



3D CNN based Alzheimer's Diseases Classification using Segmented Grey Matter Extracted from Whole-brain MRI

Bijen Khagi^a, Goo-Rak Kwon^{a,*}

^a Department of Information and Communication Engineering, Chosun University, Dong-Gu, Gwangju 501-759, Republic of Korea

Corresponding author: *grkwon@chosun.ac.kr

Abstract—A recent study from MRI has revealed that there is a minor increase in cerebral-spinal fluid (CSF) content in brain ventricles and sulci, along with a substantial decrease in grey matter (GM) content and brain volume among Alzheimer's disease (AD) patients. It has been discovered that the grey matter volume shrinkage may indicate the possible case of dementia and related diseases like AD. Clinicians and radiologists use imaging techniques like Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scan, and Positron Emission Tomography (PET) to diagnose and visualize the tissue contents of the brain. Using the whole brain MRI as the feature is an on-going approach among machine learning researchers, however, we are interested only in grey matter content. First, we segment the MRI using the SPM (Statistical parameter mapping) tool and then apply the smoothing technique to get a 3D image of grey matter (later called as grey version) from each MRI. This image file is then fed into 3D convolutional neural network (CNN) with necessary pre-processing so that it can train the network, to produce a classifying model. Once trained, an untested MRI (i.e. its grey version) can be passed through the CNN to determine whether it is a healthy control (HC), or Mild Cognitive Impairment (MCI) due to AD (mAD) or AD dementia (ADD). Our validation and testing accuracy are reported here and compared with normal MRI and its grey version.

Keywords—Alzheimer disease (AD); Magnetic resonance imaging (MRI); Statistical parameter mapping (SPM); Convolutional neural network (CNN).

Manuscript received 29 Jan. 2021; revised 22 Mar. 2021; accepted 4 Apr. 2021. Date of publication 30 Jun. 2021.
International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

In regards to Alzheimer's disease, it is a neurodegenerative disease that influences the functional and constructional parts related to the brain. It is one of the most familiar forms of dementia that develops problems with memory, behavior, thinking, and other intellectual abilities disturbing personal and socio-economic aspects as well.

A recent study from MRI has revealed that there is a minor increase in CSF content in brain ventricles and sulci, along with a substantial decrease in GM content and brain volume among AD patients [1]. The segmented tissue content reveals the volume of each type, and as AD is a neurodegenerative disease, the shrinking brain volume may alarm the case of a possible diagnosis of brain atrophy that may cause dementia and finally AD.

MRI is a magnetic-field gradient-based neuroimaging biomarker technique that provides anatomic and physiological information for diagnosis [11] of different parts of the body including the brain. It uses a strong magnetic field

and radio-wave to generate a higher-quality picture of the structure and volume of the brain. The high quality and greater contrast image of the anatomical structures along with functional images of various organs helps the medical professionals to obtain maximum data and information without any physical operation of the participant [12].

Formerly Convolutional Neural Network (CNN) was designed for object recognition and later found its use in image classification, signal prediction, image-segmentation, pattern recognition etc. Due to its autonomous functioning nature, it has been exploring as an important deep learning tool in the field of artificial intelligence (AI) and advanced computer vision. In 2012 A. Krizhevsky et al. [2] were able to successfully engage CNN was in the larger database classification of natural images with the lowest top 5 error rate i.e. 15.3% in the ImageNet database with one thousand classes of image types. Later various advanced variants of CNN were proposed by deep learning researchers for object recognition and image classification including the one of residual network Resnet50 [3], inception network GoogLeNet [4], and regional boundary box-based r-CNN [5]. Regarding AD detection

using imaging modalities various architectures have been proposed. Payan et al. [18] proposed a patch-based sparse auto-encoder (SAE-CNN) to classify the MRI scans employing dataset partitioning. Hosseini-Asl et al. [19] used a deeply supervised and adaptable 3D CNN (DSA-3D-CNN) for s-MRI classification. Oh et al. [20] proposed a convolutional auto-encoder (CAE) constructed as 3D volumetric CNN for AD vs. normal older control (NC) and also proposed sMCI vs. pMCI classification using supervised learning transfer. Later Liu et al. in 2018 [21] proposed a modest CNN architecture with concatenation done in the convolution layer. In our recent work, we have proposed diverging architecture-based CNN for proper feature extraction and classification of s-MRI [14].

