Differentiating Alzheimer's Disease from Cognitively Normal Individuals Using Convolutional Neural Networks: A Reproducible Study

Elina Thibeau-Sutre[1], Camille Brianceau[2], ChatGPT[3], and Ninon Burgos[2]

[1] Department of Applied Mathematics, Technical Medical Centre, University of Twente, Enschede, The Netherlands
e.thibeau-sutre@utwente.nl
[2] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France
ninon.burgos@cnrs.fr
[3] OpenAI, Internet, World

Abstract. Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects millions of individuals worldwide. Early and accurate diagnosis is crucial for effective intervention and patient care. This study aims to develop a reproducible deep learning model based on convolutional neural networks (CNNs) to differentiate between AD patients and cognitively normal participants using brain MRI scans.
The chosen CNN model demonstrated a high level of accuracy in distinguishing AD patients from cognitively normal controls. On the test set, the model achieved an accuracy of 87%. The attribution maps associated to the trained network highlighted regions known to be affected by the disease (medial temporal regions).
This reproducible study demonstrates the potential of convolutional neural networks in effectively differentiating Alzheimer's disease patients from cognitively normal individuals based on brain MRI scans. The high balanced accuracy achieved by the model highlight its clinical relevance and potential as a valuable diagnostic tool. The open-source code used in this study is made available to facilitate further research and ensure transparency and reproducibility in the field of neuroimaging-based AD diagnosis.

Keywords: Alzheimer's disease · Deep Learning · Magnetic Resonance Imaging.

1 Introduction

Alzheimer's disease (AD) affects over 20 million people worldwide. Neuroimaging provides useful information to identify AD [1], such as the atrophy due to gray matter loss with anatomical magnetic resonance imaging (MRI). A major interest is then to analyze those markers to identify AD at an early stage. Machine

learning and deep learning methods have the potential to assist in identifying patients with AD by learning discriminative patterns from neuroimaging data [2].

As the most widely used architecture of deep learning, convolutional neural networks (CNN) have attracted huge attention thanks to their great success in image classification [3]. Contrary to conventional machine learning, deep learning allows the automatic abstraction of low-to-high level latent feature representations. Thus, one can hypothesize that deep learning depends less on image preprocessing and requires less prior on other complex procedures, such as feature selection, resulting in a more objective and less bias-prone process [4].

The purpose of this paper is to explain the results of a deep learning network trained to differentiate Alzheimer's disease patients from cognitively normal participants. The source code for the experiments and models described in this paper will be made available on GitHub and is attached to this submission during the review process.

2 Materials

We used T1w-MR images from the clinical routine of the MST hospital in Enschede [5][4]. The medical team identified which scans were associated to the AD label (210) and to the CN label (103).

3 Methods

3.1 Preprocessing of T1-weighted MRI

Images were processed with the N4 bias field correction to remove scanner imperfections or variations in tissue properties. The images were then non-linearly registered to the MNI template with the SPM toolbox. Finally the intensity were rescaled between the 5th and 95th percentage. This led to images of size $145\times145\times145$ with a voxel size of 1mm^3.

3.2 Deep learning network

Architecture The architecture of our network is shown in Figure 1. It includes 5 convolutional layers and 3 fully-connected layers. After each of these layers a ReLU activation layer is inserted, as well as a dropout layer with a rate of 0.7. The dimension reduction occurs because of the convolutional layers with a stride of 2.

[4] Data set available on request

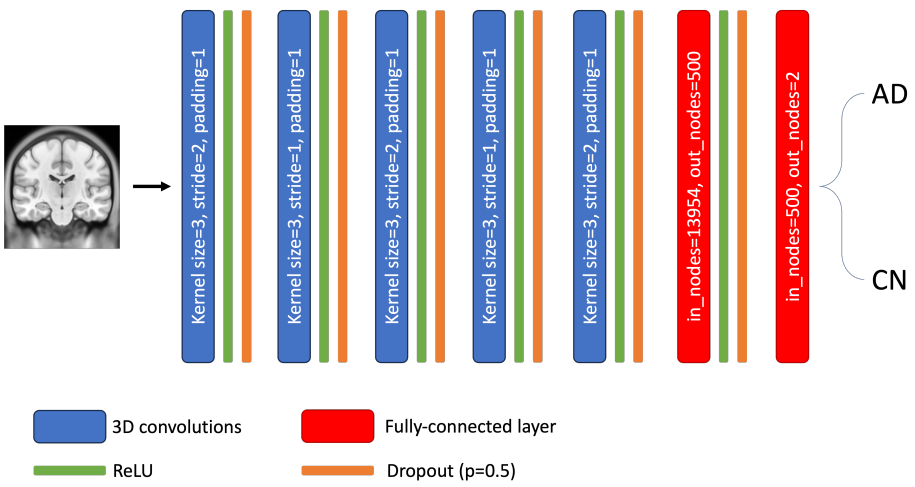


Fig. 1: Architecture of the deep learning network.

Training Training a deep learning network involves a process of optimizing its parameters to learn meaningful representations from the given data. At its core, this process hinges on the minimization of a defined loss function, in our case cross-entropy loss. During training, the network's weights and biases are iteratively updated using an optimization algorithm, Adam, which is a variant of the stochastic gradient descent, with a learning rate of 0.001.

