Alzheimer detection using Deep Learning

Roxana Zachman-Țîșcă
Babeș-Bolyai University
roxana.zachman@stud.ubbcluj.ro

Abstract

Detecting Alzheimer's disease as early as possible is one of the most important challenges in today's medicine. Although no disease-modifying agents capable of reversing the initial pathological changes are currently available, it may be possible to prevent or delay the development of dementia in a proportion of the population by modifying exposure to common risk factors [11]. For any such treatment to be effective, it is important that the patient is not diagnosed too late. In this paper we study the effectiveness of scanning brain MRI scans using deep learning techniques.

1. Introduction

Deep learning has proven extremely useful in a plethora of areas, including healthcare [14]. Deep learning has been successfully used in medical imaging, healthcare data analytics, drug discovery, genomics analysis and many more. It has been used in the context of the COVID-19 pandemic to analyze CT scans, predict intensive care unit admission, and to estimate the need for mechanical ventilation. It has even been used in protein folding [4]. It has now become quite clear that deep learning is a great aid to medical professionals. We will attempt to use the techniques of deep learning (more specifically CNNs - Convolutional Neural Networks) in the context Alzheimer detection from brain MRI scans.

2. Related work

A systematic review of machine learning classification methods for assisted diagnosis is provided in [9] by Pellegrini et al, including literature works published during the years 2006–2016. Most works use traditional image processing methods for image pre-processing and feature extraction and traditional machine learning classification methods. Only 2 out of the 111 relevant studies considered were making use of a deep learning-based approach. Since 2016, as mentioned in [9], there has been a significant increase on the number of publications (with a big presence in conferences) using deep learning-based methods for the automatic classification of Alzheimer's disease on MRI images. In this sense, Vieira et al [13] provide the first review focused on deep learning (DL) methods, however, not only for AD classification but also for other brain-based disorders.

In general, there are two types of DL approaches: to combine Convolutional Neural Network (CNN) with traditional image processing and machine learning methods (for feature selection and/or classification) or to develop full end-to-end CNN DL solutions. With respect to the first method, the one which combines approaches, we can mention works like the ones by Khagi et al [5] or by Nawaz et al [8], where the reported accuracies are close to 99%. While impressive, these works provide few details about the selection of cases in the dataset, meaning that the reported results can be biased and unrealistic Also relevant in this category is the work done by Puente et al [10], where an average accuracy of 86.81% was reported. This is also interesting and reproducible, although no balanced accuracy (BAC) metric is provided, which is a more realistic metric to evaluate the accuracy of the approach in the presence of the high data imbalance of this dataset. Other metrics provided are accuracy, precision, recall and specificity, reporting 0% precision and recall for mild dementia and moderate AD (Alzheimer's Disease) classes due to the small number of cases in the dataset.

If we look at the other approach, the one with full DL solutions, we can find methods such as the one proposed by Islam et al [3]. Here, they proposed a 2D architecture which consists of an ensemble of three homogeneous and slightly different models containing convolutional, batch normalization, Rectified Linear Unit (ReLU) and pooling operations. For data augmentation, they proposed a cropping strategy, where three crops of size 112×112 pixels are extracted from each database sample, one from each image plane. They made use of the full OASIS-1 dataset, dedicating 70% for training, 10% for validation and 20% for testing. The model was trained independently and the 'softmax' classification layer with cross entropy was added to solve a four class classification problem considering the CDR information on the dataset. A cost-sensitive training strategy is used for dealing with data imbalance, using a cost matrix to modify the output of the last layer of the networks for giving more importance to the underrepresented classes, assigning weights dependable to the number of samples of each class. Besides, models were optimized with the Stochastic Gradient Descent (SGD). algorithm and early stopping regularization. Then, individual models' answers are ensembled using a majority voting strategy, where the class with the majority of the votes is assigned as the output answer. The mean accuracy of the proposed architecture is 93.18%, with 94% precision (67% for mild dementia and 50% for moderate dementia) and 93% mean recall (33% for very-mild dementia and 50% for moderate dementia).

3. Our approach

Our approach is heavily inspired from the principles of AlexNet [7] and VGG [12], with specific changes to fit our exact data.

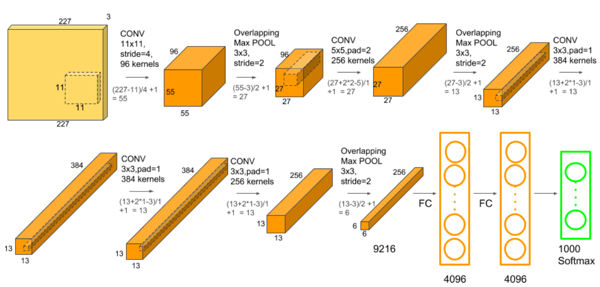

Figure 1. AlexNet architecture

Similarly to AlexNet and VGG, we followed the pattern of increasing the number of feature maps while decreasing the resolution of these feature maps as the depth increases. We do this by applying convolutional layers (with increasingly many filters from layer to layer), followed by a maxpool layer which decreases the resolution of the feature maps. After each convolutional layer, we apply BatchNorm [2] followed by the ReLU activation function. At the end of the architecture, we flatten the feature maps (which by this point have dimension 1024x6x5 in $channels$x$height$x$width$) into tensors of size 30720. This is followed by 2 fully connected layers of 2048 and 1024 neurons respectively, and then another fully connected layer mapping to our 4 classes (NonDemented, VeryMildDemented, MildDemented, ModerateDemented).

