Transfer Learning to Predict Early Stages of Alzheimer's Disease Using DenseNet

Hoang (Mark) Nguyen
Department of Computer Science and Electrical Engineering
University of Missouri at Kansas City
hdnf39@mail.umkc.edu

N. N. Chu
CWLab International
nchu@cwlab.com

***Abstract*—In recent years, Deep Learning (DL) in Convolutional Neural Network (CNN) has caught significant attention from research in Alzheimer's Disease (AD) detection, especially in processing brain image data administered by the AD NeuroImage Initative (ADNI.) The increasing number of publications using DL/CNN models has reported accuracy above 93%, with little practical means for validation due to mostly inadequate model description. However, AD remains a challenging topic for DL experiments with the high hope to increase their clinical value. DL is particularly questionable when the quantity of dataset is below statistical significance. In this paper, we propose using transfer learning, a popular technique in DL, to predict transitional stages of AD in terms of Mild Cognitive Impairment (MCI). Our modeling process makes use of 1,854 MRI images of Early stage of MCI (EMCI) and Late MCI (LMCI,) transfer learning from 2,233 MRI images based of Cognitive Normal (CN) and AD states. Our DL/CNN modeling results in performance metrics showing 90.59% in accuracy, 0.947 in Area Under Characteristic Curve (AUC,) and 10.6% Equal Error Rate (EER) in CN-AD classification. The model with transfer learning outperformed the one without pre-trained network with 0.932% in accuracy, 0.932 in AUC, and 14.2% EER.**

I. Introduction

Alzheimer's Disease (AD), a neurodegenerative disorder, is one of the leading causes of death in developed countries. It is a progressive and irreversible disease without a known cure or medicine to reverse the brain deterioration. AD symptoms are also obvious, and it's too late to be treated. Therefore, early-stage diagnosis is much more crucial than AD diagnosis. The initial stage of AD is named Mild Cognitive Impairment (MCI).

In the AD NeuralImage Initiative (ADNI) dataset, each patient is classified into one out of 6 labels: Cognitive Normal (CN), Significant Memory Concern (SMC), EMCI, MCI, LMCI, and AD. The first two labels belong to the healthy group, the next 3 are transitional MCI group, and the last one is definitive AD. MCI is the transition stage of AD, and the conversion time has observed from 6 to 36 months. To provide proper medical assistance for any patients to prolong the transition to AD, it is crucial to make accurate predictions on which MCI stages appear. We propose using transfer learning to predict AD's transitional stages using only MRI images in the ADNI dataset. In this paper, we focus on classifying two labels: EMCI and LMCI, since they are strongly correlated to each other.

Because of a large number of parameters employed in a conventional CNN model dataset, e.g., ImageNet, with 14 million images categorized into 20,000 classes, this Deep Learning approach makes it feasible to generalize any given image information. Unfortunately, there is not many publicly available datasets for AD that meet such large quantity criteria. Transfer learning is a popular technique in the CNN domain that transfers knowledge from the source domain into the target domain. The models are first pre-trained in a large dataset (source domain) and then trained on the target dataset where the size of dataset is usually much smaller. ImageNet pre-trained CNN models is a standard for transfer learning for computer vision research and is widely available in DL libraries such as Tensorflow or Pytorch. Transfer learning can be imposed on multiple domains in a chain. If not starting with millions of images, we consider tens of thousands desirable, or thousands as a minimum in a legitimate forecast.

Magnetic Resonance Imaging (MRI) feeds into ADNI 3D image dataset that is more complicated than the 2D image dataset in ImageNet. No dataset standardizes pre-trained 3D CNN. Therefore, we propose to properly segregate the ADNI dataset to perform transfer learning. AD-CN classification may not be clinically valuable in most pursuit, but it could be a potential source domain. CNN model learns to distinguish the difference between AD and CN, which is useful for initial AD stage classification. The training is divided into two steps: the model is pre-trained on AD and CN data as a source domain and trained again on EMCI and LMCI. We propose to use 3D DenseNet as our CNN model.

II. Literature Review

ResNet and DenseNet are popular CNN models. In [1], Liu proposed to multi-model for AD classification and achieved 88.9% accuracy for classifying AD-CN subjects and 76.2% for MCI-CN subjects. The system consists of a ResNet based segmentation model and a DenseNet model. The former extracts feature from the hippocampus using segmentation masks, and the latter captures the discriminative features from the whole brain for AD diagnosis. Another fully connected layer is added to fuse two features by fine-tuning to make the final decision.