The goal of this paper is to prove that the tissue segmented MRI can be effective than a normal MRI with diverse pixel value for CNN-based. Our finding on a limited dataset to some extent supports our hypothesis. More study in the larger dataset is still under study.

II. MATERIAL AND METHODS

SPM was used to perform the segmentation of the brain into 3 tissue types so a separate 3D image file is obtained in Nifti format for Grey, White, and CSF parts. Being Grey matter most suspicious part of our study. The training and testing MRI files were obtained from National Research Centre for dementia (NRCD) Korea [6] now also known as Gwangju Alzheimer's disease and Related Dementias (GARD) center. From the total dataset pool, only a few were selected for our experiment. The dataset consists of 42 Alzheimer's disease dementia (ADD), 42 NC, and 39 MCI due to AD (mAD). ADD consists of 24 males and 18 females of mean ages 76.25 ± 3.33 and 75.03 ± 6.29 , respectively. NC consists of 24 males and 18 females of mean age's 76.26 ± 4.57 and 69.66 ± 3.09 , respectively. mAD consists of 24 males and 15 females of mean ages 74.75 ± 3.588 and 72.06 ± 2.89 , respectively. The reason behind using fewer datasets is to test the implementation of our idea i.e. grey version may work better than processed MRI, in a simpler way as much as possible.

Firstly, the raw MRI file is pre-processed using the co-registration function available in the same SPM, then skull stripped and segmented following the segmentation pipeline of bias correction and spatial normalization using TPM (tissue probability map) from ICBN brain template [7] [8]. The obtained grey version is transformed as shown in Figs. 1 and 2.

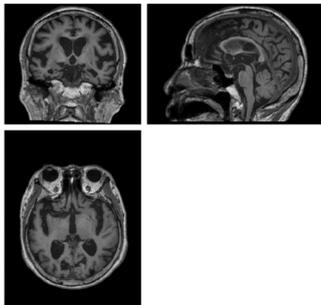


Fig. 1 MRI scan of a normal subject as obtained from the NRCD dataset. The figure shows the MRI in the x-y plane as the 2D image in coronal, sagittal, and axial plane, respectively.

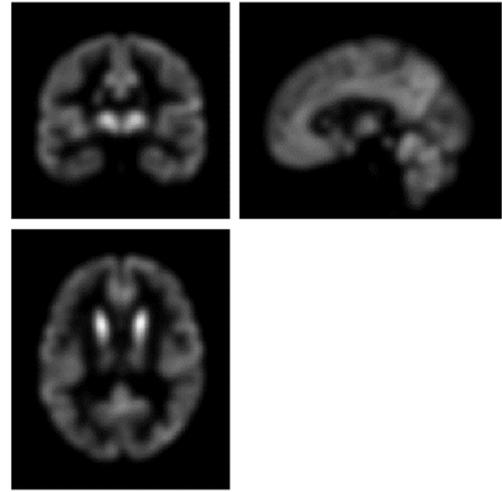


Fig. 2 MRI Smooth grey version after pre-processing steps in SPM for MRI input of Fig. 5. The figure shows the processed, segmented, and smooth MRI in the x-y plane as the 2D image in coronal, sagittal, and axial plane respectively.

A. SPM based segmentation

The major procedure in VBM includes i) spatial normalization and diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL) registration, ii) modulation and segmentation, and iii) smoothing. Spatial normalization transforms all the participants' volume to the identical stereotactic space for uniformity. This is achieved by registering each of the images to its identical template image, by reducing the residual sum of squared differences amongst them using affine transformation [13] and nonlinear registration for the global brain shape difference. Consequently, modulation is performed to compensate for volume changes owing to spatial normalization.