4. Dataset

The dataset we used is the Alzheimer dataset from Kaggle. It consists of 6400 brain MRI scans, each belonging to one of the classes: NonDemented, VeryMildDemented, MildDemented, ModerateDemented. The dataset was hand collected from various websites with each and every labels verified. All the images have the same shape (208x176 in height x width), which determined us to not use any data augmentation such as cropping or rescaling.

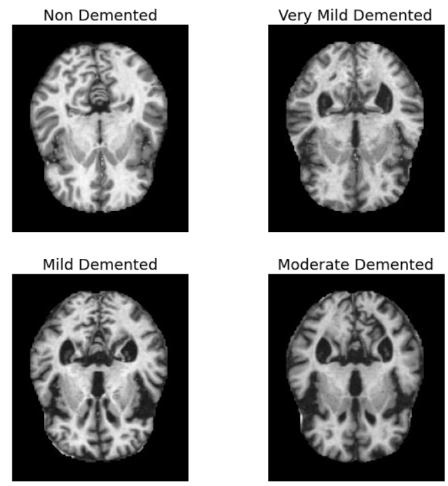

Figure 2. Different type of patients

While the data is of high quality, we can still notice some problems with it. One thing is that it is not that much data to begin with (only 6.4k images), and the other and more important thing is that the dataset is not balanced across classes. The dataset consists of 3.2k "NonDemented" samples, 2.2k "VeryMildDemented" samples, 900 "MildDemented" samples, and 64 "ModerateDemented" samples. We expect that an improvement in the dataset would greatly improve performance.

5. Results and future improvements

We trained the model from scratch, without using pretraining for around 75 epochs with early stopping. As a loss function we used Cross Entropy, and Adam [6] as the optimization technique. The resulting multi class accuracy we reached is around 75%. While this number is lower than the methods previously discussed, there are many improvements to be made. Some of which are:

- Using a pre-trained model instead of training from scratch
- Using a deeper model with residual connections [1]
- Using more augmentation techniques: random cropping, resizing, flipping, lighting changes, etc
- Improving the dataset: more samples and a more balanced dataset
- Using test-time augmentation like some of the models discussed used

Even though the model is simple and doesn't have the best performance, the results show that it is clearly learning something and not just randomly guessing the responses and thus is provides a good baseline to start adding improvements to. With the above improvements we expect the performance to improve significantly.

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 2

[2] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015. 2

[3] Jyoti Islam and Yanqing Zhang. Brain mri analysis for alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain informatics*, 5(2):1–14, 2018. 1

[4] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021. 1

[5] Bijen Khagi, Goo-Rak Kwon, and Ramesh Lama. Comparative analysis of alzheimer's disease classification by cdr level using cnn, feature selection, and machine-learning techniques. *International Journal of Imaging Systems and Technology*, 29(3):297–310, 2019. 1

[6] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 2

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017. 2

[8] Hina Nawaz, Muazzam Maqsood, Sitara Afzal, Farhan Aadil, Irfan Mehmood, and Seungmin Rho. A deep feature-based real-time system for alzheimer disease stage detection. *Multimedia Tools and Applications*, 80(28):35789–35807, 2021. 1

[9] Enrico Pellegrini, Lucia Ballerini, Maria del C Valdes Hernandez, Francesca M Chappell, Victor González-Castro, Devasuda Anblagan, Samuel Danso, Susana Muñoz-Maniega, Dominic Job, Cyril Pernet, et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10:519–535, 2018. 1

[10] Alejandro Puente-Castro, Enrique Fernandez-Blanco, Alejandro Pazos, and Cristian R Munteanu. Automatic assessment of alzheimer's disease diagnosis based on deep learning techniques. *Computers in Biology and Medicine*, 120:103764, 2020. 1

[11] Jill Rasmussen and Haya Langerman. Alzheimer's disease–why we need early diagnosis. *Degenerative neurological and neuromuscular disease*, 9:123, 2019. 1

[12] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 2

[13] Sandra Vieira, Walter H.L. Pinaya, and Andrea Mechelli. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience Biobehavioral Reviews*, 74:58–75, 2017. 1

[14] Fei Wang, Lawrence Peter Casalino, and Dhruv Khullar. Deep Learning in Medicine—Promise, Progress, and Challenges. *JAMA Internal Medicine*, 179(3):293–294, 03 2019. 1