Early Diagnosing and Transformation Prediction of Alzheimer's Disease Using Multi-Scaled Self-Attention Network on Structural MRI Images with Occlusion Sensitivity Analysis

Xinxin Fan^{a,b,1}, Haining Li^{c,1}, Lin Liu^b, Kai Zhang^a, Zhewei Zhang^a, Yi Chen^a, Zhen Wang^d, Xiaoli He^e, Jinping Xu^{a,*}, Qingmao Hu^{a,*} and Alzheimer's Disease Neuroimaging Initiative² ^aInstitute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China ^bUniversity of Chinese Academy of Sciences, Beijing, China ^cDepartment of Neurology, General Hospital of Ningxia Medical University, Yinchuan, China

Department of Neurology, General Hospital of Ningxia Medical C

^dZhuhai Institute of Advanced Technology, Zhuhai, China

^eDepartment of Psychology, Ningxia University, Yinchuan, China

Handling Associate Editor: Zhijun Yao

Accepted 13 November 2023 Pre-press 27 December 2023

Abstract.

Background: Structural magnetic resonance imaging (sMRI) is vital for early Alzheimer's disease (AD) diagnosis, though confirming specific biomarkers remains challenging. Our proposed Multi-Scale Self-Attention Network (MUSAN) enhances classification of cognitively normal (CN) and AD individuals, distinguishing stable (sMCI) from progressive mild cognitive impairment (pMCI).

Objective: This study leverages AD structural atrophy properties to achieve precise AD classification, combining different scales of brain region features. The ultimate goal is an interpretable algorithm for this method.

Methods: The MUSAN takes whole-brain sMRI as input, enabling automatic extraction of brain region features and modeling of correlations between different scales of brain regions, and achieves personalized disease interpretation of brain regions. Furthermore, we also employed an occlusion sensitivity algorithm to localize and visualize brain regions sensitive to disease.

*Correspondence to: Dr. Jinping Xu, Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. E-mail: jp.xu@siat.ac.cn and Dr. Qingmao Hu, Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. E-mail: qm.hu@siat.ac.cn.

¹These authors contributed equally in this work.

²Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). As such, the investigatorswithin the ADNI contributed to the design and implementation of ADNI and/or provided databut did not participate in analysis or writing of this report. A complete listing of ADNIinvestigators can be found at: http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Results: Our method is applied to ADNI-1, ADNI-2, and ADNI-3, and achieves high performance on the classification of CN from AD with accuracy (0.93), specificity (0.82), sensitivity (0.96), and area under curve (AUC) (0.95), as well as notable performance on the distinguish of sMCI from pMCI with accuracy (0.85), specificity (0.84), sensitivity (0.74), and AUC (0.86). Our sensitivity masking algorithm identified key regions in distinguishing CN from AD: hippocampus, amygdala, and vermis. Moreover, cingulum, pallidum, and inferior frontal gyrus are crucial for sMCI and pMCI discrimination. These discoveries align with existing literature, confirming the dependability of our model in AD research.

Conclusions: Our method provides an effective AD diagnostic and conversion prediction method. The occlusion sensitivity algorithm enhances deep learning interpretability, bolstering AD research reliability.

Keywords: Alzheimer's disease, deep learning, explainable deep learning, multilevel feature learning, structural MRI

INTRODUCTION

Alzheimer's disease (AD) is a progressive and irreversible mental disorder that typically affects older individuals. Currently, over 30 million people worldwide are affected by dementia disease, 60% to 70% of which were furtherdiagnosed as AD [1]. The ongoing global aging trend is expected to result in a significant increase in the number of people affected by AD, thereby further straining public healthcare services worldwide. Although drug intervention and memory optimization are helpful to AD, they can only delay the progression of AD rather than effective cure. It is highly urgent to make early diagnosis of AD [2].