DARTEL [13] registration template is used to perform spatial registration. DARTEL template was created from 555 IXI participants between the ages of 20–80 years. It provides an SPM12 extension tool for achieving a more precise inter-participant registration of brain images. The extension tool uses default tissue probability maps (TPMs) as a reference map to perform the initial spatial registration steps and later tissue wise segmentation of brain. This TPM is a reformed version of the ICBM Tissue probabilistic Atlases (from 452 T1 weighted Human brain scans) provided by the International Consortium for Brain Mapping [8]. Moreover, an optimized shooting approach was used for the adaptive threshold and lower initial resolutions to acquire a good trade-off between accuracy and calculation. The segmented tissue was smoothed to suppress noise and effects due to the residual difference in functional and gyral-related anatomies during inter-participant averaging. The final image was smoothed using [8 8 8] the Gaussian smoothing kernel. Hence, the normalized-modulated-segmented smooth image of voxel 1.5 mm and dimensions $121 \times 145 \times 121$ was finally formed for each tissue volume, i.e., GM, white matter (WM), and CSF. We have considered the GM and WM volume as the major input for the further mapping process.

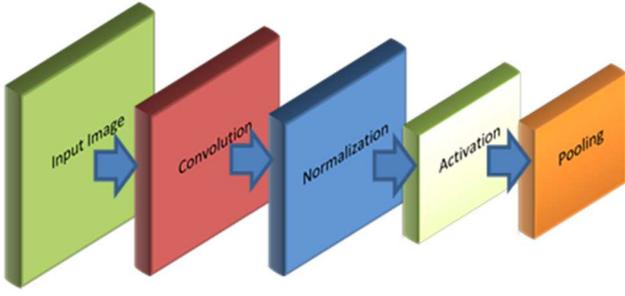


Fig. 3 Convolutional layer architecture in CNN. The block shows the basic layers used in CNN i.e. convolution operation followed by a batch or channel normalization layer which is followed by an activation layer and finally a pooling layer for feature down-sampling.

B. CNN design for classification

Conventionally CNN contains many convolutional layers that transform their input with convolution filters initialized differently with various size and stride of a small extent that runs over each image to pass the extracted feature vector to the succeeding layers. CNN is a supervised training phenomenon as it requires user-defined target values generally called a label, ground truth, or target value. Based on the error between the predicted value and target value, the loss function performs the iterative-training for different epochs using backward propagation until the parameters of all layers participating in training remain constant or almost constant with the minimum error between the predicted value and target value. Here it is to be noticed that, training a CNN is directly affected by the number of training materials and the quality of label or ground truth. The network performs accordingly how it is trained hence called supervised network. But firstly before going into detail into the work, it will be helpful if we go through some major layer-wise mathematical operations used inside the CNN network as shown in Fig. 3.

1) Convolutional layer: This is a learnable layer with multi-dimensional filters (kernels) of a specified size that runs across the input signal (image). Mathematically kernel is a 2D or 3D square matrix to be operated with the input signal. The hyper-parameter step or stride controls the area of reception for the filter to convolve through the input signal by sliding the window with each stride size. The convolution operation of the input signals with the kernel follows equation (1)

$$y_l = \sum_n^{N-1} x_n w_{l-n} \quad (1)$$

The convolution operation follows as above equation: where x_n is the signal input for layer l , w_l is its filter weight, and ' n ' is the number of elements in x . For the next preceding layer input, the output vector y_l becomes the input. The subscripts represent the n^{th} element of the feature vector.

The output of the convolution is a reduced version of the input image known as the feature map or feature vector. Here, one important consideration is the initial constituent of the filter also known as filter-weight, which is normally a random value however different initialization techniques have been proposed to enhance the convergence process of the network.

2) Pooling layer: The feature vector or feature map obtained from convolution is bulkier in dimension due to the larger

number of filters used hence pooling operation is performed to select a representative feature map. The pooling layer works as the down-sampling layer, eventually decreases the size of output of feature vectors from the convolutional layer which may cause extra memory-hardware consumption and overfitting. Various types of pooling action may be average pooling, max-pooling, min-pooling which selects the average, maximum and minimum value from the selected pool size filter respectively. The generally used pooling is the max-pooling function which forwards the maximum value from its selected window, for generating a feature map [15].

