On the design of CNN for automatic detection of Alzheimer's disease

Sheng Liu Chhavi Yadav Carlos Fernandez-Granda* Narges Razavian*

New York University

Introduction

Clinical trials for AD need to target patients at earlier stages before significant brain atrophies. But diagnosing the disease at an early stage is challenging. In this work we focus on learning to differentiate between cognitively normal aging (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD), using structural brain MRI (T1-weighted scans).

Motivation: The performance provided by traditional hand-crafted features is limited.

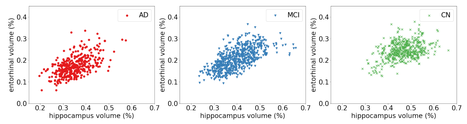


Figure 1: Visualization of intracranial normalized hippocampus and entorhinal volumes of AD, MCI, and CN subjects.

Dataset and preprocessing

Dataset Preprocessed T1-weighted structural MRI scans from the Alzheimer's Disease Neuroimaging Initiative(**ADNI**) dataset.

Preprocessing

- **Clinica** software platform to perform registration.
- Split Train/Validation/Test in subject level.
- Demographics are shown in Table 1. Some patients are double counted when their label transited.

Split	Class	# subjects	# Scans	Mean Age (std)
Train	CN	140	567	77.0 (5.4)
	MCI	248	840	75.9 (7.3)
	AD	193	527	76.7 (7.4)
Val	CN	33	126	77.2 (5.6)
	MCI	39	138	73.3 (7.2)
	AD	41	124	76.1 (8.3)
Test	CN	24	105	79.0 (6.1)
	MCI	43	140	76.7 (6.5)
	AD	45	135	76.4 (5.1)

Table 1: Demographics of our training, validation and test sets after preprocessing.

Challenges

- In contrast to natural images, all scans are registered and have very similar structure.
- The number of examples is orders of magnitude smaller than datasets used to benchmark computer vision tasks.

Method

- **Instance normalization**
- **Small-sized** initial kernels and small strides.
- **Wider** architecture that is not too deep.
- Incorporating **age** information through sinusoids encodings.

Block	Layer	Type	Output size
	Inputs		$96 \times 96 \times 96$
1	Conv3D InstanceNorm3D ReLU	k1-c4-f-p0-s1-d1	$96 \times 96 \times 96$
	MaxPool3D	k3-s2	$47 \times 47 \times 47$
2	Conv3D InstanceNorm3D ReLU	k3-c32-f-p0-s1-d2	$43 \times 43 \times 43$
	MaxPool3D	k3-s2	$21 \times 21 \times 21$
3	Conv3D InstanceNorm3D ReLU	k5-c64-f-p2-s1-d2	$17 \times 17 \times 17$
	MaxPool3D	k3-s2	$8 \times 8 \times 8$
4	Conv3D InstanceNorm3D ReLU	k3-c64-f-p1-s1-d2	$6 \times 6 \times 6$
	MaxPool3D	k5-s2	$5 \times 5 \times 5$
FC1		1024	
FC2		3	
Softmax		3	

Table 2: The backbone architecture. k = kernel size, c = number of channels as a multiple of the widening factor f, p = padding size, s = stride and d = dilation. The age encoding, if used, is forward propagated through two linear layers with layer normalization before being added to the output of FC1.

Comparison to other methods

- Input resolution: $96 \times 96 \times 96$
- Baseline: 3D AlexNet and ResNet-18

Our proposed model outperforms previously reported results by $\sim$ **14%**.

Method	Accuracy	Balanced Acc	Micro-AUC
ResNet-18*	50.8%	-	-
ResNet-18 pretrained*	56.8%	-	-
ResNet-18 3D◇	$52.4 \pm 1.8\%$	53.1%	-
ResNet-18 3D	$50.1 \pm 1.1\%$	$51.3 \pm 1.0\%$	$71.2 \pm 0.4\%$
AlexNet 3D	$57.2 \pm 0.5\%$	$56.2 \pm 0.8\%$	$75.1 \pm 0.4\%$
proposed•	$66.9 \pm 1.2\%$	$67.9 \pm 1.1\%$	$82.0 \pm 0.7\%$
proposed• + Age	$\mathbf{68.2 \pm 1.1\%}$	$\mathbf{70.0 \pm 0.8\%}$	$\mathbf{82.0 \pm 0.2\%}$

* *Results on 2D ResNets initialized with or without pretrained weights on Imagenet reported by Valliani and Soni (2017).*
◇ *3D ResNet with mild modifications, see Fung et al. (2019) for details. The balanced accuracy is computed using the confusion matrix in the paper.*
• *The backbone model showed in Table 2 with a widening factor of 8.*

Table 3: Comparison of the published models to our best proposed models. + Age means that the model incorporates age encodings.

Ablation study

- **Instance Normalization** (IN) outperforms **Batch Normalization** (BN) for different architectures.

