

Hybrid CNN-SVM for Alzheimer's Disease Classification from Structural MRI and the Alzheimer's Disease Neuroimaging Initiative (ADNI)

Lan Lin^{1,a,*}, Baiwen Zhang^{1,b}, Shuicai Wu^{1,c}

¹ College of Life Science and Bioengineering, Beijing University of Technology, Beijing, 100124, China

^a lanlin@bjut.edu.cn, ^b zbwrhj@126.com, ^c wushuicai@bjut.edu.cn, *corresponding author

Keywords: Alzheimer's disease; classification; transfer learning

Abstract: Alzheimer's disease (AD) is a progressive neurological disorder among the elders, which results in memory-related issues in subjects. An accurate classification of patients with AD and mild cognitive impairment (MCI) from healthy control subjects (HC) based on structural magnetic resonance imaging (MRI) is of critical clinical importance. In this paper, good intermediate representations of MRI are obtained from a pre-trained convolutional neural network (CNN). Principal component analysis (PCA) and sequential feature selection (SFS) are applied for feature selection, while a support vector machine (SVM) is adopted to evaluate the classification accuracy. 422 Alzheimer's Disease Neuroimaging Initiative (ADNI) baseline MRI were used for development and validation of our proposed method. As a result, this paper achieved a classification accuracy of 90% for binary classification of AD and HC, 81% for AD and MCI and 72% for MCI and HC.

1. Introduction

AD is a popularly known neurodegenerative disease in the elderly population worldwide, which damages the brain cells, and eventually destroys regions of the brain that are responsible for memory, learning, thinking, behaviour and other cognitive functions[1]. According to a recent estimate, about 5.3 million Americans of all ages were suffering from AD in 2015. The affected number is predicted to triple by the year of 2050 [2]. MCI is an intermediate disease stage before the onset of early AD, and subjects with MCI have a high risk of developing AD. No treatment has been reported so far to reverse or stop the pathological damage of AD on the brain. Structural MRI has received considerable attention in supporting the diagnosis of AD.

So far, many machine learning and computer-aided diagnostics methods have been successfully applied to structural MRI analysis to make the diagnosis of different stages of AD more efficient and accurate [3-5]. The concepts of deep learning, inspired from the cognitive processes of human brains, have shown promising performance in extracting features and learning patterns from complex data. It models abstraction from large-scale data at multiple layers, which starts from simpler concepts to more abstract ones. Despite the success of traditional machine learning, deep learning has several attractive features. First, it can directly uncover features from the training data. Therefore, the features don't need to be identified by an expert anymore. As a result, the feature selection process can be simplified. Second, the performance of algorithm increases as the scale of data increases. Third, deep learning solves the problem end-to-end, and therefore we don't need to break the problem down into different parts and solve them individually.

A deep network is a high-capacity machine learning model with many hidden layers of artificial neurons and millions of parameters. State-of-the-art networks are often trained on the ImageNet dataset with over 10 million annotated general images, which is infeasible for most neuroimage studies. Questions have been raised about how to train such a model from scratch. A popular way to solve this problem is a process called transfer learning, which adopts highly refined features from a pre-trained CNN, where the topmost hidden layer is somewhat specialized to the task that it was trained for, while the features from intermediate layers accommodate within-class variance and possess discriminative information that can be extracted and trained with commonly used classifier such as a SVM. This technique can largely reduce the feature learning time and boost the

classification ability compared with training a CNN from scratch.

The aim of this study is to exploit the transfer learning idea into building a CNN to learn invariant and more transferable feature from MRI, which can be applied to classify structural brain MRI as AD, MCI or HC, aiming for individual patient diagnosis. The rest of this paper is organized as follows. We particularize the proposed method in the second part. The results of this study are reported in the third part. We discuss the results and conclude the present study in the last part.

2. Materials and Methods

2.1. Data

ADNI [6] is a public-private partnership program launched in 2003 for collecting data of candidate biomarkers to promote the development of AD research. The data used in this study were selected from the ADNI Grand Opportunity (ADNI-GO) and ADNI-2 studies in ADNI database (adni.loni.usc.edu). The subjects are divided into AD (105 patients), MCI (123 patients) and HC (194 subjects), with average age around 73. Because of the imbalance in the number of subjects in each group, 75 randomly selected subjects from each group were designated as the training subjects, and the remaining subjects were used for testing. The structural MRI data were collected according to the ADNI acquisition protocol using three Tesla scanner. Characteristics of the included subjects are summarized in Table 1.

Table 1 Demographic and behavioral information. MMSE: Mini-Mental State Examination; CDR: Clinical Dementia Rating. The values are denoted as mean \pm standard deviation.