3) Activation layer: It a common practice to uses various activation functions to transform the feature between each layer so that the convolution process gets smoother and faster without losing important information. Mostly used is:

(a) Rectified linear activation unit (ReLU): Rectifier linear units [16] add non-linearity during training the network and select only the non-negative numbers as activated features as shown in equation (2).

$$f(x) = \max(0; x) \quad (2)$$

As the equation suggests it misses the negative weights to maintain a range of $[0, x]$ but a slightly different ReLU called leaky rectifier linear unit (LeakyRelu) [26] proved better than the original ReLU itself. This may be due to its characteristics which add nonlinearity, sparsity in the convolutional network resulting in network robustness to minor fluctuations such as noise present along with the input. Similarly, exponential linear unit (ELU) also keeps a minimum threshold for negative inputs. All activation functions shown as graph in Fig. 4. Equation (3) represents the single input LeakyRelu function.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \quad (3)$$

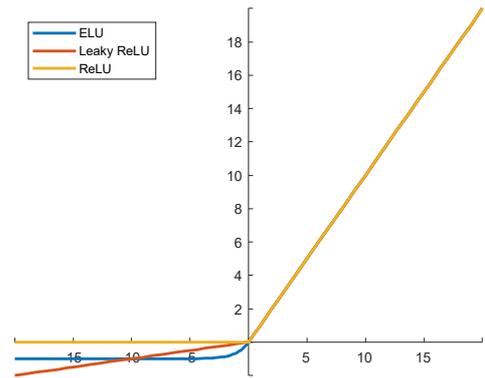


Fig. 4 Activation function graph for ELU, Leaky ReLU and ReLU.

4) Softmax: This function or classifier is a class-based prediction operation to find the probability distribution (PD) scores of the ' k ' targets i.e. the number of output classes. Hence the final classification layer uses the softmax function to predict the final class of the input MRI image. For $i=I$,

2... k number of classes with an input feature vector x_i , the i^{th} probability score p_i is shown in Equation (4)

$$p_i = \frac{e^{x_i}}{\sum_1^k e^{x_k}} \quad (4)$$

Here, p_i being a value between 0 to 1; hence the i^{th} class with maximum probability score wins the race [17]. One of the problems using the 2D CNN is in the selection of the appropriate slice/slices along with its orientation as training inputs i.e. to select in the axial, sagittal, or coronal axis. Recent literature proposes the ‘best scan’ or the ‘best multiple slices’ [22]-[25] for an effectual performance however, this makes the region of interest (ROI) area and patch selection process more indeterminate. It becomes difficult and unfeasible every time. We might lose some important information if we emphasize only specific scans or the orientation. Therefore, the safest and the best tactic to accommodate all the pixel information can be using the all brain slices or whole MRI volume. This comes with 3D values (i.e. pixel values for the x, y, and z dimensions in a planar geometry). Furthermore, the process of choice of slice/slices is still ambiguous. In comparison to 2D, 3D CNN has an extra depth feature extraction capability because of its 3rd dimension, which makes the convolution operation more computation. The addition of the depth or channel in the 3rd dimension helps to accumulate the feature along the x, y, and z dimensions. The used equation is as shown in (5).

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv_{.3} (w_{ik}^{l-1}, s_i^{l-1}) \quad (5)$$

where $conv_{.3}$ is a fixed 3-D convolution i.e. $N \times N \times N$ without zero paddings on the edges. In reference to equation (5), x_k^l is the input b_k^l is the bias of the k^{th} neuron at layer l , and s_i^{l-1} is the output of the i^{th} neuron at layer $l-1$. w_{ik}^{l-1} is the kernel (weight) from the i^{th} neuron at layer $l-1$ to the k^{th} neuron at layer l .