[2] proposed using an ensemble of 3D DenseNet to predict different stages of AD. The paper experimented with different settings of DenseNet, such as dropout, depth, growth

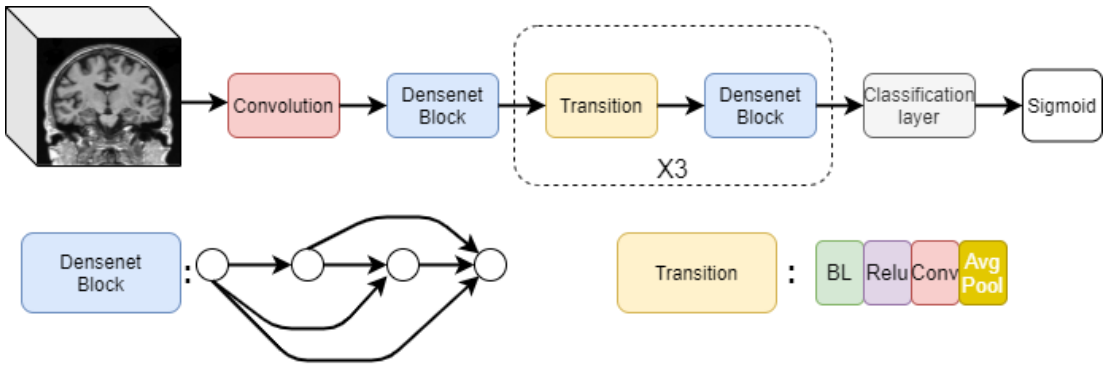

Fig. 1. DenseNet Architecture

rates. The predictions from each model were then combined using probability-based fusion to produce the final prediction. This proposed method achieved 98.83% accuracy in AD vs. CN classification and 97.53% of accuracy in AD/MCI/CN classification.

A few problems have observed with previous publications using DL/CNN with the ADNI dataset. First, they failed to acknowledge that the dataset quantity is not sufficient enough for deep learning. Second, the applied CNN architecture is large in the number of parameters, while ADNI data are high in the number of dimensionalities. Moreover, none took on the same set of data from ADNI. Thus, most of their results, including ours, are not reliable when comparing performance in AD detection. Different from these prior publications, we use only MRI images with a single DenseNet model. We used all the available MRI, 4087 images, in the ADNI dataset, which is at least 5 times the number of data used in prior publications. Because of the curse of dimensionality, we avoid adding more information to DL models, which might hurt our model's learning capability.

III. Proposed Method and Dataset

a) CNN architecture: In this work, we experimented the transfer learning using 3D DenseNet architecture as shown in Fig. 1, DenseNet Architecture 1. This is the 3D version of the DenseNet model in [3]. The model employs the dense block in which the nth layer's feature is obtained by concatenating all the preceding layers. The model implementation was modified from [1] which is a project from [4]. In our DenseNet architecture, the growth rate is 32, the number of filters to learn in the first convolution layer is 64, the multiplicative factor for the number of bottleneck layers is 4. We decided to add a sigmoid layer at the end of the model.

b) Experiment Setup: We trained the DenseNet model with 10e-2 initial learning rate, which was reduced by a factor of 10 if the validation loss did not drop by 10 consecutive epochs. The model was trained for 100 epochs and used the weights from the epoch that achieved the lowest validation loss. The momentum and weight decay were set to 0.9 and 10e-4, respectively. The training was divided into two phases: the first phase was trained on AD-CN data. The second phase was trained on the transitional, EMCI and LMCI labels.

[1]https://github.com/kenshohara/3D-ResNets-PyTorch/blob/master/models/densenet.py

TABLE I
Dataset distribution in 2 phases of transfer learning.

	Phase 1		Phase 2	
	AD	CN	EMCI	LMCI
Training	1081	483	859	440
Validation	142	81	120	65
Testing	295	151	241	129
Total	1518	715	1220	634

c) Dataset: We downloaded 4,087 MRI images from the ADNI dataset. The dataset consists of 4 different labels: AD, CN, EMCI and LMCI. In two phases, the data were divided into 70% training, 10% validation, and 20% testing. Tables I shows the distribution of images in each label in our experiments.