	HC	Patients with MCI	Patients with AD
Number	194	123	105
Male/Female	90/104	67/56	59/46
Age (years)	72.4 \pm 4.9	72.1 \pm 5.5	74.4 \pm 6.2
Education	16.4 \pm 2.3	16.4 \pm 2.1	16.0 \pm 2.5
MMSE	29.2 \pm 2.1	27.4 \pm 1.8	22.9 \pm 2.0
CDR	0.03 \pm 0.21	0.55 \pm 1.26	0.90 \pm 0.41

2.2. MRI data pre-processing

The preprocessed steps are performed with a widely used software package Statistical Parametric Mapping (SPM 8) (<http://www.fil.ion.ucl.ac.uk/spm/>, Wellcome Department of Cognitive Neurology, University College London, UK) that is implemented on MATLAB 7.6 (Mathworks, Natick, MA, USA) environment. All parameters were left at their default values unless specified. The raw MRIs were first inspected to verify they were free from anatomical abnormalities, and were approximately aligned with Montreal Neurological Institute (MNI) space. Then, tissue segmentation was achieved using the New Segment toolbox under the SPM8 distribution, which is essentially the same as the Unified Segmentation model, except for a different treatment of the mixing proportions. T1 images are segmented into three main classes: grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). After that, the registration algorithm, Diffeomorphic Anatomical Registration Through Exponential Lie Algebra (DARTEL) [7] imported GM generated in the previous step and iteratively generated flow fields and a series of template images. GM maps are then modulated by the Jacobian determinants to ensure that regional differences in the total amount were conserved, smoothed using an 8-mm full-width-at-half-maximum (FWHM) Gaussian kernel, and normalized to their average templates and further to the Dartel template in MNI space. The voxel size of normalized data is 2*2*2 mm³.

2.3. Features from AlexNet

In this work, we used the AlexNet [8] architecture, which was trained on a very large dataset containing about one million images labelled with 1000 categories. The dimensions of the smoothed, modulated, normalized GM maps were scaled with trilinear interpolation to $227 \times 227 \times 65$ (65 slices of size 227×227 for each subject), which matches the size of the input layer in AlexNet. Then, the images were exported in a lossless PNG format from NIfTI format. This step produced a total number of 27430 images, with 6825 belonging to the AD group, 7995 belonging to the MCI group, and the remaining 12610 belonging to the HC group.

Table 2 The Data and Parameters Dimension of AlexNet.

Layers	Kernel size	Number of kernels	Number of bias	Total
Conv1	$11 \times 11 \times 3$	96	96	34,944
Conv2	$5 \times 5 \times 48$	256	256	307,456
Conv3	$3 \times 3 \times 256$	384	384	885,120
Conv4	$3 \times 3 \times 192$	384	384	663,936
Conv5	$3 \times 3 \times 192$	256	256	442,624
FC1	$6 \times 6 \times 256 \times 4096$	1	4096	37,752,832
FC2	4096×4096	1	4096	16,781,312
FC3	4096×1000	1	4096	4,097,000

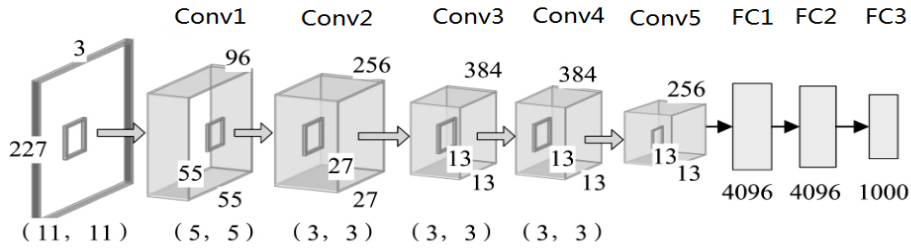


Figure 1 Network Structure

The architecture of the model was described previously[8], which is composed of eight layers: five convolutional layers followed by three fully connected layers (Figure 1). The first convolutional layer takes as input the image of size $227 \times 227 \times 3$, where three is the number of channels for RGB images. Every convolutional layer is set up by four steps, which are sorted as convolution, ReLU, pooling and normalization. Kernel size and kernel number of each convolutional layer and the size of weight matrix of each fully-connected layer are listed in the Table 2. By using transfer learning, the considerable cost of developing large CNNs can be avoided. The features extraction was based on Caffe [9], an efficient framework exploited by Berkeley Vision and Learning Center (BVLC). The intermediate layers capture features that are neither too specific to the dataset the network was originally trained on, nor too general to not contain any representative information from images. Features from conv3, conv4 and conv5 of the AlexNet were extracted using the pre-trained network. For the conv3 and conv4 which do not have a pooling layer followed, we perform overlapping max-pooling within 3×3 regions and step size of 2 to get the final feature vectors. The features from 65 slices of each subject were concatenated together to form the feature representation of the subject.

