasing model diagnosis accuracy.

This is notion is further supported when observing the study conducted by Islam and Zhang (2018). They also used the OASIS dataset to train their model, employing a similar convolutional neural network and achieving an accuracy of 73.5% with their image classifier. The main reason for this difference in accuracy most likely comes from the difference in the construction of our models. The model constructed by Islam and Zhang had 2D architecture, relying on 2-D convolution to process images. In contract, the model constructed in this study uses 3-D convolution. By keeping the extra layer of dimensionality, the data from 3-D MRI scans can be analyzed more in depth. Although this added complexity increases training time, it results in better accuracy for the model's diagnoses as the machine is exposed to more information.

The study conducted by Yagis et al. took a similar approach with 3-D convolution on the OASIS dataset. However, their model achieved a lower accuracy of 69.9% ± 0.06 for the classification of AD subjects. The most likely cause for this lower accuracy is their smaller dataset. Yagis et al. chose to choose a sample of 100 healthy controls and 100 patients with AD (patients with CDR > 0 were considered to have AD, a labelling method similar to the one used in this study). This smaller data set resulted in a lower model accuracy, even with a similar approach. This suggests that their model had insufficient data to learn from, indicating that larger datasets and more images improve model performance.

## Potential Sources of Error

All patients that had CDR values > 0 were considered to have a positive diagnosis for dementia. This labeling method could lead to errors with patients that have CDR values of 0.5. A CDR score of 0.5 may not be a completely accurate indicator of cognitive impairment in a patient (Wada-Isoe, 2019). This could lead to false labelling of a patient, assigning them a positive diagnosis when they have no clear signs of impairment. This could lead to issues in the model training: the model could be exposed to a false positive or false negative which can decrease the accuracy of the model. Errors might also stem from the OASIS-3 Dataset itself, as Islam and Zhang (2018) reported that several NIfTI files from the dataset had slices swapped along the x and z axes. Since we essentially assume that all slices used are on the same axis, these swapped files can negatively impact the model's performance. Slices on different axes present in the dataset may reduce the accuracy of the model.

### Future Studies

There are several ways to improve the accuracy of the model in future projects. These options include potential modifications of both the model and the dataset.

1. Compiling a larger data set is the clearest way to improve on the model's performance. The process of diagnosing dementia is extremely complex and filled with many intricacies. Having a larger data set with more examples for the computer to learn from would likely improve model performance.
2. Although several preprocessing steps were taken to increase the quality of the input data, there are more steps than can be taken to further increase the quality. Steps like Bias field correction and image registration would theoretically yield cleaner training data and a higher model accuracy.

### Conclusion

Using complex analysis with 3-D convolution, a model has been developed that demonstrates machine learning's applicability in Dementia Diagnosis. Although certain aspects can be improved in future studies, this study yields promising results with a model that diagnoses dementia with 76.25% accuracy. Results from this study represent a step towards the full automation of dementia diagnosis, a feat that would drastically increase the efficiency and speed of the diagnosis process.

## Acknowledgements

Data was provided by OASIS-3: Longitudinal Multimodal Neuroimaging: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P30 AG066444, P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, R01 EB009352. AV-45 doses were provided by Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly.

## References

Buvari, S., & Pettersson, K. (2020, June 23). A Comparison on Image, Numerical and Hybrid based Deep Learning for Computer-aided AD Diagnostics. Retrieved March 15, 2023, from https://www.diva-portal.org/smash/get/diva2:1463276/FULLTEXT01.pdf.

Chandra, A., Dervenoulas, G., & Politis, M. (2018, August 17). *Magnetic Resonance Imaging in alzheimer's disease and mild cognitive impairment - journal of neurology*. SpringerLink. Retrieved March 23, 2023, from https://link.springer.com/article/10.1007/s00415-018-9016-3.