In recent years, researchers have conducted extensive studies on the diagnosis, etiology, and treatment of AD using various biomarkers [3-12]. However, early diagnosis of AD is challenging due to its insidious onset and lack of early symptoms. Objective biomarkers that can be used to identify patients at early stages of AD are useful for early interventions and better disease management. With the increasing processing power of computers, deep learning methods have been increasingly utilized for auxiliary diagnosing a variety of diseases [6, 7, 13-18]. One such application is the use of deep learning on structural magnetic resonance imaging (sMRI) data and for computer-aided diagnosis (CAD) of AD and its prodromal stage (i.e., mild cognitive impairment (MCI)). Currently, CAD methods can be broadly classified into three categories: 1) region-based, 2) whole-brain-based, and 3) patch-based.

Previous studies [19–23] on AD has predominantly utilized region-based methods using several common processing steps: 1) manually or semiautomatically extracting of features from local brain regions; 2) applying dimensionality reduction algorithms to process MRI data; and 3) using machine learning or other classifiers to predict outcomes based on the reduced-dimensional data. In the study conducted by Sørensen et al. [24], they extracted the shape and structural features of bilateral hippocampus as input features for support vector machine (SVM) to perform classification task. Gutman et al. extracted brain cortex and hippocampal atrophy from predefined regions of interest (ROI) as features to early diagnosing of AD [19, 25]. Furthermore, Zhang et al. utilized SVM to classify individuals with MCI from AD using gray matter volumes of nonoverlapping regions across the entire brain [26]. With the development of brain atlases, Koikkalainen et al. extracted brain region features from multiple atlases and constructed a classifier for predicting MCI conversion [27, 28]. However, manual region extraction is time-consuming and experience-driving. Additionally, these extracted regions are based on prior knowledge, which are not enough to sufficiently represent all the information in the brain.

Compared to traditional region-based methods, deep learning methods have demonstrated advantages by utilizing multiple convolutional kernels to extract various features. As a result, researchers do not need to focus excessively on how the models extract and utilize specific features. However, training deep learning models require significant computational resources. To reduce computational costs, Qiu et al. implemented a patch-based method. They randomly sampled $64 \times 64 \times 64$ voxel patches from 3D MRI images and used them as input features for a 3D convolutional neural network (CNN) for the diagnosis of AD [29]. Their task of CN versus AD classification achieved an accuracy (0.83), sensitivity (0.76), and specificity (0.89). Meanwhile, Leela et al. used a uniform sampling technique by selecting 27 voxel blocks that were $50 \times 41 \times 40$ in size from MRI images. They assigned each patch corresponding to a dedicated 3D CNN network [30]. The results of all 3D CNN networks were fed into a fusion layer for classification, which achieved accuracy of 0.97 for the CN versus MCI versus AD classification task. Lian et al. adopted a similar patch-based strategy and divided a complete image into multiple patches [31]. However, they assigned different weights to different patches using potential pruned sub-networks. In the CN versus AD classification task, their method attained better performance with accuracy (0.90), sensitivity (0.82), and specificity (0.96). For the pMCI versus sMCI classification task, their method resulted in an accuracy (0.81), sensitivity (0.52), and specificity (0.78). Nevertheless, the patch-based method seems to be a compromise way to computational resources. It sacrifices a significant amount of patch-level information and inter-patch correlations, thus the global information derived from whole-brain MRI images has been overlooked.

With the continuous advancement of GPU memory technology, computational resources are more abundant than ever. Consequently, researchers focused more on studying whole-brain models, which used features from the entire brain image in network models. Features have been automatically extracted through operations such as convolution. Moreover, utilizing the whole-brain method enables the retention of global information in the images, effectively reducing the loss of global information that results from patch-based methods. Lim et al. proposed a modified 16-layered visual geometry group (VGG) network for AD classification using sMRI images [32]. Li et al. proposed a deep network with residual blocks for AD diagnosis employing 1,776 sMRI images from the ADNI database [33]. Hoang et al. applied a vision transformer to slice whole-brain MRI images [34]. Lian et al. proposed an attention-guided model for classification of AD based on the whole brain [35]. The model utilized pre-trained full-brain MRI images and visualization of attention maps as guidance to achieve a secondary diagnostic model for AD.

