2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Unsupervised detection of individual atrophy in Alzheimer's disease

Shichen Jin, Peini Zou, Ying Han, Jiehui Jiang* and the Alzheimer's Disease Neuroimaging Initiative

Abstract—**Background: To realize precision medicine, it is important to realize the detection of the individual atrophy of Alzheimer's disease (AD) patients. Our objective is to find individual brain regions of interest (ROIs) in AD patients via an unsupervised deep learning network.**

Methods: **This study used structural Magnetic Resonance Imaging (sMRI) scans with the 732 healthy control (HC) subjects and 202 AD patients from the Alzheimer's disease Neuroimaging Initiative (ADNI), and the 105 HC subjects were collected at the Xuanwu Hospital. An unsupervised deep learning network based on Adversarial Autoencoders (AAE) was proposed to delineate the individual atrophy of AD patients. In the proposed model, Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN) were combined to learn the potential distribution and train a generator. In this step, the 530 HCs from ADNI were applied as the training dataset and the 105 HCs from Xuanwu Hospital were applied as an external validation dataset. The structural similarity (SSIM) was used to judge the robustness of the proposed model. Then, ROIs of the 202 AD patients were detected. In order to verify the clinical performance of these ROIs, other 202 HCs were selected from ADNI and a multilayer perceptron (MLP) was used to classify AD versus HC by 5 folder cross-validation. In the comparative experiments, we compared our model with three other previous models.**

Results: **The SSIM reached 0.86 in both training and external validation datasets. Eventually, the classification accuracy of our model achieved 0.94±0.02. In the meanwhile, the classification accuracies were 0.89±0.01, 0.85±0.04 and 0.91±0.03 for the three previous methods.**

Conclusion**: Our deep learning model could detect individual atrophy in AD patients. It may be a useful tool for AD diagnosis in clinics.**

Keywords—**Atrophy detection, Alzheimer's disease, Magnetic Resonance Imaging, Adversarial autoencoders, Multilayer perceptron**

I. INTRODUCTION

Alzheimer's disease (AD) is the most common neurodegenerative disease in the world. It is primarily characterized by progressive memory loss and accompanied by several kinds of cognitive and functional impairment. As an important clinical tool, magnetic resonance imaging (MRI) has been used to reflect the brain structure and characteristic cerebral changes noted in AD.

AD patients are with individual variations, therefore it is important to acquire individual atrophy morphometry information. Currently, a few MRI imaging markers were proposed. One example is hippocampus atrophy. The hippocampus volume is the gold standard to detect atrophy, but it is costly and time-consuming because of the necessity of professional physicians and manual selection of the region of interest (ROI)(1). Alternatively, automatic segmentation tools such as voxel-based morphometry (VBM)(2) were required(3). However, these methods were based on group analysis or the templates, which made it difficult to identify the precise individual atrophy.

Recently, to realize the detection of individual atrophy, machine learning and deep learning methods have been applied. For instance, Raouia Ayachi and Nahla Ben Amor et al., detected tumors by Support Vector Machines (SVM)(4). D. Zikic used context-sensitive classification forests(5). Moreover, many popular deep learning models such as Generative Adversarial Networks (GAN) were also applied to segment tumors in brain MRI images (6). However, previous methods could hardly be applied to atrophy detection. To our knowledge, studies on detecting individual atrophy in AD patients are still immature. Therefore, this study proposed a novel unsupervised deep learning model based on Adversarial Autoencoders (AAE)(7). This novel model was trained by MRI data of healthy controls (HCs). Using the MRI data of AD patients as the inputs, our model reconstructed the MRI images, and automatically calculated the residual images between the original and reconstructed images. The residual images were considered as the individual atrophy ROIs. In order to verify the diagnosis performance of the individual atrophy, we generated a new residual mask as the Regions of Interest (ROI), extracted ROI images from the AD and HC groups, and calculated the classification accuracy of ROI images. We also compared the classification results of our method with three previous methods.

II. MATERIALS AND METHODS

A. participants and preprocessing of sMRI images

The data of this study were from the Alzheimer's disease Neuroimaging Initiative (ADNI) and Xuanwu Hospital, Capital Medical University, Beijing, China. ADNI is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early prediction and tracking of Alzheimer's disease. The sMRI images of 530 HCs from ADNI were used to train the AAE model and the 105 HCs from Xuanwu Hospital were served as an external validation group. The 202 AD subjects and matched number of HC subjects randomly selected from ADNI were used in the classification task. Tab. 1 showed the demographics and characteristics of all data. The data in column 1 and column 2 are used for model training and validation while the rest from ADNI are applied in classification. There were significant differences between HC and AD patients in age, education, Mini-mental State Examination (MMSE) and APOE4 gene information ($p<0.005$)

Shichen Jin, Peini Zou Jiehui Jiang are with the Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, China (corresponding author to provide phone: +86 139-1892-0926; e-mail: jiangjiehui@shu.edu.cn).