IV. Results and Discussion

Our results in the testing set are shown in Table II: AUC, EER, and Accuracy of our experiments. In the AD-CN experiments, the model achieved 0.947 AUC, 10.6% EER, and 90.59% accuracy which out-performed prior state-of-the-art publications [5]. With transfer learning, the model performed 0.898 AUC, 17.8 % EER, and 87.03 accuracy in EMCI vs LMCI classification. The results were better than the same model without utilizing transfer learning approach. Further examination incorporating sensitivity and specificity of the Confusion Matrices for these three experiments is shown in Fig. 2. Our DenseNet model achieved promising results on surface as most publications in AD-CN classification. Moreover, our proposed approach using transfer learning outperformed the non-pretrained model in predicting transition stages of AD. However, counting all aspects of a performance metrics, it is hard to correlate the high AUC figures to any reliable outcome without having a larger dataset size to begin with.

In conclusion, our exercise of DL/CNN with transfer learning highlights the impact that:

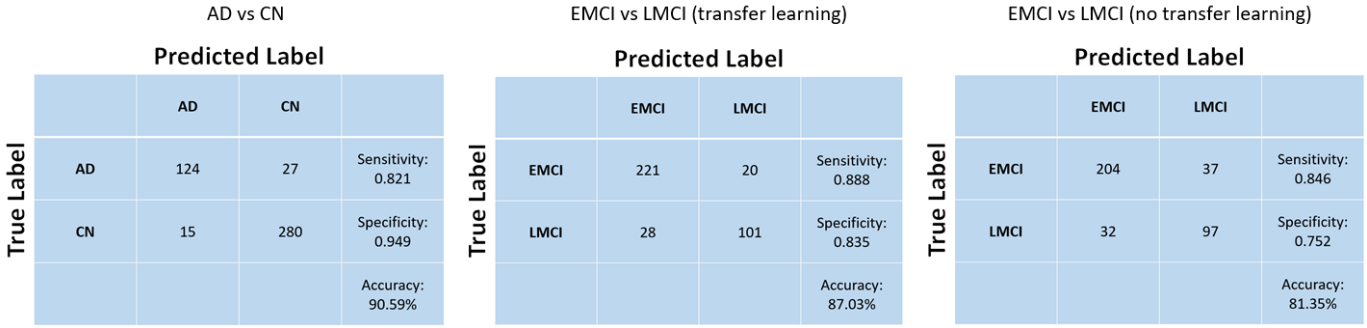

Fig. 2. Confusion Matrix

TABLE II
AUC, EER, AND ACCURACY OF TRANSITIONAL STAGES OF AD CLASSIFICATION.

Performance Metric	Phase 1 AD vs CN	Phase 2 EMCI vs LMCI	P2 w/No Transfer Learning EMCI vs LMCI
AUC	0.947	0.932	0.898
EER (%)	10.6	14.2	17.8
Accuracy (%)	90.59	87.03	81.35

- Forecast of the transitional process toward AD brings the potential to introduce more timely treatment than detection of the set AD.
- CNN promises have to be examined with proper performance metrics including the Confusion Matrix.
- Transfer learning can achieve useful results in improving the decision making capability of the deep learning model.

V. ACKNOWLEDGMENT

The authors would like to thank Saumil Dhankar for his dedication and volunteering effort during the initial 5 months of our experiments for preparing the ADNI datasets, CNN literature review and other essential tasks.

REFERENCES

[1] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, M. Xu, A. D. N. Initiative *et al.*, "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease," *NeuroImage*, vol. 208, p. 116459, 2020.

[2] S. Wang, H. Wang, A. C. Cheung, Y. Shen, and M. Gan, "Ensemble of 3d densely connected convolutional network for diagnosis of mild cognitive impairment and alzheimer's disease," in *Deep Learning Applications*. Springer, 2020, pp. 53–73.

[3] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[4] H. Kataoka, T. Wakamiya, K. Hara, and Y. Satoh, "Would mega-scale datasets further enhance spatiotemporal 3d cnns?" *arXiv preprint arXiv:2004.04968*, 2020.

[5] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-González, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, O. Colliot *et al.*, "Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation," *Medical Image Analysis*, p. 101694, 2020.