As deep learning is a model for automatic feature extraction and contains a large number of parameters, explaining how deep learning models work is still an exploratory task. However, objective interpretation of deep learning models can still be achieved with methods such as class activation mapping (CAM) [36]. This approach enables exploration of important brain regions and plays significant roles in classification tasks of AD versus CN. CAM methods use a principle similar to the embedding method in feature selection. By extracting the weight parameters of a fixed feature layer, regions linked to the classification category are visualized. However, this method showed two drawbacks: 1) the feature map can only be uti-

lized in the final layer; and 2) information interpreted is biased towards semantic information, lacking figurative boundaries and relationship between brain regions. To solve this problem, Selvaraju et al. [37] proposed the Grad-CAM method, which optimizes the visualization region by using the gradient weights of the feature maps on top of CAM. However, since the probability map resulted from Grad-CAM is in the middle layer of the model, its resolution and sample size will be smaller than the original input image. Moreover, both CAM and Grad-CAM share the same limitation that the deeper visualizations tend to favor semantic information over structural information, resulting in weaker presentations of the latter. Zeiler et al. summarized several CNN visualization methods to modify and adapt the model's architecture [38]. The occlusion sensitivity method was mentioned in this study, which systematically masks different parts of the input image with a gray square to answer the question and monitor the classifier's output. The model locates objects in the scene because the probability of correctly recognizing an object decreases significantly when it is occluded. The output images resulted from this method are consistent with the size of the original input image, and the sensitivity of each region to the classifier can be clearly identified.

When using deep learning methods to classify AD, the issue of interpretability must be considered. Qiu et al. borrowed the idea of CAM and proposed an AD classification explanation model based on fully convolutional networks (FCN) [29], which is a type of deep neural network that employs full convolutions. It plays a significant role in the field of image processing. Lian et al. visualized the brain regions that contribute to AD classification using an attention-guide method [35]. Chen et al. is aware of the unexplainability of deep learning and manually extracted brain regions with special significance, and established a correlation matrix to explore the impact of various brain regions on AD classification [39].

Considering the advantages and disadvantages of the above three methods and the interpretability of deep learning, we proposed a multi-scale self-attention network (MUSAN) that directly uses whole-brain sMRI as input and introduced an occlusion sensitivity algorithm to explain the predictions of the model in the current study. First, we hypothesized that different brain regions of different scales show different effects on AD classification, and there are interrelated relationships among brain regions. Therefore, we designed a multi-scale extraction scheme and integrated self-attention mechanisms to model the correlations between brain regions. Then, the sensitivity of each brain region to AD was predicted using occlusion. Finally, we evaluated our method on three datasets for the classification tasks of CN versus AD and sMCI versus pMCI.

MATERIALS AND METHODS

Participants

Data used in the preparation of this article were obtained from the Alzheimer's DiseaseNeuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner,MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see https://www.adni-info.org

T1 images of 1,080 subjects were used in this work, which were obtained from the ADNI, including: 1) ADNI-1 [40], 2) ADNI-2 [40], and 3) ADNI-3 [40]. The demographic information of the subjects is summarized in Table 1.

ADNI-1: The baseline ADNI-1 dataset consists of 1.5T T1-weighted sMRI scans. These subjects were divided into three categories (i.e., AD, MCI, and CN) depends on the standard clinical criteria, including Mini-Mental State Examination (MMSE) scores and Clinical Dementia Rating. According to whether MCI subjects would convert to AD within 36 months from the baseline assessment, the MCI subjects were further subdivided into stable MCI(sMCI) that were always diagnosed as MCI at all time points (0–96 months), and progressive MCI (pMCI) that converted to AD within 36 months of baseline. Above all, the baseline ADNI-1 datasets contain 189 AD subjects, 154 pMCI subjects, 178 sMCI subjects, and 229 CN subjects.