Ying Han is with Center of Alzheimer's Disease, Institute for Brain Disorders, Beijing, China.

Table 1. DEMOGRAPHICS AND CHARACTERISTICS

	ADNI	Xuanwu	ADNI		
	HC (n=530)	HC (n=105)	AD (n=202)	HC (n=202)	P-value
Age	73.99±1.10	66.75±4.90	74.75±8.15	73.99±6.18	<0.005
Sex	281/249	69/36	123/77	115/87	0.125
Education	16.38±2.73	12.28±3.12	15.58±2.77	16.24±2.66	<0.005
MMSE	29.10±1.10	28.75±1.5	23±2.23	29.10±1.12	<0.005
APOE4	160(30.21%)	31(29.52%)	135(66.85%)	71(35.14%)	<0.005

Note: Age and MMSE are given as mean ± standard deviation, Gender is given as Female/Male, APOE4 is given as the number of carriers (percentage). An independent two-sample two-tailed t-test is conducted for quantitative variables and the variables are maintained consistent.

The preprocessing of sMRI images mainly included the following steps: 1. We converted the DICOM format into a 3D image in the Neuroimaging Informatics Technology Initiative (NIfTI) format using the DCM2NII (https://people.cas.sc.edu/ rorden/mricron/dcm2nii.html) tool; 2. All volumes were spatially normalized to the Montreal Neurological Institute(MNI) template by Statistical Parametric Mappin (SPM8; https://www.fil.ion.ucl.ac.uk/ spm/ software/spm8) voxel size is 2x2x2; 3. The whole brain template was used to eliminate the interference of noise in the background 4. To accelerate the model solution, the voxel values of the images were normalized into -1 to 1; 5. The three-dimensional images were sliced from the axial direction and interpolated to a size of 128*128 in order to meet the input requirements of the deep learning model.

B. *Framework of this study*

Fig.1 showed the framework of this study. The left part showed the schematic of the deep learning framework. The reconstructed images and residual images could be automatically generated as the outputs. The right part showed the structure of the classification model. With the generated 'residual mask', the pixels in ROIs were selected to feed into the Multilayer perceptron (MLP) classifier for the classification task.

C. *Deep learning model*

The sMRI images of 530 HCs from ADNI were used to train the AAE model. For the characteristics of AD atrophy, the unsupervised deep learning model was realized by two stages. In the first stage, an AAE network was trained to extract the data characteristics of HC images $x_h \in X^{d*w*h}$, which was represented as a predetermined standard distribution P(z). Then, a reverse process was applied to decode the latent distribution with its original image. The robustness of this step can be judged by the loss of model and structural similarity (SSIM), which is an index to measure the similarity of the generated image and its input image. In the second stage, images of patients with AD were fed into the model. For that our model was only trained with HC images, the AD images were reconstructed as HC images. As a result, the individual atrophy could be generated by calculating the residual images between the original and reconstructed images.

Deep learning model construction. The AAE network used in this method consisted of one GAN and one autoencoder network(8). This network was an improvement of traditional Variational Auto-encoder (VAE). The formula of optimization can be expressed as_(1)

$$\min_G \max_D V(D,G) = \mathrm{E}_{z\sim P(z)}[\log \mathrm{D}(z)] + \mathrm{E}_{x\sim P(x_h)}\left[\log\left(1-\mathrm{D}(\mathrm{G}(x))\right)\right] \quad (1)$$

We used Jensen-Shannon (JS) divergence to align the implicit variable z to prior P(z) rather than traditional Kullback-Leibler (KL) divergence. JS divergence was useful to solve the primary problems of VAE, such as sparse manifolds or coarse reconstructions after dimension reduction.

As shown in Fig.1, the GAN consisted of two parts: a generative model G, and a discriminative model D. G was used to approach the real data with a lower dimension latent code z and D was used for distinguishing the generated data from the real data. This employment reached the state of "Nash balance" in the process of alternative optimization.

