This full text paper was peer-reviewed at the direction of IEEE Instrumentation and Measurement Society prior to the acceptance and publication.

A Deep CNN based Multi-class Classification of Alzheimer's Disease using MRI

Ammarah Farooq[1], Syed Muhammad Anwar[2], Muhammad Awais[3], Saad Rehman[1]

[1]Department of Computer Engineering, C.E.M.E, National University of Sciences and Technology (NUST), Islamabad, Pakistan
[2]Department of Software Engineering, U.E.T, Taxila, Pakistan
[3]Center for Vision, Speech and Signal Processing, Surrey University, UK

***Abstract*—In the recent years, deep learning has gained huge fame in solving problems from various fields including medical image analysis. This work proposes a deep convolutional neural network based pipeline for the diagnosis of Alzheimer's disease and its stages using magnetic resonance imaging (MRI) scans. Alzheimer's disease causes permanent damage to the brain cells associated with memory and thinking skills. The diagnosis of Alzheimer's in elderly people is quite difficult and requires a highly discriminative feature representation for classification due to similar brain patterns and pixel intensities. Deep learning techniques are capable of learning such representations from data. In this paper, a 4-way classifier is implemented to classify Alzheimer's (AD), mild cognitive impairment (MCI), late mild cognitive impairment (LMCI) and healthy persons. Experiments are performed using ADNI dataset on a high performance graphical processing unit based system and new state-of-the-art results are obtained for multiclass classification of the disease. The proposed technique results in a prediction accuracy of 98.8%, which is a noticeable increase in accuracy as compared to the previous studies and clearly reveals the effectiveness of the proposed method.**

***Keywords*—Alzheimer's disease; deep learning; MRI; multi-class**

I. INTRODUCTION

Alzheimer's disease (AD) is a neurological brain disorder which causes permanent damage of brain cells associated with the ability of thinking and memorizing. It begins with mild deterioration of nerve cells which gradually leads to intense form of dementia making a patient unable to perform their simple day to day tasks. According to Alzheimer's association, AD is the sixth leading cause of death in United States [1]. It is caused by many factors among which age is the most significant. People with age greater than 65 are at a high risk of suffering from this disease [2]. The brain cortex of the patient shrivels up and severe shrinkage occurs particularly in the hippocampus area. This region is involved in thinking, reasoning and making new memories. Brain ventricles, which produce cerebrospinal fluid, also become larger in an AD patient. A well-timed diagnosis of this disease is crucial and requires good clinical assessment based on patient's medical history, several neuropsychological tests such as mini-mental state examination (MMSE), neuropsychiatric inventory questionnaire, clinical dementia rating and other pathological evaluations. National Institute of Aging Alzheimer's Association developed the first clinical criteria for AD diagnosis [3]. In addition to these clinical methods, modern techniques use various imaging technologies for the purpose. Magnetic resonance imaging (MRI) and positron emission tomography (PET) techniques are popular and non-invasively characterize the structure of brain. Cerebrospinal fluid (CSF) analysis is also used in practice.

Tong et al. [4] proposed a multi-modal framework which first calculates pairwise similarity of features for each modality separately and then combines similarities from different modalities into a unified graph for classification using non-linear graph fusion technique. Similarly, Sorenson et al. [5] combined various MRI biomarkers; cortical thickness measurements, volumetric measurements, hippocampal shape, and hippocampal texture, for multi-class classification. Zhang et al. [6] used a combined kernel technique with SVM classifier to classify AD and normal. The combined kernel is made by a fusion of features from the above mentioned three modalities.