ADNI-2: The baseline ADNI-2 dataset consists of 3T T1-weighted sMRI scans. According to the same clinical criteria as those used in ADNI-1, these subjects were divided into two categories as 62 AD subjects and 130 CN subjects.

ADNI-3: The baseline ADNI-3 dataset consists of 3TT1-weighted sMRI scans. And those subjects were divided into two categories as 41 AD subjects and 97 CN subjects depends on the same clinical criteria as ADNI-1 and ADNI-2.

MRI data acquisition

Raw unprocessed T1-weighted MRI images were downloaded from the ADNI database, which were scanned using different MRI scanners at multi-sites. Details about data acquisition protocol can be found in ADNI official webpage (https://adni.loni.usc.edu/methods/documents/).

Data preprocessing

We pre-processed all sMRI scans using a standard pipeline as shown in Fig. 1. Specifically, the direction and origin coordinates of the MRI scans were normalized using Simple ITK [41]. Then, the brain skull and dura were stripped by Deep-brain [42]. The FLIRT [43] method in the FSL(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) package was used to linearly align all sMRI to the MNI152 template to remove the global linear difference and also to resample all images to the same spatial shape

	Baseli	ne demographic informa	tion of the subjects	
Dataset	Group	Gender (Male/Female)	Age $(Mean \pm SD)$	$\begin{array}{c} \text{MMSE} \\ \text{(Mean} \pm \text{SD)} \end{array}$
ADNI-1	AD pMCI sMCI CN	180 (96/89) 154 (98/56) 178 (117/61) 229 (119/110)	$75.33 \pm 7.58 74.32 \pm 7.14 73.71 \pm 7.62 74.32 \pm 6.26$	$\begin{array}{c} 22.59 \pm 2.97 \\ 26.58 \pm 1.70 \\ 27.34 \pm 1.81 \\ 28.98 \pm 1.16 \end{array}$
ADNI-2	AD CN	62 (36/26) 130 (67/63)	$\begin{array}{c} 75.94 \pm 7.92 \\ 74.72 \pm 6.34 \end{array}$	$20.64 \pm 4.23 \\ 28.75 \pm 1.38$
ADNI-3	AD CN	41 (22/19) 97 (36/61)	$\begin{array}{c} 72.75 \pm 9.63 \\ 69.91 \pm 6.74 \end{array}$	$22.76 \pm 3.56 \\ 29.09 \pm 1.11$

 Table 1

 Baseline demographic information of the subjects

AD, Alzheimer's disease, sMCI, stable mild cognitive impairment; pMCI, progressive mild cognitive impairment; CN, cognitive normal; MMSE, Mini-Mental State Examination; SD, standard deviation.



Fig. 1. Image preprocessing steps using an AD case as example.1) Skull - striping image, 2) MNI152_1mm as matching template, 3) N3 Correction, and 4) Normalized image.

 $(181 \times 218 \times 181)$. Finally, all images were corrected for intensity using the N3 algorithm [44] and the intensity of image was normalized to a range of 0 to 1.Retico mentioned that using whole-brain will be able to obtain higher area under curve (AUC) [45], and the algorithm using whole-brain can also reduce the resources consumed by segmentation, the need for additional design of segmentation models, and the loss of accuracy caused by segmentation leading to bias in the classification model for downstream tasks. Therefore, after preprocessing the whole-brain data, we do not perform any segmentation operation and retain all the features.

Multi-scale self-attention network (MUSAN)

As illustrated in Fig. 2, we proposed a comprehensive MUSAN to capture multi-level discriminative knowledge from the whole-brain sMRI images. Initially, we utilized ResNet [46] as the fundamental classification backbone. Afterwards, we enhanced model's proficiency in feature extraction by employing multi-scale feature extraction and built the relationship between different scaled feature maps by self-attention. Finally, the global linear layer was used to classify the subjects through the extracted feature maps.