To improve the poor expression ability of the original network, the Resnet structure was applied to the generator in our model. The encoder and decoder both consisted of one convolution layer and five Resnet blocks. Additionally, the usage of the Batch Normalization (BN) layer solved the problems of gradient disappearance and gradient explosion. Then, an additional regularization was proposed with an extra loss shown as $|z_o - z_g|^2$ to keep the generator restoring the original features as more as possible. In the training step, the learning rate was set as 0.0002, the batch size was 32. Furthermore, the t-distributed Stochastic Neighbor Embedding (TSNE) distribution was used so that the distribution relationship between the generated image and the input image could be roughly observed.

Validation of the deep learning model. To test the robustness of our model, MRI images of 105 HC subjects from Xuanwu hospital were used as an external test group. The mean SSIM was calculated as the index to judge the effectivity of our model. Both SSIM values from the training group and test group were calculated in this study with 5 folder cross-validation in1000 times.

Visualization of residual images for AD patients. To visualize residual images and remove noise, the density-based spatial clustering of applications with noise (DBSCAN) cluster method was applied, in which the parameters 'Eps'，'MinPts' were set as 4 and 10(9).

Residual mask generation. Using the deep learning model, each corresponding individual atrophy can be obtained by calculating the residual images between the original and reconstructed images. To verify the clinical value of these individual atrophy images, we generated a residual mask and extracted ROI images for classification: Firstly, a universally residual mask was generated by combing pixels that appeared in more than 75 percent residual images of AD patients; Secondly, we used the residual mask as an ROI, and extracted ROI images from 202 ADs and 202 HCs for further study.

D. *Classification experiments*

With ROI images from the above step, we used multilayer perceptron (MLP) as the classifier. MLP is a classical and simple neural network. The ROIs from 202 AD and 202 HC subjects served as the inputs of MLP. The 5-fold cross-validation was applied to ensure the rationality of the results.

To verify the clinical performance of these ROIs, we compared our model with three previous models: (1) principal component analysis (PCA) combined with support

vector machines (SVM) based on the whole brain images; (2) The traditional CNN model based on the whole brain images; (3) we extracted hippocampus regions with the Anatomical Automatic Labeling (AAL) template, used pixels of hippocampus regions as inputs, and trained an MLP classifier. All experiments were done with a 5-folder cross-validation and repeated experiment 200 times. The accuracy, sensitivity, and specificity were calculated to measure the classification performance.

E. *Statistical analysis*

In the project, all statistical analyses were performed in SPSS Version 22.0 software (SPSS Inc., Chicago, IL). For example, a two-sample *t*-test was used to compare demographic characteristics between AD and HC groups. The P-value less than 0.05 was considered significant.

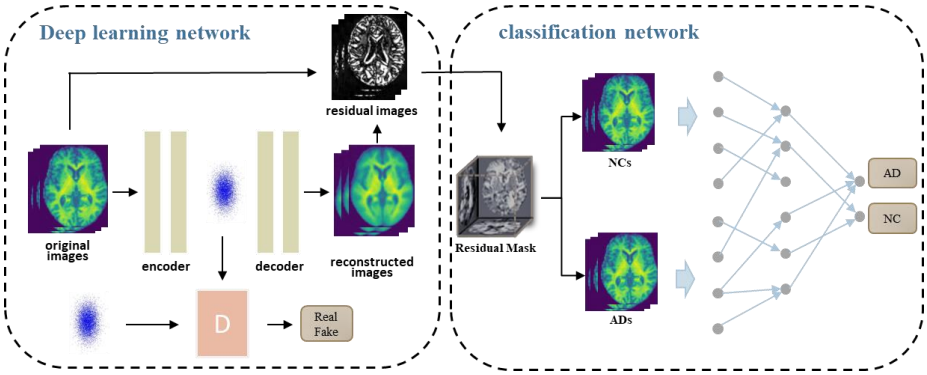

Figure 1. The framework of this study.

III. RESULTS

A. *Results of deep learning model*

Deep learning model construction. The advantage of our deep learning model was that it could generate the latent distribution map of features and then gained residual images of AD patients. To prove this, all features after TSNE embedding, including features of original HC and AD images, and generated HC and AD images were visualized in Fig.2. As a result, original AD and HC slices were widely distributed for large intra-group variability. These distributions were not concentrated at one location but were slightly linear. This was because the slice distribution at disparate positions was relatively different when the difference between adjacent slices was not significant. As seen in Fig.2A and Fig.2B, the image distributions generated by the HC group and AD group were close, which was consistent with our previous hypothesis. Fig.2C showed the situation that all distributions were focused, indicated that related slices were not the individual ROIs. This dimensionality reduction diagram also proved that the feature distribution in this experiment was reasonable.