TABLE I
CNN ARCHITECTURE LAYERS

Layers	Specification
Image Input	64×64×64×1 images with 'zerocenter' normalization
Convolution	16 3×3×3×1 convolutions with stride [1 1 1] and padding 'same'
Batch Normalisation	Batch normalization layer for 16 channels
Activation	ReLU
Pooling	2×2×2 max pooling with stride [2 2 2] and padding [0 0 0; 0 0 0]
Convolution	32 3×3×3×16 convolutions with stride [1 1 1] and padding 'same'
Activation	Batch normalization with 32 channels
Pooling	2×2×2 max pooling with stride [2 2 2] and padding [0 0 0; 0 0 0]
Convolution	32 3×3×3×32 convolutions with stride [1 1 1] and padding 'same'
Batch Normalisation	Batch normalization layer for 32 channels

Activation	ReLU
Pooling	2×2×2 max pooling with stride [3 3 3] and padding [0 0 0; 0 0 0]
Convolution	64 3×3×3×32 convolutions with stride [1 1 1] and padding 'same'
Activation	ReLU
Pooling	2×2×2 max pooling with stride [3 3 3] and padding [0 0 0; 0 0 0]
Fully connected	1000 fully connected layer
Dropout	50% dropout
Fully connected	1000 fully connected layer
Activation	ReLU
Fully connected	3 fully connected layer
Softmax	softmax
Classification output	crossentropy with 'ADD', 'mAD' and 'HC' classes

III. RESULT AND DISCUSSION

CNN was designed with architecture resemblance to U-net [9] encoder architecture with necessary modification as listed layer-wise with details in Table I. Only 2 layers of encoder were used, to obtain the final features, supported by the fully connected layers (FCLs) and subsequently with a softmax classification layer for getting the performance result. Here, another important factor is in the selection of 3D max-pool and Batch normalization layer, in between each encoder layer for normalization of each weight within a fixed scale. ReLU has been used as an activation function to pass only non-negative values. Once the network is trained using 50% of data and 20% for validation, the remaining is used for testing. The obtained result is reported in Table III. The result is from a random ratio, so we have reported the best accuracy from 5 consecutive experiments. The low accuracy may be due to the use of less training MRI, as deep learning model performance heavily depends on the number of training material.

TABLE II
HYPERPARAMETER AND TRAINING CONDITION FOR CNN

Hyperparameter and Training condition	Selected
Training optimization function	Adam
Mini-batch size	8
Maximum epoch	100
Gradient Threshold method	$l-2$ normalization
Initial Learn Rate	1e-3
Learn Rate Drop Factor	0.95
Validation Frequency	50
Number of Iterations	700
Learn Rate Drop Period	10
Weight initializer	Glorot
Training:Validation:Test ratio	5:2:3
Gradient decay Factor	0.9000

IV. CONCLUSION

Test on other bulkier datasets like ADNI [10] is also in progress but not clear remarks could be found till now. And then we are looking more importantly to develop a general architecture that can work in almost all types of MRI, unlike its origin or obtained procedure. The used hyperparameters are tabulated in Table II. To test the effect of a wider architecture we tested our recently proposed architecture [14]. The experiment was re-run using the ‘divNet’ architecture as proposed in [14], we found out the accuracy of NRCG_Grey_MRI vs. NRCG_nifti_MRI to be around 42.31% and 40.5% respectively. All experiments were simulated on MATLAB R2019a academic software and the hardware is NVIDIA GeForce RTX 2070 GPU with 24 GB RAM. Network models were trained on GPU whereas the trained model was tested in Intel® Core™ i5-9600K CPU operating at @ 3.70 GHz frequency with 32 GB memory.