3D convolutional neural networks (CNNs)

Deep CNNS offer an end-to-end learning method that learns features through multiple layers of convolutional operations and multi-level feature combinations [47]. Essentially, it combines highlevel and low-level features and enriches them on the feature level through multi-level feature stacking. In our work, we utilized 3D CNNs, which can perform convolutional operations on 3D volumes. Compared to 2D CNNs, 3D CNNs can filter features in the x, y, and z directions, allowing improved integration of spatial information in images. In this network, a 3D CNN with a standard convolutional kernel size of $7 \times 7 \times 7$ was used to filter the input images.

Scale pyramid convolution (SPC)

Despite the fact that current 3D CNNs are capable of achieving multi-scale feature extraction by combining high-level and low-level features, they can only extract information from a single receptive field within the same level of feature extraction. Therefore, they are not suitable for extracting features from the brain regions. To address this issue, we took inspiration from the pyramid feature extraction method [48] and proposed a novel feature extraction module named scale pyramid convolution (SPC). In this module, we divided the features into multiple branches, each with a channel number of C. By utilizing parallel computing of tensor information in different branches, our network worked rapidly to extract and integrate multi-scale features. For each branch, we have utilized convolutional kernels of different sizes to extract information from different receptive fields, obtaining different spatial resolutions and depths. Finally, we fused the computation results of these



Fig. 2. The architectural diagram of the multi-scaled self-attention network (MUSAN). MUSAN is comprised of 3D convolution and multiscaled self-attention (MUSA) blocks. Each MUSA block consists of a residual-structured 3D convolution and a MUSA module. The MUSA module incorporates spatial pyramid convolution (SPC) and self-attention mechanisms.

branches together using concatenation and obtained feature maps with the same size as the original input. This indicates that the SPC module is a plug-and-play deep learning model that can be conveniently applied to various image processing tasks.

Specifically, we proposed a new criterion for selecting group size without using additional parameters. We eventually divided the number of channels of each convolutional kernel into multiple groups, and the size of each group is evenly distributed according to the total number of groups. This criterion can be expressed as:

$$g = C//G$$

where the quantity C is the channel of input feature map, the G is the total number of groups, and g is the group size. Note that C should be divisible by G. Finally, the multi-scale feature map generation function is given by:

$$F_i = Conv(k_i \times k_i \times k_i, g)(X) \quad i = 0, 1, 2 \dots G-1$$

where the *i*-th kernel size $k_i = 2 \times (i+1)+1$, and $F_i \in \mathbb{R}^{C \times H \times W \times D}$ donates the feature map with different scales. The whole brain feature map could be

obtained by concatenation way as:

$$F = Cat([F_0, F_1, \dots, F_{G-1}])$$
 $i = 0, 1, 2 \dots G-1$

Self-attention layer

In this study, we utilized the SPC module to extract multi-scale spatial information. However, CNN is limited to handle solely local receptive fields, and the inability to obtain global information emerged as a challenge. Recently, Wang et al. [49] introduced the self-attention module to model the correlation between multi-scale information. This method can better capture the relationship between different scales, thereby improving the performance and accuracy. In fact, the self-attention module is capable of mapping the input features into query, key, and value vectors. The key and value vectors represent the feature information extracted by each convolutional block from sMRI images, while the query vector determines ROIs that needs to be focused on during the learning process. This module enables the extraction of important features from sMRI images for classification tasks. By using the $1 \times 1 \times 1$ convolution filter, the key, query, and value will be transformed to vectors and can be denoted by k(x),

q(x), v(x) as follow:

$$Key : k(x) = W_k x$$

$$Query : q(x) = W_q x$$

$$Value : v(x) = W_v x$$

Here, $x \in \mathbb{R}^{C \times N}$ is the feature from SPC module. C is the number of channels and N is the location embedding of features from SPC output. W_k , W_q , W_v are all $1 \times 1 \times 1$ convolution filters. So the selfattention map (a_{ij}) can be calculated as:

