

Automated brain atrophy quantification and evaluation using spatial resolution enhancement

Yonglai Zuo, Xiaohan Hao, Mengdie Song, Fulang Qi, Bensheng Qiu* and Xiaoxiao Wang*

Center for Biomedical Imaging
University of Science and Technology of China
Hefei, China

*Email: {bqiu, wang506}@ustc.edu.cn

Abstract—Brain atrophy is one of the most common features of neurodegenerative diseases and is particularly critical in the early diagnosis of conditions like Alzheimer’s and multiple sclerosis. Automated segmentation and quantification are highly desirable in brain atrophy evaluation but existing methods require high-quality MRI scans with isotropic resolution. However in practice, clinicians usually choose to reduce the number of slices to save time, and because of their anisotropic resolution, the morphometric analysis cannot be implemented. Here we propose the use of an inter-slice interpolation network to increase spatial resolution to isotropic to accomplish morphometric analysis of brain atrophy, which is not feasible to perform with clinical scans. To achieve fully automated quantification, we set up a series of indicators that cover multiple aspects such as holistic and localized, structural and cortical, based on our clinical experience. As a result, a fully automated diagnostic process for pathological brain atrophy is completed. Our results on the IXI and ADNI datasets show that the slice spacing method produces realistic and reliable anatomy and achieves high diagnostic accuracy, and our indicators are well suited to clinical application scenarios.

Keywords—medical imaging, structural MRI, brain atrophy, super-resolution, 3D CNN

I. INTRODUCTION

Brain atrophy, a phenomenon in which the brain tissue itself undergoes organic lesions leading to atrophy, is considered one of the most important biomarkers of neurodegenerative disease [1] and is critical to the diagnosis of a wide range of conditions such as Alzheimer’s disease (AD). An accurate and timely diagnosis of pathological atrophy can give patients the best chance to develop a treatment plan that will improve disease progression and reduce symptoms [2]. Open-source neuroimaging software, such as FreeSurfer[3], enables researchers to automate the way of conducting brain studies. The automatic analysis tools usually require scans with high isotropic resolution (1mm) to reconstruct accurate 3d volumes [4] and obtain accurate and trustworthy analysis results. However, the scanning time in practice is very lengthy when the voxel is set to 1 mm, and physicians usually opt for 2d scanning sequences to obtain images with low through-plane resolution and high in-plane resolution, which typically have less than 30 slices. The adoption of this scanning pattern has reduced time cost and improved patient comfort but has also resulted in most clinical scans being unavailable for morphologic measurements of brain atrophy [5].

In recent years, deep learning has shown great potential for applications in medical image analysis [6]. It opens up the

possibility of utilizing widely available clinical MRI resources for large-scale morphology-based disease screening and risk prediction. Zhang [7] addressed the problem of limited brain data using a data augmentation strategy and proposed a new CNN classification model. Zhang [8] came up with a patch-based classification network for discriminative atrophy localization and AD classification instead of widely used voxel-based methods, which significantly reduced the number of patches required for achieving a good classification performance. Silva [9] proposed a Biomechanical Network that models the effects of age, disease status, and scan interval to regress regional patterns of atrophy, hence effectively integrating clinical information. Laso [10] presented a new tool to quantify brain volumes directly using anisotropy images and a specialized segmentation method to segment these anisotropy scans. However, there exist some limitations including: 1) Most of these methods use end-to-end learning, which lacks interpretable diagnostic information and specific quantification and, therefore is difficult to be accepted by clinicians. 2) They are not adapted to the characteristics of clinical scans, and cannot be implemented on clinical data, as a result, fail to comply with clinical procedures.

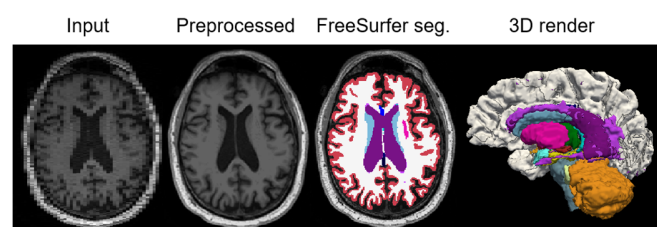


Fig. 1. Example of a T1 input and corresponding preprocess result and segmentation.