The latest researches in computer vision and machine learning are motivated by neural networks and deep learning. Deep Learning is a representation learning technique which allows a machine to learn distinctive representations from raw data. The reason behind the popularity is the hierarchal and layered structure of the network. Convolutional neural networks (CNN) are inspired by human visual system and learns the features through compositional hierarchy of object starting with simple edges and moving towards more complex forms. A stack of convolutional and pooling layers achieves this. Each convolutional layer finds local conjunctions of features from the preceding layer and pooling combines the similar features into one by downsizing the feature map [7]. Deep learning has also stimulated the neuroscience researchers and they started to find their solutions for problems associated with neuro-imaging. Lui et al. [8] used zero-masking strategy for fusing data from different modalities and then trained a stacked autoencoder (SAE) network for classification. In another work, Lui et al. [9] used the idea of using multi-phase features followed by SAE and a linear as well as softmax classifier. They took MMSE as low level

978-1-5386-1620-8/17/$31.00 ©2017 IEEE

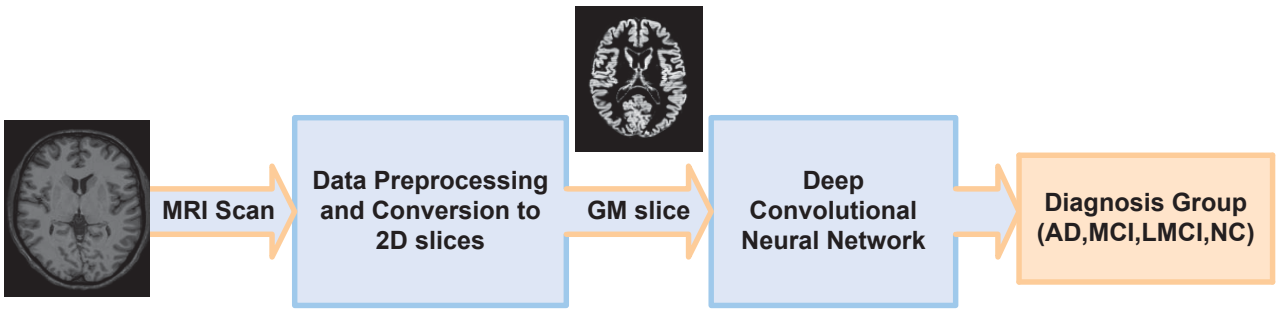

Fig. 1. The proposed deep learning pipeline for 4-way classification of Alzheimer's into AD, MCI, LMCI and NC

feature and high level feature included multimodal neuroimaging data. Shi et al. [10] implemented multimodal stacked deep polynomial networks (SDPNs) for classification. Two separate SDPN learnt features from MRI and PET data and the outputs were then fused and passed to a final stage SDPN. They achieved about 97% accuracy for diagnosing AD from normal. However, results of these researches were as good as results from using image processing techniques for multi-class classification.

A good improvement in classification accuracy is seen in the work by Gupta et al. [11], where they first learnt bases from natural images and MRI scans using a SAE network and then convolved these bases with MRI data to learn features while training. They obtained 85% accuracy but using one-vs-all strategy for three classes. Payan et al. [12] used similar method; performed 3D convolution of pre-trained bases with MRI data and passed the feature maps to fully connected network. They achieved 89.5% accuracy for classifying AD, MCI and normal in a single setup. Hosseini et al. [13] expanded the concept of using pre-trained bases and implemented 3D convolutional autoencoders (CAE) with three different scales to learn feature bases for the three convolutional layers of CNN architecture. Multiple fully connected layers were used on top of convolutional layers for class evaluations. This method improved the accuracy to 94%. However, Sarraf et al. [14] were the first to implement the diagnosing pipeline with pure CNN and without any pre-training. They trained their network for MRI as well as resting state functional MRI data separately using LeNet and GoogLeNet models and achieved best results for binary classification.

In this work, a CNN based pipeline is implemented for multi-class classification of Alzheimer's disease. Multi-class pipeline is a 4-way classifier which classifies AD, early and late stages of mild cognitive impairment as MCI and LMCI respectively and normal cognitive as NC. The rest of the paper presents methodology and implementation in Section 2, experiments and results in Section 3, and conclusion in Section 4.

II. Methodology and Implementation

Figure 1 shows the proposed pipeline which consists of two steps; preprocessing and network training. The details are presented in the following subsections.

A. Data Preprocessing

MRI scans are provided in the form of 3D Nifti volumes. At first, skull stripping and gray matter (GM) segmentation is carried out on axial scans through spatial normalization, bias correction and modulation using SPM-8* tool. GM volumes are then converted to JPEG slices using Python Nibabel package. Slices from start and end which contain no information are discarded from the dataset. Processed slices for each class is shown in Figure 2.