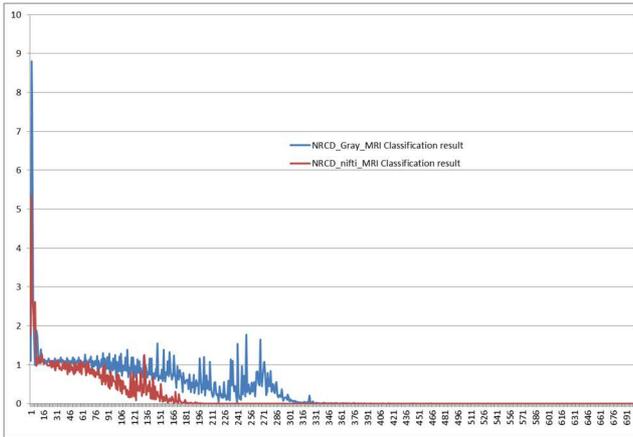


Fig. 5. Training loss (y-axis) is plotted against each iteration (x-axis). The red curve is the loss for the original MRI whereas the blue curve represents the loss of its grey version.

TABLE III
RESULT OF CLASSIFICATION FROM ORIGINAL MRI AGAINST ITS GREY VERSION

NRCG_nifti_MRI Classification result	NRCG_Grey_MRI Classification result
K: 3	K: 3
N: 37	N: 38
Degrees of freedom: 4	Degrees of freedom: 4
Random Accuracy: 0.4054	Random Accuracy: 0.4211
Fall: 0.4235	Fall: 0.3933
Gall: 0.4967	Gall: 0.4248
Cohen-Kappa: 0.0915	Cohen-Kappa: 0.1264
Informedness: 0.1737	Informedness: 0.1147
Markedness: 0.0988	Markedness: 0.1617
correlation: 0.1310	correlation: 0.1362
Loglikelihood: 1.4425	Loglikelihood: 2.4423
Mutual-Information: 0.0390	Mutual-Information: 0.0643
pearsonXsq: 1.2315	pearsonXsq: 1.6216

To conclude, we have performed an initial test of whether grey matter content MRI volume is the efficient training material for deep layered CNN or not? Detail feature extraction and analysis are still under the subject of study. Although the overall classification result is not very high which may be due to the use of limited training materials. As deep learning networks are data-hungry network, which highly depends on the quantity and quality of its training material for good performance. However, from Table III, it is clear that the accuracy from the Grey version is increased by almost 2-3% than its MRI version when we conduct the test in a similar environment. Besides, other performance metrics like Cohen-Kappa, Informedness, Mutual information, etc. are also included in Table III. The training loss is relatively shown between the two versions in Fig. 5.

The obtained result supports our idea, of using the grey version for a better classification result. Although during random test sometimes the result is not supporting, however in average, the result supports our idea. One of the drawbacks of our proposed method is we need to perform a manual segmentation task of each MRI using an additional tool of SPM. Extra effort and time are required for the segmentation process, so in general, this idea is quite a tedious process. However, if we can combine the segmentation and smoothing algorithm along with the classification task in a single CNN, this might be helpful in a more sophisticated way. As of now, we are only showing the grey matter version is helpful in classification tasks between ADD, mAD, and HC.

Regarding future works, we are working to develop better and optimized CNN models for MRI classification with a higher performance ratio. Deep neural networks have certain limitations and weaknesses. Like it is easily prone to overfitting and lacks generalization. It means the ratio of correct classification of images with different features like orientation, color difference, contrast difference is comparatively low so we are working to reduce this generalization problem. Besides, the used dataset in this study is limited, so we plan to test our idea in a bigger sample size and other multiple sources available for public use.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No. NRF-2019R1A4A1029769). And this study was supported by research funds from Chosun University, 2021.

Data collection and sharing for this project was funded by the ADNI (National Institutes of Health grant number U01 AG024904) and the DOD ADNI (Department of Defense grant number W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and the generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen

Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research provides funding to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The Northern California Institute for Research and Education is the grantee organization, and the study was coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuroimaging at the University of Southern California.