Here we present our automated diagnostic method for analyzing brain atrophy. We employ a spatial super-resolution method based on deep learning as preprocessing to enhance the resolution of the images to isotropy (1mm), facilitating morphometric measurements. Simultaneously, we quantify the degree of brain volumetric atrophy and cortical thickness atrophy by establishing a set of criteria. These criteria are then combined with clinical scenarios to generate the diagnostic results. Our method allows for the utilization of a large number of clinical scans in scientific analysis, filling the gap in clinical diagnostic criteria for brain atrophy and fully automating the entire process. To the best of our knowledge, we are the first to apply the spatial super-resolution method to the quantification and diagnosis of brain conditions. Our main contributions are as

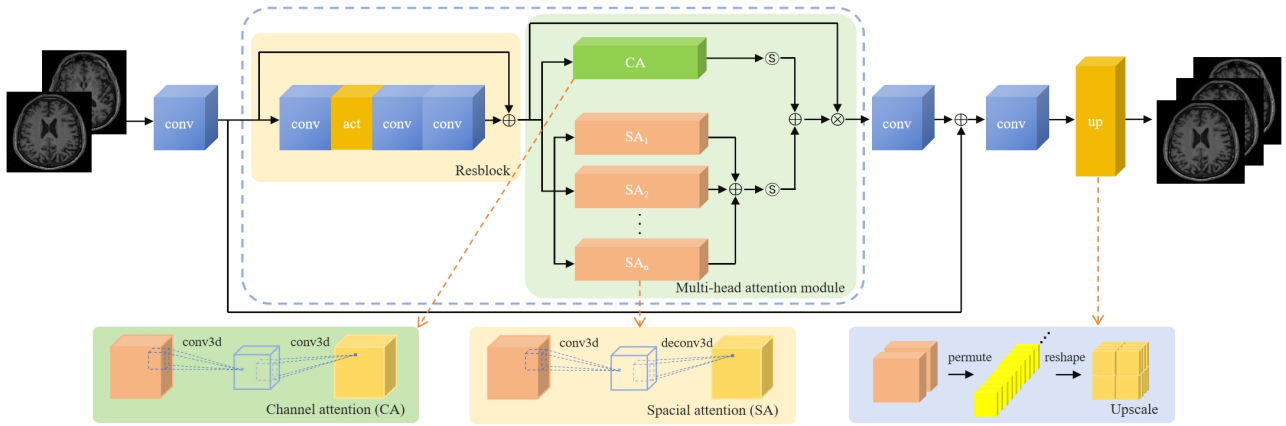


Fig. 2. The structure of the proposed preprocess network to yield isotropic scans.

follows:

1. A deep convolutional network is designed and implemented to enhance the spatial resolution of anisotropic scans to isotropic, which produces realistic and authentic results and addresses the problem that clinical scans cannot be used for research-grade analysis.

2. By focusing on key regions of interest (ROIs) and selecting the most representative ones, a series of brain atrophy quantification indicators are proposed for comprehensive and fine-grained assessment of pathological brain atrophy and cortical atrophy.

3. The proposed method has been evaluated and visualized using real datasets. The accuracy, sensitivity, and specificity of the biomarkers demonstrate that the evaluation results are accurate and our method improves the automated clinical diagnostic process, aiding doctors in clinical scenarios.

II. METHODS

We first need to address the low spatial resolution of conventional clinical scan sequences. Therefore, we have designed a dedicated deep-learning network for inter-slice reconstruction to bridge the resolution gap for research purposes. The result of the data preprocessing is illustrated in Fig. 1. Brain atrophy is a crucial biomarker, representing an essential feature across various diseases. To the best of our knowledge, no studies are focusing on measures of brain atrophy that distinguish it from other neurodegenerative diseases, so we have integrated the overall volume of ventricles, white matter, and smaller organs, to develop a series of quantitative criteria to improve the automated analysis process. We have also developed criteria for cortical atrophy based on important ROIs cortical thickness information, serving as valuable clinical reference points.

A. Spatial Super-resolution Network

Our super-resolution network is depicted in Fig. 2, with the aim of pre-enhancing images to research-level spatial resolution. The input comprises low anisotropic resolution clinical T1 scans to be enhanced. Input images undergo a module that includes a Resblock [11] cascading a multi-head attention module [12], and enhanced features are integrated through residual concatenation into original images. The image matrix is up-sampled to

isotropic resolution using a 3D upscale module we designed for a 3d matrix.