$$a_{i,j} = \frac{exp(k(x_i)^T q(x_j))}{\sum\limits_{i=1}^N k(x_i)^T q(x_j)}$$

Here, a_{ij} donates the relationship between *i*-*th* region and *j*-*th* region. By adopting this method, it becomes feasible to model the inter-relationships between diverse features, thereby facilitating the computation of correlations between multi-scale features within the SPC module. Consequently, this leads to the realization of information exchange across multiple scales of brain regions. The output of attention layer is $O = (O_1, O_2, \dots, O_N) \in \mathbb{R}^{C \times N}$ where

$$so_j = W\left(\sum_{i=1}^N a_{i,j}h(x_i)\right)$$

In order to keep the same number of channels as the inputs, *W* is the $1 \times 1 \times 1$ convolution filter.

Multi-scaled self-attention block (MUSA block)

We proposed a novel MUSA block that combines the SPC module with self-attention mechanisms to enable efficient extraction and correlation modeling in sMRI images. The MUSA block is easy to integrate into any 3D network thanks to its scalable design. The input is first processed by a 3D CNN to extract non-linear features, and then fed into the MUSA module. Within the MUSA module, the SPC module initially processes the input features to extract multiscale feature information. The output of SPC is then subjected to self-attention mechanism for modeling multi-scale correlations of brain regions. The output is a 3D image with the same size as the input image. After applying MUSA blocks to the input, pooling is performed to down-sample the output, followed by residual connections to avoid issues such as gradient vanishing.

Occlusion sensitivity

To investigate the impact of brain regions on AD, we introduced the concept of occlusion sensitivity to get the effect of local brain regions on AD as shown in Fig. 3. First, we employed a sliding window approach to occlude the sMRI images by setting the voxel values within the occlusion blocks to 0. In this experiment, we utilized occlusion blocks of size $8 \times 8 \times 8$. In order to ensure sufficient smoothness in the occlusion sensitivity map, we set the overlap rate of the sliding window to 0.25. The occluded images are then fed into our pre-trained model to obtain occluded score_i. Subsequently, these scores are compared with the predicted scores score orig obtained from the unoccluded original images, which yielded a different measure (diff) between the two scores. Finally, we arranged the scores of all occlusion blocks in the order of occlusion to reconstruct the occlusion sensitivity map of the sMRI images. When the diff value approaches to 0, it indicates a lower sensitivity in the corresponding region, suggesting that changes in that region would not significantly impact disease diagnosis. On the other hand, when the diff value approaches to-1 or 1, it indicates that our model exhibits sensitivity to that particular region in the individual, implying that alterations in these regions would have an impact on disease diagnosis.

Experimental setting

The proposed MUSAN was implemented on a single GPU (NVIDIA TITAN 24GB) using Python3.9 based on Pytorch 1.12 package. The input of our network has been prepared into the same spatial shape $128 \times 128 \times 128$ by spatial padding and resize. The model was trained by the cross-entropy loss with the Adam optimizer for 200 epochs and L2-Regularization value of le^{-4} . The training set was augmented online by the combination of: 1) randomly flipping the brain images, 2) randomly rotate the brain images, and 3) randomly contrast transform the brain images. During training, resampling is used to balance the proportion of training categories that will be.

To evaluate the generalization ability of various methods, we used ADNI-1 as the training set and ADNI-2 and ADNI-3 as the test sets during the experiment's evaluation phase. Quantitative assessment of



Fig. 3. The computational workflow of the occlusion sensitive algorithm. The algorithm employs region-wise occlusion by inputting occluded images into the network to predict the difference between the predicted results of the modified image and the original image.

performance was assessed using accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristic curve (AUC). The AUC measures the model's capacity to distinguish positive and negative samples by calculating the area under the AUC. In addition, we conducted comparative experiments with state-of-the-art wholebrain-based approaches in our empirical evaluation, and further performed ablation studies on the proposed modules to demonstrate the superiority and correctness of our model.