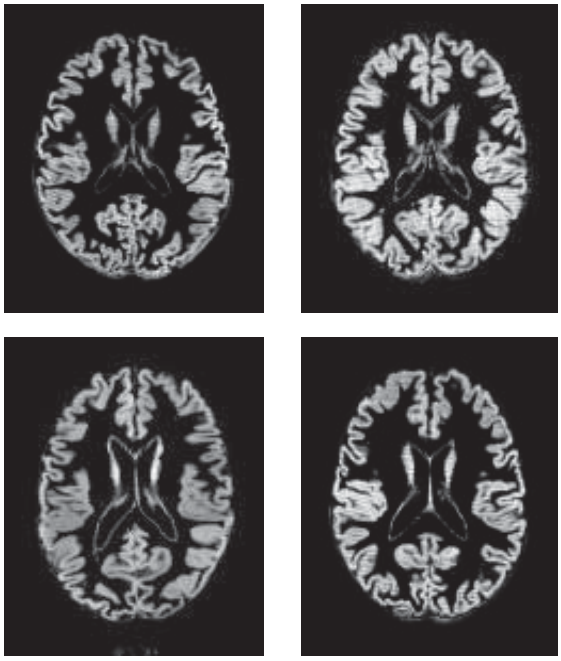

Fig. 2. Gray Matter slices for each diagnosis group after pre-processing. (*Top left: AD, right: LMCI, bottom left: MCI, right: NC).* Images are very similar and require highly discriminative features for correct diagnosis.

*available at (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/)

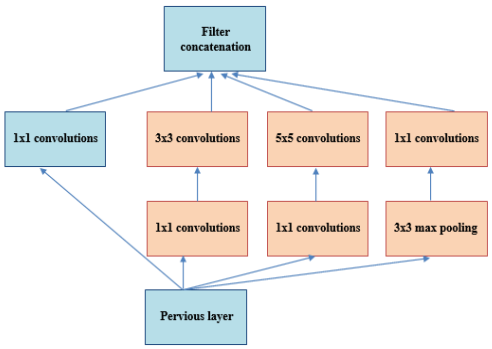

Fig. 3. GoogLeNet Inception module

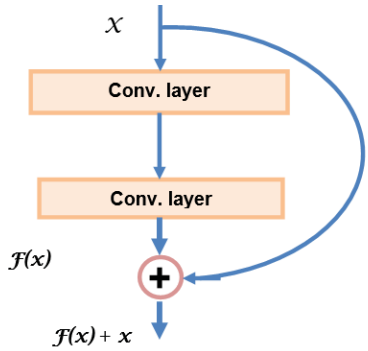

Fig. 4. Residual learning block

B. *Network Architecture*

As mentioned earlier, human visual cortex inspires convolutional neural networks. Each neuron responds to stimuli occurring in its receptive field. This operation is similar to convolution and input image patch works as its small receptive field. When input pattern matches with the convolving filter, response is produced in the form of feature maps. Apart from convolution, CNN also include other layers like rectified linear unit (ReLU) activations, pooling layers, normalization and fully connected (FC) layers. The obtained feature maps are passed through several conv-ReLU-pool operations and reach the final single or multiple fully connected layers. Convolution operation considers local connectivity, parameter sharing and shift invariance, hence, becoming more powerful than learning feature through fully connected neurons.

Highly complex CNN architectures have been developed so far including AlexNet, ZFNet, VGG-Net, GoogleNet and ResNet from ImageNet challenge [15]. These models are successfully being used in numerous applications like object recognition, detection, segmentation and more. The proposed framework is also experimented upon two state-of-the-art models; namely GoogLeNet and ResNet the details of which are as follows.