We splice multiple attention heads after repeated Resblocks to reinforce the extracted local features. Due to the complex structure of the human brain, there are few global feature patterns such as repetition, symmetry, isometric scaling, etc., so we only focus on different local scales. The attention module consists of a Channel Attention (CA) [13] module in parallel with multiple Spatial Attention (SA) [14] modules using a multi-path design, where the size of the convolution kernel in the n -th branch SA_n is $2n-1$. By changing the number of SA branches, the network can extract features from volumes of different sizes. Feature maps are multiplied with the residual path to apply the weights to the original image matrix. In the end, the feature maps are converted to slices with a 3D upscale block to increase to the specified resolution. We modified all the modules to adapt them to 3d images.

TABLE I. NETWORK PERFORMANCE AT DIFFERENT ENHANCEMENT SCALES

Voxel	Scale	PSNR	SSIM	RMSE
2mm	x2	49.06	0.9825	1.542
4mm	x4	43.64	0.9548	2.585
8mm	x8	40.35	0.9120	3.533

B. Quantitative Indicators for Brain Atrophy

Since the assessment of brain atrophy is a preliminary task for many disease diagnoses, there are currently no morphometric quantification methods for this specific task [15]. Compared with neurodegenerative diseases that focus on the cortex thickness, pathologic brain atrophy itself is more concerned with volume changes which necessitates a distinction between the two and a separate set of criteria for brain atrophy [16]. Structural volumetric atrophy refers to a reduction in the actual volume of brain regions, such as the hippocampus or the entire gray matter. This reduction indicates the real shrinkage or disappearance of brain tissue in three-dimensional space. Specifically, we propose using the following indicators:

- **Ventricle-to-brain ratio:** Indicates the proportion of the ventricles in the total brain structure, and their enlargement means shrinkage of brain volumes. A ventricles-to-brain volume ratio of roughly 25% is usually considered normal, while a ratio higher than 40% indicates the presence of brain atrophy.
- **Percentage reduction in volume of hippocampus, frontal lobe, and temporal lobe:** The Hippocampus, frontal lobe, and temporal lobe are common locations of atrophy in brain regions. A decrease in their volume of more than 15% relative to normal value can suggest pathologic atrophic changes.
- **Gray-to-white (G/W) ratio:** An increase in gray-to-white ratio may indicate a loss of white matter, which may be due to neurodegenerative diseases such as Alzheimer's disease or Parkinson's disease. Conversely, a diminished gray-to-white ratio might imply a loss of gray matter, possibly linked to factors such as brain injury or cerebral ischemia. The normal range of the gray-to-white ratio should be between 1.1 and 1.5 [17].
- **Highly contributing cortices include the entorhinal cortex, superior and inferior temporal cortex, hippocampal gyrus, anterior and posterior cingulate gyrus, and temporal pole.** These regions are involved in a variety of cognitive functions and are common sites of lesions in neurodegenerative diseases. A 30% reduction in cortical thickness in these areas and a 15% reduction in average thickness can be considered cortical atrophy.

Based on the Brodmann partition, we use the cortex at the above sensitive locations as ROIs. Each of these four indicators covers a different aspect and, when used individually, can each represent a certain category of atrophy. For example, the Ventricle-to-brain ratio can reflect lesions associated with cerebrospinal fluid (CSF), while reduced cortical thickness is a common clinical biomarker for AD. Given that the third and fourth indicators are based on a more in-depth assessment of the first and second indicators, serving as refinements and extensions, we determine that the patient has pathologic brain atrophy when any of the first two indicators and any of the last two indicators are satisfied simultaneously.

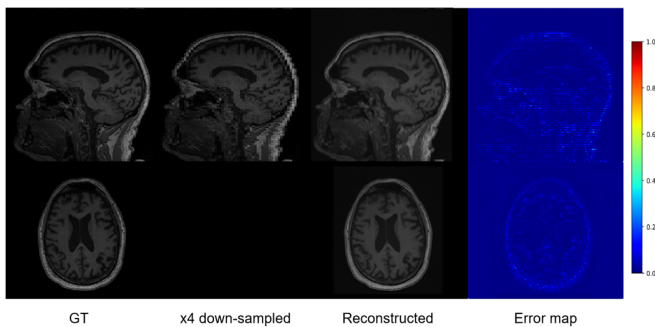


Fig. 3. The network's reconstruction result of a subject x4 down-sampled from axial direction and its corresponding ground truth.

C. Diagnostic Accuracy Metrics

We chose the following commonly used metrics to assess diagnostic performance: accuracy (ACC), sensitivity (SEN), and

specificity (SPE), with accuracy used as the primary metric. These metrics are as follows.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$SEN = \frac{TP}{TP + FN} \quad (2)$$

$$SPE = \frac{TN}{TN + FP} \quad (3)$$

where $TP = TruePositive$, $TN = TrueNegative$, $FP = FalsePositive$, $FN = FalseNegative$.