- ResNet, a conventional 3D CNN, has been extensively used to tackle the problem of gradient vanishing in deep networks by incorporating residual modules. In this study, Liu et al. [50] employed ResNet to classify AD from healthy controls using features from hippocampal regions. Furthermore, Bae et al. [51] applied transfer learning using ResNet for the classification of sMCI and pMCI on MRI images.
- 2) SEResNet, a neural network architecture that utilizes ResNet as its backbone and introduces a squeeze-and-excitation (SE) block module. The SE blocks learn a set of weights that capture channel-wise dependencies within the feature maps. These weights are subsequently employed to scale the feature maps before they are propagated to the subsequent layer. Ji et al. [52] applies the SeResNet architecture to obtain features of gray matter and white matter for the classification of CN, MCI, and AD.
- 3) DenseNet [53] addresses the vanishing gradient problem in very deep neural networks by introducing dense connections between layers. In a dense block, each layer is connected to every other layer in a feed-forward fashion, creating

a dense connectivity pattern. This allows feature reuse and encourages feature propagation throughout the network, leading to better gradient flow and improved model performance. He et al. [54] utilized DenseNet for the classification of AD and CN on MRI images.

RESULTS

Classification results of CN versus AD

In the task of CN versus AD, we compared our MUSAN to three competing methods. We trained the model in ADNI-1 and tested them in ADNI-2 and ADNI-3, respectively. We quantified the performance of our model using four different metrics, including ACC, SEN, SPE, and AUC in Table 2. Importantly, our proposed MUSAN outperforms these methods by $1\% \sim 2\%$ in accuracy on the ADNI-2 dataset, while attaining acceptable ACC (0.956), SEN (1.000), SPE (0.853), and AUC (0.968) on the ADNI-3 dataset. Although the MUSAN approach yielded a sensitivity score of 1.00, which might highlight some bias in the model, there is no need to worry much about the deviation introduced by the model as the specificity achieves a commendable level of 0.85. Furthermore, the AUC also reaches a value of 0.968, highlighting the model's discriminative capacity.

Classification result of sMCI versus pMCI

In the task of sMCI versus pMCI, we also compared our MUSAN network to three competing methods. The model was trained on ADNI-1, 10% subjects were randomly selected for validation and another 10% for evaluation. Performance metrics,

whole-brain-based methods								
Methods		AD	NI-2			AD	NI-3	
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
ResNet50 [51]	0.911	0.857	0.887	0.948	0.906	0.889	0.780	0.906
SEResNet50 [52]	0.901	0.907	0.790	0.928	0.913	0.939	0.756	0.900
DenseNet121 [54]	0.922	0.873	0.877	0.949	0.920	0.941	0.781	0.935
ResAttnNet50 [57]	0.927	0.900	0.871	0.945	0.913	0.914	0.780	0.926
MUSAN (Ours)	0.932	0.962	0.822	0.953	0.956	1.000	0.853	0.968

Table 2 Result of CN versus AD classification on ADNI-2 and ADNI-3, trained on ADNI-1 with whole-brain-based methods

Table 3 Result of pMCI versus sMCI classification on the ADNI-1 test set, trained on the ADNI-1 training set with whole-brain-based methods

Methods		AD	NI-1	
	ACC	SEN	SPE	AUC
ResNet50 [51]	0.755	0.667	0.588	0.745
SEResNet50 [52]	0.735	0.750	0.353	0.645
DenseNet121 [54]	0.775	0.750	0.529	0.773
ResAttnNet50 [57]	0.782	0.761	0.516	0.787
MUSAN (Ours)	0.848	0.743	0.839	0.863