Visualization of residual images for AD patients. Fig.3 showed one example of residual images from one AD patient. This patient was 68 years old; APOE4 was positive, and MMSE value was 20. As shown in Fig.3, the atrophy marked in the residual image was mainly concentrated in the hippocampus region, which was consistent with previous research on AD. The result showed that our deep learning model could be effective.

Validation of the deep learning model. After 120000 times training of our deep learning model, the loss of the anomaly prediction model tended to be stable while the performance on validation was consistent with the training set without overfitting. Finally, SSIM reached 0.86 in both ADNI and Xuanwu's datasets.

Residual mask generation. Fig.4 presented the residual mask. As it shows, main ROIs were found in the hippocampus region, which was in line with current research results.

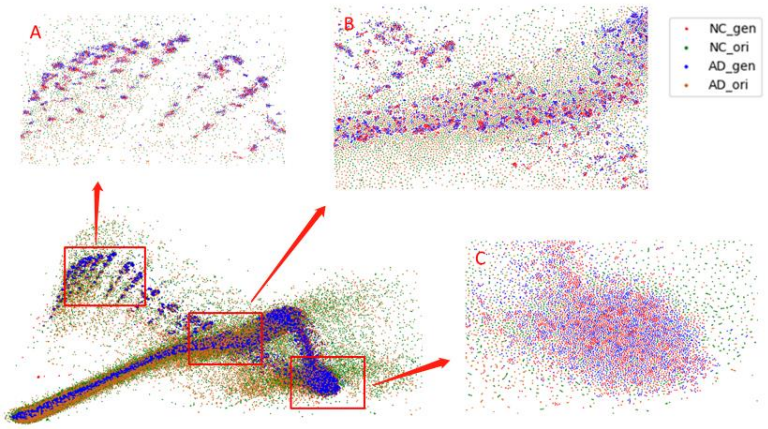

Figure 2. Image shows TSNE embedding slices of original HC images (green), original AD images (chocolate), generated HC images (red) and generated AD images(blue)

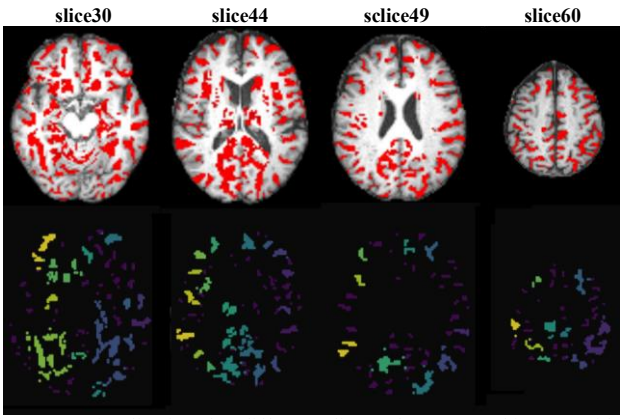

Figure 3. Visualization for different slices of an AD subject's residual images. row 1 is the residual images from the view of axial and is sectioned at 30,44,49, and 60 respectively. The hippocampus is shown at slices of 44 and 49. Row 2 is gained from row 1 after threshold selection. Row 3 is marked by DBSCAN cluster, showing the obvious focused location of difference for effective observation.

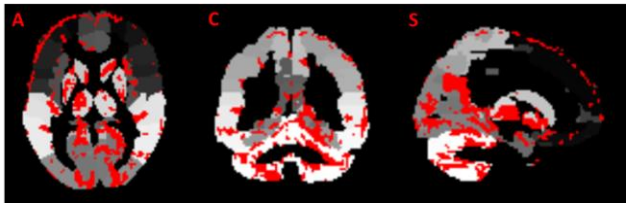

Figure 4. The 'residual mask' from the view of Axial, Coronal, and Sagittal is marked with 'A',' C','S' with the AAL template as background. The red part indicated the selected pixels that more than 70 percent of AD subjects appeared in the residual images.