TABLE II. CORRELATIONS AND P-VALUES OF AUTOMATED ANALYSIS RESULTS WITH GROUND-TRUTH VOLUMES.

Voxel	Regions						
	W.m.	Vent.	Hipp.	Amyg.	Thal.	Put.	Pall.
2mm	0.99 (10^{-13})	0.99 (10^{-9})	0.97 (10^{-8})	0.96 (10^{-8})	0.94 (10^{-7})	0.94 (10^{-7})	0.94 (10^{-8})
4mm	0.97 (10^{-10})	0.98 (10^{-8})	0.94 (10^{-7})	0.95 (10^{-7})	0.88 (10^{-6})	0.94 (10^{-7})	0.90 (10^{-6})
8mm	0.88 (10^{-8})	0.92 (10^{-8})	0.56 (10^{-3})	0.74 (10^{-7})	0.91 (10^{-11})	0.77 (10^{-7})	0.61 (10^{-4})

III. EXPERIMENTAL EVALUATION

We used 581 3D structural MRI scans from the IXI dataset for training the spatial resolution enhancement network, along with 354 cases from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [18] to test its generalization and validate the classification performance of our indicators. We implemented inter-slice reconstruction on 178 patients, and 176 normal controls (NC), assessed the reconstruction through FreeSurfer analysis and conducted diagnoses based on our quantitative metrics.

A. Super-resolution Reconstruction Results

We set up three slice spacing scenarios to increase the 2mm, 4mm, and 8mm voxels to 1mm, i.e., to increase the spatial resolution by a factor of 2, 4, and 8 respectively to test the performance of our network (as shown in Table I). Employing commonly used image metrics such as PSNR, SSIM, and RMSE, we ensured accurate structural representation. Visualization of the outputs is shown in Fig. 3, and correlation tests were conducted to validate the realism and reliability of our reconstruction results in the next section. Compared to the ground truth (GT), errors come mainly from structures outside the brain, such as the skull, and the brain's anatomy is preserved intact. Our reconstructed images exhibit realistic and satisfying visual effects in both 2D and 3D. To simulate real clinical scenarios, we used x4 and x8 enhanced results for subsequent experiments, where x8 corresponds to scanning sequences with an initial 32-layer configuration.

B. Pearson Correlation Test of Reconstruction Results

After segmentation and measurement, we selected a series of representative volumes to demonstrate the results of our correlation test. As shown in Table II, anatomy was preserved

intact under x2 and x4 enhancements, achieving a general similarity of approximately 0.9 for each part. Even at x8 enhancement, the correlation between tissue classes (white matter, etc.) remained strong (0.88) and perfect for the ventricles (0.92). Since small individual subcortical ROIs are more difficult for inter-slice reconstruction and lack other contrasts, they have moderate correlations (average 0.72). However, it's still worth noting that all the correlations are strongly significant ($P < 0.05$), providing robust evidence for the authenticity and credibility of our reconstruction results.

TABLE III. DIAGNOSTIC ACCURACY ON ANDI 3 DATASET

Voxel	Scale	ACC	SEN	SPE
4mm	x4	0.9251	0.9231	0.8992
8mm	x8	0.8683	0.8846	0.8539

C. Results of Diagnostic Accuracy

We performed diagnostic experiments on 354 data from the Adni 3 dataset, of which 178 were atrophy subjects and 176 were NC. The subjects are grouped based on age, with each group spanning a ten-year difference. The normal value range of the patient's age group is used as reference. The subjects are classified according to our quantitative indicators and the results are shown in Table III. It can be seen that the accuracy, sensitivity, and specificity of the diagnostic results based on our criteria are satisfying, with an accuracy of 0.92 under x4 enhancement. The accuracy of x8 enhanced results also reaches 0.86, which is deemed satisfactory given the inherent complexity of the task. We have visualized the classification results to intuitively grasp the degree and pattern of the patient's atrophy as shown in Fig. 4. The defects and cavities within the cortex are all marked in red to assist clinicians in gaining a more intuitive understanding of atrophy patterns.