1. *GoogLeNet*

It is 22-layer model introduced in 2015 by Google team and was the winner of ILSVRC-2014 [16]. It is the first architecture that differs from traditional CNN architectures i.e., stacking more layers on top sequentially. This model takes care of the computational budget, while allowing processing in depth as well as width. Increasing depth on one hand may produce better results but on other hand it also increases learnable parameters and chances of over-fitting in case of small labeled data available. Uniqueness of this network lies in its "Inception module", formed based on Hebbian principle and multi-scale processing as shown in Figure.3. The idea behind this module is considering maximum information coming from the input patch. Different sizes of filter extract information on fine scale as well as cover large scale, whereas max-pool and 1×1 convolution layers reduce dimensionality. The information is aggregated and passed to the next layer. Overall, the model take control over the number of learnable parameters and has 12 times fewer parameters than AlexNet [16].

2. *ResNet*

Residual Network (ResNet) was developed by Kaiming et al. [17] and was winner of ILSVRC-2015. The architecture included a central concept of shortcut or skip connections which are added as bypass to convolutional layers of regular feed-forward network; making the block a residual block as shown in Figure 4. The main idea is that at each step instead of learning features from a function F(x), the features are learnt from F(x) plus original x which makes optimization easier and makes the model to converge faster. For normal feed-forward networks, the prediction accuracy decreases as the depth of network increases. Many factors are responsible for such results including vanishing gradient problem, saturation, size of training data and over-fitting. Residual learning allows network depth to become as deep as more than thousand layers. During backward pass, skip connections make flow of gradient easy and solve vanishing gradient problem. A ResNet with 152 layers is 8 times deeper than VGG network but still have lower computational complexity [17].

III. EXPERIMENTS AND RESULT DISCUSSION

A. *Dataset*

The dataset used in this study is provided by Alzheimer's disease Neuroimaging Initiative (ADNI)* [18]. The demographic information of the dataset is presented in Table 1. The selected set of MRI scans is acquired from 3T scanners. A subject is scanned at different point of times in different visits i.e., baseline, after one two and three years.

*dataset available at (http://adni.loni.usc.edu/)

Each such scan is considered as a separate subject in this work. The dataset consists of 33 AD, 22 LMCI, 49 MCI patients and 45 healthy controls which makes a total 355 MRI volumes. A balanced multiclass data is prepared for classification by using data augmentation for the class which has smaller number of images to get balanced with other classes. Augmentation is done by simply flipping the image along horizontal axis. This operation is valid because of the left and right symmetry of brain regions. The balanced set includes a total of 9506 images for each class, and a total of 38024 images for all classes.

TABLE 1.
DEMOGRAPHIC INFORMATION OF DATASET

Classes	*No. of Subjects*	*Sex*	*Age*	*No. of MRI Volumes*
AD	33	22 F, 11 M	74 ± 8.3	73
LMCI	22	7 F,15 M	73 ± 7.4	61
MCI	49	19 F, 30 M	75 ± 8.2	84
NC	45	27 F, 18 M	75 ± 3.7	137

B. Experimental setup and Evaluation

Networks are trained from scratch on data for 100 epochs with each epoch consisting of 1000 batches. All experiments are performed by splitting data into 25% as test and 75% as train data. 10% data from train set is used as validation set. Xavier initialization of weights is adopted to speed up the training of deeper networks. Implementation is done using the Torch7 framework [19]. Table 2 shows detailed results of the proposed CNN framework. The comparison of proposed pipeline with the literature along with techniques and data modalities is discussed in Table 3. All of the mentioned approaches have worked with ADNI dataset. Performance is evaluated in terms of accuracy (ACC), sensitivity (SEN), specificity (SPE) and positive predictive value (PPV) or precision. For ResNet model, we have obtained results for 18 layers and 152 layers. In terms of classification accuracy, the presented framework has

TABLE 2.
DETAILED PERFORMCE OF PROPOSED FRAMEWORK FOR EACH MODEL

Model	*Class*	*SPE (%)*	*SEN (%)*	*PPV (%)*	*Overall ACC (%)*
GoogLeNet	**AD**	99.2	**97.9**	97.6	98.88
	LMCI	**99.9**	**99.9**	**99.8**	
	MCI	99.6	**97.9**	98.6	
	NC	**98.6**	96.3	**97.6**	
ResNet-18	**AD**	**99.5**	97.1	**98.5**	98.01
	LMCI	**99.9**	**99.9**	99.7	
	MCI	**99.8**	96.4	**99.3**	
	NC	97.9	**97.9**	93.94	
ResNet-152	**AD**	99.3	97.6	98.1	98.14
	LMCI	**99.9**	**99.9**	**99.8**	
	MCI	99.6	97.4	99.1	
	NC	98.4	97.2	95.4	