2.4. Feature selection and SVM

The concatenated feature maps still have too many features; principal component analysis (PCA) and sequential feature selection (SFS) were used for the precise selection of the best features. In this work, we use the average miss-classification error (MCE) of 10-fold cross-validation with a supervised SVM classifier as the objective function of the search. The forward search is used because the number of selected features ($p = 20$) is much fewer than the total PCs. SVM

implementation was taken from the LibSVM toolkit [10]. As the performance obtained with the features extracted was optimal using the linear kernel, we had not explored additional higher dimensional kernels in this study. To give an overall evaluation of the proposed method, besides the prediction accuracy (ACC), we also utilize sensitivity (SEN) and specificity (SPE).

3. Results

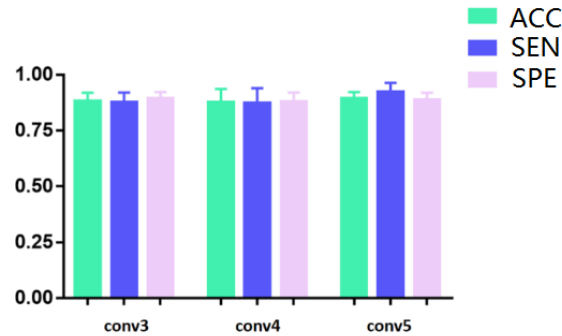


Figure 2 Comparison of classification performance for features extracted from Conv3, Conv4 and Conv5 of a pre-trained AlexNet.

We examined performance of three convolutional layers (Conv3, Conv4 and Conv5) for discrimination of patients with AD from HC subjects without any feature selection (Figure 2). There were not any statistically significant differences between mean classification performances of those layers as determined by one-way ANOVA ($p=0.633$). For the rest of this study, we used the features extracted from the conv3 layer. Table 3 displays the overall performances of the conv3 layer. Feature selection has improved the accuracy from 88% to 90%.

Table 3 Performance of conv3 after feature selection

	Accuracy	Sensitivity	Specificity
AD vs HC	90%	87%	91%
AD vs MCI	81%	70%	88%
MCI vs HC	72%	69%	74%

4. Discussion

One interesting finding in this study was that three convolutional layers offer close performance. Images from the ImageNet dataset consist of 1000 categories, while the MRI images of our AD dataset contain relatively simple and uniform patterns with slight variations. In this sense, deep learning architecture with three or four layers may be sufficient to capture the high-level abstractions. This is the reason why we select conv3 for feature reduction.

In this paper, we present a new classification method to automatically discriminate patients with AD (or MCI) from HC based on MRI data. The feature parameters that showed a discriminating power between patients with AD, MCI and HC were extracted from a pre-trained AlexNet, and we used this feature parameter to create a classifier that can predict a subject's group. Accuracy for classification on AD and HC reaches 90%. Overall, our method has obtained a competitive result compared to previous studies in the literature. Although the proposed approach showed the favorable results, it has downsides as well. The difference between ordinary images and MRI is huge. Not only the colours, but also objectiveness in the two domains is quite different. Generic features in pertained CNN models may not be the best features for an AD classification task. Thus, images needed to fine-tune a CNN trained on ordinary image set to MRI based AD classification is still large. One important future direction is to use unsupervised learning strategy, such as generative adversarial networks (GAN) which reduce the amount of required training samples in

fine-tuning a CNN.

Acknowledgements

This research was financially supported by the Scientific Research General Project of Beijing Municipal Education Committee(KM201810005033).

References

- [1] Lin, L., Fu, Z., and Wu S. (2015) Mouse brain magnetic resonance microscopy: Applications in Alzheimer disease. *Microsc Res Tech.*, 78(5), 416-24.
- [2] Beheshti, I., Demirel H., and Alzheimer's Disease Neuroimaging Initiative. (2016) Feature-ranking-based Alzheimer's disease classification from structural MRI. *Magn. Reson. Imaging*, 34(3), 252–263.
- [3] Tripoliti, E.E., Fotiadis, D.I., and Argyropoulou, M. (2008) A supervised method to assist the diagnosis and classification of the status of Alzheimer's disease using data from an fMRI experiment. *Engineering in Medicine and Biology Society, 30th Annual International Conference of the IEEE*, 4419–4422.
- [4] Zhu, X., Suk, H.I., Lee, S.W., and Shen D. (2016) Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification. *IEEE Transactions on Biomedical Engineering*, 63(3), 607–618.
- [5] Hu, K., Wang, Y., Chen, K., and et al. (2016) Multi-scale features extraction from baseline structure MRI for MCI patient classification and AD early diagnosis. *Neurocomputing*, 175, 132–145.
- [6] Jack, C.R. , Bernstein, M.A., Fox N.C., and et al. (2008) The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging*, 27, 685-691.
- [7] Ashburner, J. (2007) A fast diffeomorphic image registration algorithm. *Neuroimage*, 38, 95-113.
- [8] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems*, 60, 1097-1105.
- [9] Jia, Y., Shelhamer, E., Donahue, J. S. and et al. (2014) Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the ACM international conference on multimedia*, ACM, 675-678.
- [10] Chang C.C., and Lin C.J. (2011) LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2 (3), 1-27