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *Journal of Alzheimer's disease: JAD*, *41*(3), 685–708. https://doi.org/10.3233/JAD-131928

Gill, S., Mouches, P., Hu, S., Rajashekar, D., MacMaster, F. P., Smith, E. E., Forkert, N. D., Ismail, Z., & Alzheimer's Disease Neuroimaging Initiative (2020). Using Machine Learning to Predict Dementia from Neuropsychiatric Symptom and Neuroimaging Data. *Journal of Alzheimer's disease : JAD*, *75*(1), 277–288. https://doi.org/10.3233/JAD-191169

Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., Wick, A., Schlemmer, H. P., Heiland, S., Wick, W., Bendszus, M., Maier-Hein, K. H., & Kickingereder, P. (2019). Automated brain extraction of multisequence MRI using artificial neural networks. *Human brain mapping*, *40*(17), 4952–4964. https://doi.org/10.1002/hbm.24750

Islam, J., & Zhang, Y. (2018). Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain informatics*, *5*(2), 2. https://doi.org/10.1186/s40708-018-0080-3

Khan, T. K. (2016). *Clinical dementia rating*. Clinical Dementia Rating - an overview | ScienceDirect Topics. Retrieved March 15, 2023, from https://www.sciencedirect.com/topics/medicine-and-dentistry/clinical-dementia-rating#:~:text=The%20CDR%20is%20based%20on,impairment%20(CDR%20%3D%203)

LaMontagne, P. J., Benzinger, T. L. S., Morris, J. C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A. G., Raichle, M. E., Cruchaga, C., & Marcus, D. (2019, January 1). *Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease*. medRxiv. Retrieved March 18, 2023, from https://www.medrxiv.org/content/10.1101/2019.12.13.19014902v1.

Lang, N. (2022, October 24). *Using convolutional neural network for Image Classification*. Medium. Retrieved March 15, 2023, from https://towardsdatascience.com/using-convolutional-neural-network-for-image-classification-5997bfd0ede4#:~:text=The%20Convolutional%20Neural%20Network%20(CNN,suited%20for%20this%20use%20case.

Mayo Clinic. (2022, October 12). *Dementia*. Mayo Clinic. Retrieved March 17, 2023, from https://www.mayoclinic.org/diseases-conditions/dementia/diagnosis-treatment/drc-20352019.

NHS. (2021, April 8). *Is there a cure for dementia?* NHS choices. Retrieved March 16, 2023, from https://www.nhs.uk/conditions/dementia/cure/#:~:text=There%20is%20currently%20no%20%22cure,and%20dementia%20with%20Lewy%20bodies.

Stanford Medicine. (2017, September 11). *Brain Scans*. Stanford Health Care (SHC) - Stanford Medical Center. Retrieved March 23, 2023, from https://stanfordhealthcare.org/medical-conditions/brain-and-nerves/dementia/diagnosis/brain-scans.html.

Sun, X., Shi, L., Luo, Y., Yang, W., Li, H., Liang, P., Li, K., Mok, V. C. T., Chu, W. C. W., & Wang, D. (2015, July 28). *Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions - biomedical engineering online*. BioMed Central. Retrieved April 3, 2023, from https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-015-0064-y.

Wada-Isoe, K., Kikuchi, T., Umeda-Kameyama, Y., Mori, T., Akishita, M., Nakamura, Y., & ABC Dementia Scale Research Group (2019). Global Clinical Dementia Rating Score of 0.5 May Not Be an Accurate Criterion to Identify Individuals with Mild Cognitive Impairment. *Journal of Alzheimer's disease reports*, *3*(1), 233–239. https://doi.org/10.3233/ADR-190126

World Health Organization. (2023, March 15). *Dementia*. World Health Organization. Retrieved March 21, 2023, from https://www.who.int/news-room/fact-sheets/detail/dementia.

Yagis, E., Citi, L., Diciotti, S., Merzi, C., Workalemahu Atnafu, S., & Garcia Seco De Herrera, A. (2020, July 1). *3D convolutional neural networks for diagnosis of alzheimer's disease via structural MRI*. Research Repository. Retrieved March 25, 2023, from https://repository.essex.ac.uk/27801/

Zunair, H., Rahman, A., Mohammed, N., & Cohen, J. P. (2020, July 26). *Uniformizing techniques to process CT scans with 3D CNNS for tuberculosis prediction*. arXiv.org. Retrieved March 22, 2023, from https://arxiv.org/abs/2007.13224.