TABLE 3.
COMPARISON OF CLASSIFICATION PERFORMANCE OF PROPOSED FRAMEWORK WITH OTHER TECHNIQUES

Approach	**Technique**	**Modalities**	**Classification**	**Accuracy (%)**
Lui et al.	SAE-Zeromask	MRI +PET	4 way (AD/cMCI/ncMCI/NC)	53.8
Jun shi et al.	SDPN	MRI +PET	4 way (AD/cMCI/ncMCI/NC)	57
Lui et al.	MPFR	MRI + PET + Clinical	3 way (AD/MCI/NC)	59.19
Tong et al.	NGF	MRI + PET +CSF	3 way	60.2
Sorensen et al.	Combined biomarkers	MRI	3 way	62.7
Suk et al.	DW-S^2 MTL	MRI + PET + CSF	3way	62.93
Gupta et al.	SAE	MRI	3 way	85 with natural image bases, 78.2 with MRI bases
Payan et al.	SAE-CNN	MRI	3 way	89.4 with 3D convolution 85.53 with 2D convolution
Hosseni et al.	DSA- 3D CNN	MRI	3 way	94.8
Proposed	**GoogLeNet**	**MRI**	**4 way (AD/MCI/LMCI/NC)**	**98.88**
Proposed	**ResNet-18**	**MRI**	**4 way**	**98.01**
Proposed	**ResNet-152**	**MRI**	**4 way**	**98.14**

TABLE 4.
CLASS SPECIFIC PERFORMANCE COMPARISON OF PROPOSED FRAMEWORK WITH OTHER TECHNIQUES

Models	AD			MCI			NC		
	SPE	**SEN**	**PPV**	**SPE**	**SEN**	**PPV**	**SPE**	**SEN**	**PPV**
Lui et al.	-	-	61.52	-	-	61.29	-	-	49
Sorensen et al.	-	40.3	-	-	57.8	-	-	79	-
Gupta et al.	91.8	95.9	-	92.7	74.2	-	91.3	87.7	-
Hosseini et al.	-	100	100	-	80	60	-	47	70
GoogLeNet	99.2	97.9	97.6	99.6	97.9	98.6	98.6	96.3	97.6
ResNet-18	99.5	97.1	98.5	99.76	96.4	99.3	97.9	97.9	93.94
ResNet-152	99.3	97.6	98.1	99.6	97.4	99.1	98.4	97.2	95.4

clearly outperformed other approaches for all of the three models. The highest accuracy of 98.88% is achieved by GoogLeNet while ResNet-18 got 98.01% and 98.14% for ResNet-152. Although, the difference in three values is quite small but ResNet-152 performed lower in accuracy than GoogLeNet which was against the expectation made from ImageNet results. Also, moving from 18 layers to 152 layers only improved accuracy by 0.13%. This performance may account for more amount of training data for very deep networks. However, in terms of training time, ResNet-18 was fast owing to its smaller number of layers whereas other two models took similar time. Overall 4% gain in accuracy is achieved by our framework without using any pre-learnt features.

Class-wise performance comparison of each model with literature is shown in table 4. Among the three competing approaches [11, 12, 13], only [11] and [13] have evaluated class-wise performances. Sensitivity is a class accuracy in other words. Hosseini et al. [11] and Gupta et al. [13] achieved class sensitivities (AD/MCI/NC (%)) of 100/80/47 and 96/74/88 respectively. Although the sensitivity for AD class is better but for normal class, it is indicating a considerable situation where 53% and 12% are misclassified as patient. Our work also improved the class performances. We achieved the sensitivities as high as 97.9% for the three classes and even higher for LMCI. Similar pattern and comparison is present for specificity and classification precision with [11, 13]. In all cases, the proposed method improved performance of each class.