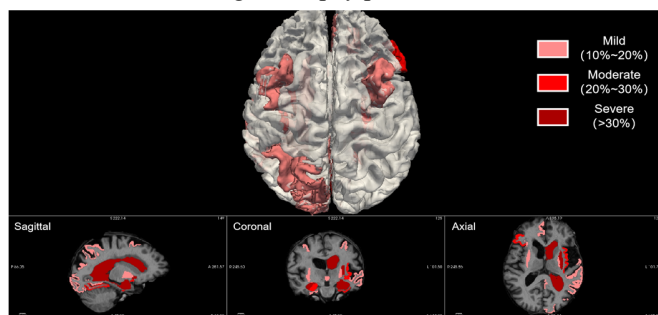


Fig. 4. Distribution of atrophied areas and their degree of atrophy or dilation.

IV. CONCLUSION

In this paper, we present a fully automated brain atrophy diagnostic method that can be applied clinically, characterized by the use of deep learning for inter-slice interpolation preprocess and the establishment of a specific set of morphological criteria to complete the entire automated diagnostic process. We first use low spatial resolution images as inputs to the network and obtain the outputs with isotropic resolution. Then conduct morphometric diagnosis through our

defined quantitative criteria. Finally, determine whether the case exhibits pathologic atrophy. The experiments show that the reconstruction results are authentic and reliable, in which the tissues and organs have a high correlation with ground truth. The diagnostic results based on the quantitative indicators are accurate and can assist doctors in clinical diagnosis. Our method has proven to be effective and flexible, accommodating various clinical procedures.

REFERENCES

- [1] C. Salvatore, A. Cerasa, P. Battista, M. Gilardi, A. Quattrone, and I. Castiglioni, "Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach," *Frontiers in Neuroscience*, pp. 307, September 2015.
- [2] J. Rasmussen and H. Langerman, "Alzheimer's disease – why we need early diagnosis," *Degenerative Neurological and Neuromuscular Disease*, vol. 9, pp. 123–130, December 2019.
- [3] B. Fischl, "FreeSurfer," *Neuroimage*, vol. 62, pp. 774–781, 2012.
- [4] Iglesias JE, et al., "SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry," *Sci. Adv.*, vol. 9, pp. 1–15, February 2023.
- [5] O. Oren, E. Kebebew, J. P. A. Ioannidis, "Curbing unnecessary and wasted diagnostic imaging," *JAMA*, vol. 321, pp. 245–246, 2019.
- [6] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on mri," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, April 2019.
- [7] Zhang F., et al., "An explainable two-dimensional single model deep learning approach for Alzheimer's disease diagnosis and brain atrophy localization," *arXiv preprint, arXiv:2107.13200*, 2021.
- [8] X. Zhang, L. Han, L. Han, H. Chen, D. Dancy and D. Zhang, "sMRI-PatchNet: A Novel Efficient Explainable Patch-Based Deep Learning Network for Alzheimer's Disease Diagnosis With Structural MRI," *IEEE Access*, vol. 11, pp. 108603-108616, 2023.
- [9] Silva M.D., et al., "Distinguishing Healthy Ageing from Dementia: A Biomechanical Simulation of Brain Atrophy Using Deep Networks," *Machine Learning in Clinical Neuroimaging*, vol 13001, 2021.
- [10] Laso Pablo, et al., "Quantifying white matter hyperintensity and brain volumes in heterogeneous clinical and low-field portable MRI," *arXiv preprint, arXiv:2312.05119*, 2023.
- [11] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.
- [12] M. -I. Georgescu, et al., "Multimodal Multi-Head Convolutional Attention with Various Kernel Sizes for Medical Image Super-Resolution," *2023 IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2194-2204, 2023.
- [13] J. Hu, L. Shen and G. Sun, "Squeeze-and-Excitation Networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 7132-7141.
- [14] S. Woo, J. Park, J.-Y. Lee and I. S. Kweon, "CBAM: Convolutional block attention module," *Proc. Eur. Conf. Comput. Vis.*, pp. 3-19, Sep. 2018.
- [15] Sastre-Garriga J., Pareto D., Rovira À, "Brain atrophy in multiple sclerosis: clinical relevance and technical aspects," *Neuroimaging Clinics of North America*, vol. 27, no. 2, pp. 289–300, 2017
- [16] Amiri H., et al., "Urgent challenges in quantification and interpretation of brain grey matter atrophy in individual MS patients using MRI," *Neuroimage Clin.*, vol. 19, pp. 466–475, 2018.
- [17] Allen J.S., Damasio H., Grabowski T.J., Bruss J., Zhang W., "Sexual dimorphism and asymmetries in the gray–white composition of the human cerebrum," *Neuroimage*, vol. 18, pp. 880–894, 2003.
- [18] C. R. Jack Jr., et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Magn. Reson. Imaging*, vol. 27, pp. 685–691, 2008.