The data is divided into batches of size 16. We used a learning rate scheduler to help control the step size during optimization. The training process continues for 25 epochs. The whole training process of this network required 18G of memory.

3.3 Machine learning method

For comparison purposes, classification was also performed with a linear SVM classifier. We chose the linear SVM as we previously showed that it obtained higher or at least comparable classification accuracy compared 2 to other conventional models (logistic regression and random forest) [6]. Moreover, given the very high-dimensionality of the input, a nonlinear SVM, e.g. with a radial basis function kernel, may not be advantageous since it would only transport the data into an even higher dimensional space. This method required less computational ressources to be trained (∼5G).

4 Results

We assessed which method was the best between machine and deep learning.

Accuracy For each method the accuracy of the classifier is computed for multiple runs (Figure 2). With a T-test we observed a significant difference between the performance of the deep learning and the machine learning method (p-value=0.043). The deep learning achieved a mean accuracy of 0.87 whereas the machine learning one achieved a mean accuracy of 0.85.

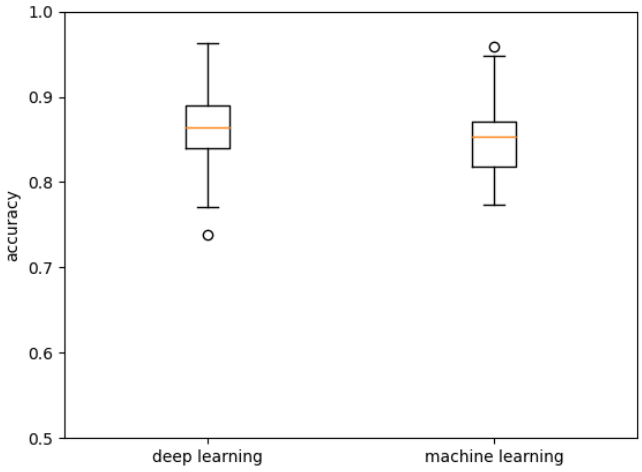
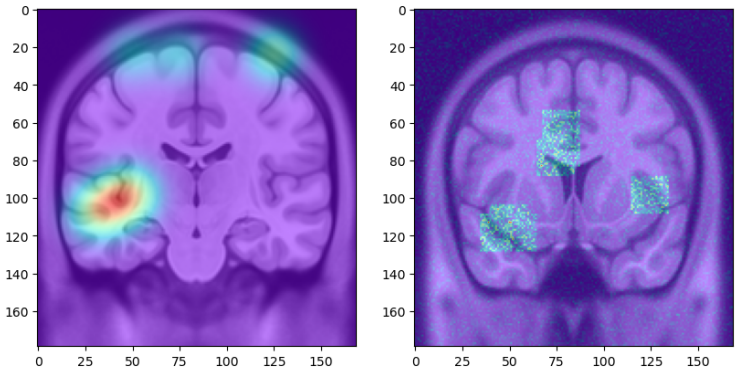


Fig. 2: Accuracy obtained over multiple runs for the deep learning and machine learning methods.

Explainability In our comprehensive study comparing the performance of deep and machine learning, we explored the generation of attribution maps as a critical component of our analysis (Figure 3). Attribution maps offer invaluable insights into the decision-making process of these machine learning models, shedding light on the regions of interest within the brain that contribute most significantly to their classifications.

Both maps highlighted specific regions of the brain, such as the hippocampus and certain cortical areas, known to be closely associated with AD pathology. These regions were illuminated as "hotspots" on the attribution maps, suggesting that the methods relied heavily on these areas to discriminate between AD and cognitively normal individuals. The main difference between the two methods could be found in regions such as the occipital lobe and cerebellum, which were more highlighted in the deep learning than the machine learning map compared to the clinically coherent regions.



(a) Deep learning (b) Machine learning

Fig. 3: Attribution maps obtained with each method.

5 Conclusion

In this article we showed that we could accurately differentiate demented patients from cognitively normal participants from their brain T1w-MRI. We found that our result depends on the chosen architecture, and especially the number of layers, but that the tendency was different for convolutional layers (a minimum of 4 were required) and the FC layers (the performance was hurt by adding more layers). The regions found in the attribution map corresponds to the ones clinically known to be affected in Alzheimer's disease, thus indicating that this network has the potential to be used in a clinical routine as a support tool for radiologists. Future work will investigate the robustness of the attribution maps towards the network hyperparameters.

References

1. Ewers, M., Sperling, R.A., Klunk, W.E., Weiner, M.W., Hampel, H.: Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. Trends in Neurosciences **34**(8) (2011) 430–442
2. Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., Dormont, D., Durrleman, S., Burgos, N., Colliot, O.: Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. Medical Image Analysis **63** (2020) 101694
3. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet Classification with Deep Convolutional Neural Networks. In: Advances in neural information processing systems. Volume 25. (2012) 1097–1105
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553) (2015) 436–444
5. Jedusor, T.E., Dumbledore, A.: T1-MR images of neurodegenerative disease in the clinical routine. Magically reproducible clinical studies **1**(1) (2022) 95–113

6. Samper-González, J., Burgos, N., Bottani, S., Fontanella, S., Lu, P., Marcoux, A., Routier, A., Guillon, J., Bacci, M., Wen, J., Bertrand, A., Bertin, H., Habert, M.O., Durrleman, S., Evgeniou, T., Colliot, O.: Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. NeuroImage **183** (2018) 504–521