IV. CONCLUSION

In this study, we presented a convolutional network based framework for classifying structural MRI images to diagnose Alzheimer's disease, its prodromal stages MCI and LMCI and normal controls. Experimental data was obtained from ADNI and total 355 volumes from 149 subjects were used. MRI scans were first preprocessed to get GM images which were then passed to CNN network. Networks were trained and tested using deep GoogLeNet and ResNet models. Results from both the models outperformed all the other methods of literature in regard of multiclass classification. About 4% increase in classification accuracy is achieved by proposed approach. As compared to most of the previous works, pre-trained feature learning is not required anymore and still the network can accurately predict the classes. Class specific performance gain is also accomplished improving performance for all classes. Hence, it demonstrates the potential of incorporating deep models directly from scratch for learning distinctive features from neuroimaging data and has high-level of implications for medical as well as neuro image processing. Such systems are reliable as well as error free. Future work may include incorporating clinical data of patients along with imaging data and dealing with multimodal data to build more robust system.

REFERENCES

[1] Alzheimer's Association, "2016 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 12(4), pp. 459-509, 2016.

[2] I. O. Korolev, "Alzheimer's disease: a clinical and basic science review," Medical Student Research Journal, vol. 4, pp. 24-33, 2014.

[3] G. M. McKhann et al., "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimers and Dementia, vol. 7(3), pp. 263-269, 2011.

[4] T. Tong, K. Gray, Q. Gao, L. Chen, D. Rueckert, Alzheimer's Disease Neuroimaging Initiative, "Multi-modal classification of Alzheimer's disease using nonlinear graph fusion," Pattern Recognition, vol. 63, pp.171-181, 2017.

[5] L. Sørensen et al., "Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry," NeuroImage: Clinical, vol. 13, pp. 470-482, 2017.

[6] D. Zhang et al., "Multimodal classification of Alzheimer's disease and mild cognitive impairment," NeuroImage, vol. 55(3), pp. 856-867, 2011.

[7] Y. LeCunn, Y. Bengio, G. Hinton, "Deep learning," Nature, vol. 521(7553), pp. 436-444, 2015.

[8] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, M.J. Fulham, “Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease,” IEEE Transactions on Biomedical Engineering, vol. 62(4), pp. 1132-1140, 2015.

[9] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, D. Feng, “Multi-phase feature representation learning for neurodegenerative disease diagnosis,” Australasian Conference on Artificial Life and Computational Intelligence, pp. 350–359, 2015.

[10] J. Shi, X. Zheng, Y. L, Q. Zhang, S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease," IEEE Journal of Biomedical and Health Informatics , vol. PP(99), 2017.

[11] A. Gupta, M. Ayhan, A. Maida, “Natural image bases to represent neuroimaging data,” International Conference on Machine Learning, pp. 987–994, 2013.

[12] A. Payan, G. Montana, “Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks,” arXiv preprint, arXiv: 1502.02506, 2015.

[13] E. Hosseini-Asl, G. Gimel'farb, A. El-Baz, “Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network,” arXiv preprint, arXiv: 1607.00556, 2016.

[14] S. Sarraf, G. Tofighi, “DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI,” bioRxiv preprint, bioRxiv: 070441, 2016.

[15] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115(3), pp. 211-252, 2015.

[16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-9, 2015.

[17] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” arXiv:1512.03385, 2015.

[18] C. R. Jack et al., “The Alzheimer’s disease neuroimaging initiative (ADNI): MRI methods,” Journal of Magnetic Resonance Imaging, vol. 27(4), pp. 685–691, 2008.

[19] R. Collobert, K. Kavukcuoglu, and C. Farabet, “Torch7: A MATLAB-like environment for machine learning,” BigLearn, NIPS Workshop, 2011.

[20] H.I. Suk, S.W. Lee, D. Shen, “Alzheimer’s Disease Neuroimaging Initiative. Deep sparse multi-task learning for feature selection in Alzheimer’s disease diagnosis,” Brain Structure and Function, vol. 221(5), pp. 2569-2587, 2016.

[21] S. Vieira, W. H. L. Pinaya, A. Mechelli, "Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications," Neuroscience & Biobehavioral Reviews, 2017.