Method	Accuracy	balanced Acc	Micro-AUC	Macro-AUC
×4 with IN	$\mathbf{63.2 \pm 1.0\%}$	$\mathbf{63.3 \pm 0.9\%}$	$\mathbf{80.5 \pm 0.5\%}$	$\mathbf{77.0 \pm 0.7\%}$
×4 with BN	$61.8 \pm 1.1\%$	$62.2 \pm 1.1\%$	$77.0 \pm 0.5\%$	$73.0 \pm 0.6\%$
×8 with IN	$\mathbf{66.9 \pm 1.2\%}$	$\mathbf{67.9 \pm 1.1\%}$	$\mathbf{82.0 \pm 0.7\%}$	$\mathbf{78.5 \pm 0.7\%}$
×8 with BN	$58.8 \pm 0.9\%$	$60.7 \pm 0.7\%$	$75.9 \pm 0.7\%$	$73.1 \pm 0.8\%$
ResNet-18 with IN	$\mathbf{52.3 \pm 0.8\%}$	$\mathbf{52.7 \pm 1.1\%}$	$\mathbf{74.1 \pm 0.7\%}$	$\mathbf{73.1 \pm 0.9\%}$
ResNet-18 with BN	$50.1 \pm 1.1\%$	$51.3 \pm 1.0\%$	$71.2 \pm 0.4\%$	$72.4 \pm 0.7\%$

- **Small-sized** initial kernels and strides in the first layer result in better performance.

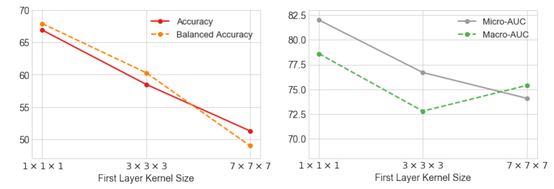


Figure 2: Comparison of the performances of different first layer kernel and stride sizes.

- **Wider** vs **Deeper**: Wider architectures achieve better performance up until a widening factor of ×4. Deeper networks only achieve marginal improvement.

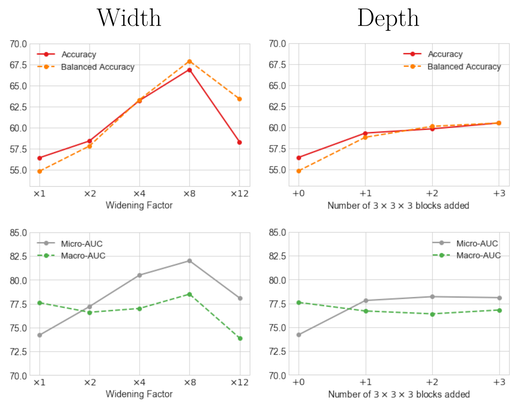
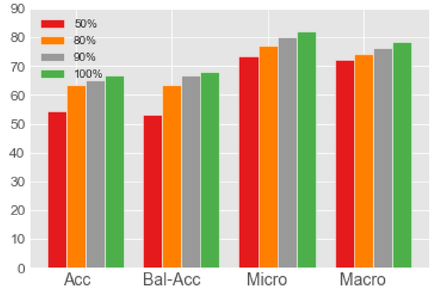


Figure 3: Performance for different widening factors (left) and numbers of added blocks (right) for backbone architecture.

- **Dataset size**: Increasing dataset size improves performance across all evaluation metrics.



Validation with independent dataset

We test the generalization ability of our proposed architecture on the Australian Imaging, Biomarkers and Lifestyle (**AIBL**) which is another longitudinal dataset of Alzheimer's disease. A similar performance achieved as on the **ADNI** data.

Method	Accuracy	Balanced Acc	Micro-AUC	Macro-AUC
proposed on ADNI	$66.9 \pm 1.2\%$	$67.9 \pm 1.1\%$	$82.0 \pm 0.7\%$	$78.5 \pm 0.7\%$
proposed on AIBL	$63.6 \pm 0.7\%$	$65.7 \pm 1.1\%$	$90.0 \pm 0.6\%$	$82.1 \pm 0.7\%$

Table 4: Comparison of the performance of the proposed model on the ADNI and AIBL datasets.

Analysis

The model focuses on gray-matter regions around the hippocampus and the ventricles, which is consistent with existing biomarkers as well as on some additional regions.

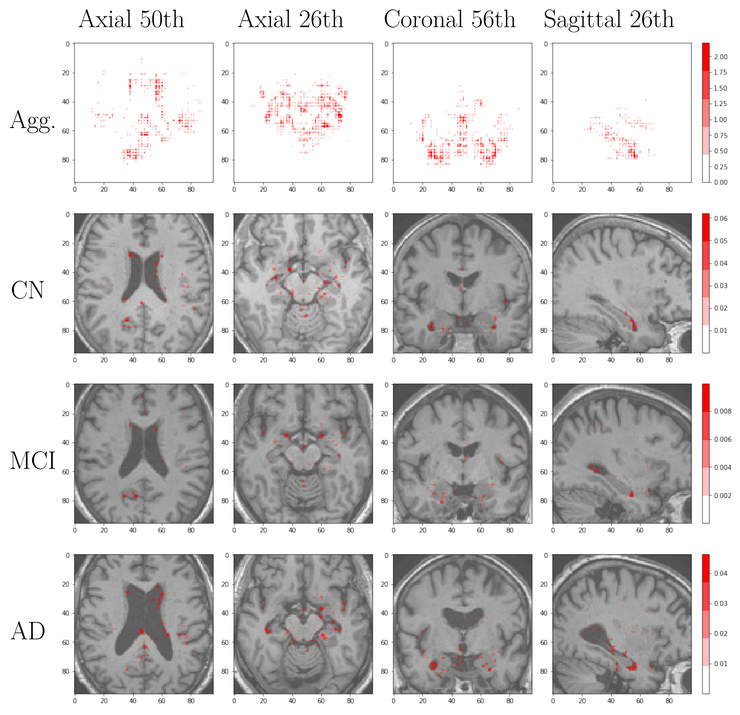


Figure 4: Visualization of class saliency maps (slices). First row: aggregated maps for all validation scans. Other rows: examples for each class.

Questions/Comments? `shengliu@nyu.edu`

Scan for paper Scan for code