REFERENCES

- [1] J. L. Tanabe, D. Amend, N. Schuff, V. DiSclafani, F. Ezekiel, D. Norman, G. Fein, and M. W. Weiner, "Tissue segmentation of the brain in Alzheimer disease", *American Journal of Neuroradiology*, 1997; 18(1):115-123.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", *Proc. NIPS*, 2012; 1097-1105.
- [3] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions", *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2015; 1-9.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition" *arXiv:1512.03385*, 2015
- [5] Girshick, Ross, "Fast r-cnn," In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448. 2015.
- [6] National Research Centre for Dementia, Korea <http://www.nrcd.re.kr/>
- [7] J. Ashburner, "A fast diffeomorphic image registration algorithm", *NeuroImage*, Volume 38, Issue 1, 2007, Pages 95-113, ISSN1053-8119, <https://doi.org/10.1016/j.neuroimage.2007.07.007>
- [8] Mazziotta J, Toga A, Evans A, et al, "A probabilistic atlas and reference system for the human brain". *International Consortium for Brain Mapping (ICBM), Philosophical Transactions of the Royal Society of London Series B*. 2001; 356(1412):1293-1322. doi:10.1098/rstb.2001.0915.
- [9] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T. and Ronneberger, O., 2016, October. "3D U-Net: learning dense volumetric segmentation from sparse annotation". In *International conference on medical image computing and computer-assisted intervention* (pp. 424-432). Springer, Cham.
- [10] Alzheimer's Disease Neuroimaging Initiative: ADNI www.loni.ucla.edu/ADNI.
- [11] Y. Zhang, S. Wang, G. Ji, and Z. Dong "Exponential wavelet iterative shrinkage thresholding algorithm with random shift for compressed sensing magnetic resonance imaging," *IEEE Transactions on Electrical and Electronic Engineering*. 2015; 10(1):116-117.
- [12] Z. Dong, Y. Zhang, F. Liu, Y. Duan, A. Kangarlu, B. S. Peterson, "Improving the spectral resolution and spectral fitting of 1H MRSI data from human calf muscle by the spread technique," *NMR in Biomedicine*. 2014; 27(11):1325-1332
- [13] J. Ashburner, "A fast diffeomorphic image registration algorithm", *NeuroImage*, Volume 38, Issue 1, 2007, Pages 95-113, ISSN1053-8119, <https://doi.org/10.1016/j.neuroimage.2007.07.007>
- [14] Khagi, B. and Kwon, G.R., 2020. 3D CNN design for the classification of Alzheimer's disease using brain MRI and PET. *IEEE Access*.
- [15] Nagi, J., Ducatelle, F., Di Caro, G.A., Cireşan, D., Meier, U., Giusti, A., Nagi, F., Schmidhuber, J. and Gambardella, L.M., 2011, November. "Max-pooling convolutional neural networks for vision-based hand gesture recognition". In *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)* (pp. 342-347). IEEE.
- [16] Nair, V. and Hinton, G.E., 2010, January. "Rectified linear units improve restricted Boltzmann machines". In *ICML*. Bishop, C.M., 2006. *Pattern recognition and machine learning*. Springer.
- [17] Payan, Adrien, and Giovanni Montana. "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks." *arXiv preprint arXiv:1502.02506* (2015).
- [18] Hosseini-Asl, Ehsan, Georgy Gimel'farb, and Ayman El-Baz. "Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network." *arXiv preprint arXiv:1607.00556* (2016).
- [19] Oh, Kanghan, Young-Chul Chung, Ko Woon Kim, Woo-Sung Kim, and Il-Seok Oh. "Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning." *Scientific Reports* 9, no. 1 (2019): 1-16.
- [20] Liu, M., Cheng, D., Wang, K., & Wang, Y. (2018). Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. *Neuroinformatics*, 16(3-4), 295-308. doi:10.1007/s12021-018-9370-4.
- [21] Yang G, Zhang Y, Yang J, et al. Automated classification of brain images using wavelet-energy and biogeography-based optimization. *Multimed Tools Appl*. 2015;75(23):15601-15617.
- [22] Jha, Debesh, Ji-In Kim, and Goo-Rak Kwon. "Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network." *Journal of healthcare engineering* 2017 (2017).
- [23] Khagi, Bijen, 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, no. 3 (2019): 297-310.
- [24] Wang SH, Zhang Y, Li YJ, Jia WJ, Liu FY, Yang MM. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimed Tools Appl*. 2016;77:1-25
- [25] Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." In *Proc. icml*, vol. 30, no. 1, p. 3. 2013.