B. Classification Results

Tab.2 showed the classification results of our proposed model and the three previous models. As shown in Tab.2, our model achieved the best accuracy (0.94 ± 0.02), sensitivity(0.99±0.02) and specificity (0.94±0.05), while the results for PCA+SVM (accuracy 0.89±0.01, sensitivity 0.92 ± 0.03, 0.87 ± 0.04), AAL+MLP (accuracy 0.85 ± 0.04, sensitivity 0.87 ± 0.05, 0.81 ± 0.18) and CNN(accuracy 0.91±0.03, sensitivity 0.95±0.06, 0.87±0.08) were worse. This result showed that the residual mask had better clinical diagnosis performance than the whole brain or certain prior knowledge ROI (hippocampus region), which indicated that the individual atrophy generated from our deep learning model may have good diagnosis performance.

Table 2. PERFORMANCE Of DIFFERENCE CLASSIFICATION APPROACHES

	Accuracy	sensitivity	specificity
PCA+SVM	0.89±0.01	0.92±0.03	0.87±0.04
AAL+MLP	0.85±0.04	0.87±0.05	0.81±0.18
CNN	0.91±0.03	0.95±0.06	0.87±0.08
Our method	**0.94±0.02**	**0.99±0.02**	**0.94±0.05**

Note：he methods are conducted with cross-validation and their results are given as mean ± standard deviation.

IV. DISCUSSION

This paper proposed an unsupervised deep learning model utilizing a prevailing model AAE to detect individual atrophy. The classification experiment results indicated the diagnosis value of individual atrophy generated from our model, and therefore the proposed model had the potential to be developed as a novel diagnostic tool in the future.

Besides, when we visualized the individual atrophy results of AD patients, such as in Fig.3, it can be observed that the results were generally consistent with previous research(3, 10). That was to say, the model indeed had the ability to identify the differences between the images of AD patients and their reconstructed HC images to some extent. By virtue of this characteristic, we can apply the model to many scenarios, even can track conspicuous brain changes by other neurodegenerative brain diseases.

It was worth noting that this study had some limitations. Firstly, the current deep learning model was not accurate enough to restore the HC images without multiscale prediction, including features of low and high scale such as the skip connection in U-Net (11). Secondly, the classification results in this study were limited. We now only compared the classification results with the AD and HC groups and in our next work, our model would be applied in other groups such as mild cognitive impairment(MCI) and subjective cognitive impairment patients(SCD). Lastly, the dataset used may not be large enough. Thus the results of this study need to be further verified by other datasets.

REFERENCES

[1] Pini L, Pievani M, Bocchetta M, Altomare D, Bosco P, Cavedo E, et al, "Brain atrophy in Alzheimer's disease and aging," Ageing Research Reviews, 2016.

[2] Ashburner, J, "A fast diffeomorphic image registration algorithm. " Neuroimage, 2007, pp. 38(1):95-113

[3] Aisen PS, Petersen RC, Donohue MC, Gamst A, Raman R, Thomas RG, et al, "Clinical Core of the Alzheimer's Disease Neuroimaging Initiative: progress and plans," 2010;6(3):239-46.

[4] Ayachi R, Amor NB, editors, "Brain Tumor Segmentation Using Support Vector Machines," European Conference on Symbolic & Quantitative Approaches to Reasoning with Uncertainty, 2009.

[5] Zikic D, Glocker B, Konukoglu E, Shotton J, Price SJ, editors, "Context-sensitive Classification Forests for Segmentation of Brain Tumor Tissues," Miccai Brats Challenge, 2012.

[6] Han C, Murao K, Noguchi T, Kawata Y, Uchiyama F, Rundo L, et al, "Learning More with Less: Conditional PGGAN-based Data Augmentation for Brain Metastases Detection Using Highly-Rough Annotation on MR Images," Proceedings of the 28th ACM International Conference on Information and Knowledge Management; Beijing, China: Association for Computing Machinery, 2019, pp. 119–27.

[7] Chen X, Konukoglu EJapa, "Unsupervised detection of lesions in brain mri using constrained adversarial auto-encoders," 2018.

[8] Makhzani A, Shlens J, Jaitly N, Goodfellow I, Frey BJCE, "Adversarial Autoencoders," 2015.

[9] Bandyopadhyay SK, Paul TUJIJoRiE, Science, "Segmentation of brain tumour from mri image analysis of k-means and dbscan clustering," 2013, pp. 1(1):48-57.

[10] Kaufer DI, Cummings JL, Ketchel P, Smith V, Macmillan A, Shelley T, et al. "Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory". 2000, pp. 12(2):233.

[11] Ronneberger O, Fischer P, Brox T, editors, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Springer, Cham, 2015.