including ACC, SEN, SPE, and AUC were used to assess the model's effectiveness (Table 3). As a result, our proposed method significantly improves the recognition accuracy of sMCI and pMCI. Specially, it can be observed that DenseNet121 performs much better than ResNet50 (Table 3). These results indicate that the performance improvement is attributed to the network depth and dense connections, which demonstrates the effectiveness of multi-scale dense information exchange for recognizing sMCI and pMCI. Moreover, the accuracy increases from 0.755 to 0.848 and the AUC from 0.745 to 0.863 when the residual modules in ResNet50 was replaced by MUSA modules. These results provide sufficient evidence for the validity of our multi-scale and brain region-related modeling method in sMCI versus pMCI classification tasks. However, SEResNet, a spatial attention mechanism network that focuses on spatial information within the same scale, performs worse in the sMCI versus pMCI task. It indicates that constraining information within the same scale will lead to inferior outcomes and hinders the interaction among multi-scale information.

Ablation experiment

To examine how multi-scale feature extraction and self-attention mechanism brain modeling affect the classification of CN versus AD and sMCI versus pMCI, ablation experiments were carried out. Our method employs ResNet50 as the baseline model, and three various experiments have been subsequently performed: 1) replacing the residual connection module with the SPC module based on ResNet50; 2) replacing the residual connection module with the self-attention module which is based on ResNet50; and 3) replacing the residual module with a concatenated module of SPC and self-attention, which is based on ResNet50.

To this end, we achieved satisfactory results on the CN versus AD classification task as compared to ResNet network (Fig. 4). However, this improvement was relatively small even combining MUSA module, indicating that the ResNet framework is already capable of handling the CN versus AD task. This also suggests that deep learning models can solve tasks with distinct anatomical structures using simple models with high accuracy. However, as shown in Fig. 5, the ResNet performs poorly on sMCI versus pMCI task. By using the self-attention module, we were able to improve the accuracy from 0.755 to 0.782 and the AUC from 0.747 to 0.787, demonstrating the usefulness of self-attention mechanisms to model brain region features for disease classification. By using the SPC multi-scale mechanism, we were able to improve the accuracy from 0.755 to 0.815 and the AUC from 0.747 to 0.804. These findings suggest that considering multi-scale fusion of brain region information can effectively identify brain regions related to disease. This also confirms that different brain regions at different scales are affected by AD. By combining SPC with self-attention, we found that the accuracy of ResNet was increased from 0.755 to 0.848 and the AUC from 0.747 to 0.863, fully demonstrating the correctness of our multi-scale brain region feature correlation modeling.

Occlusion sensitivity maps

We examined discriminative brain regions for CN versus AD and sMCI versus pMCI at the cohort and individual levels, respectively. First, we normalized our occlusion sensitivity maps onto the Anatomical Automatic Labeling (AAL) template and filtered out



Fig. 4. Result of CN versus AD on the ADNI-2 dataset, obtained by the ResNet, ResNet + Self-Attention, ResNet + SPC, and MUSAN on the ADNI-1.



Fig. 5. Result of sMCI versus pMCI on the ADNI-1 dataset, obtained by the ResNet, ResNet + SelfAttention, ResNet + SPC, and MUSAN on the ADNI-1.

the six regions with maximum sensitivity based on the occlusion sensitivity value of the brain regions mapped onto the AAL template. As shown in Tables 5 and 6, the discriminating regions of CN versus AD were mainly in the medial temporal lobe and cerebellar regions, such as the hippocampus, amygdala, and parahippocampal gyrus, whereas the discriminating regions of sMCI versus pMCI were mainly located in the posterior cingulate gyrus, the pallidum, and inferior frontal gyrus. In addition, we also showed the results of our method on individual level (Figs. 6 and 7).

DISCUSSION

Comparison with previous work

In this study, we also used some difficult-toreproduce state-of-the-art methods as comparisons, whose original data and source code were not publicly available. It is relatively difficult for us to reproduce the model using the details disclosed in the literature. Although it would be somewhat unfair to rely directly on the original authors' results for comparison, it is worth discussing and analyzing. Table 4 shows the outcomes of three traditional methods and four deep learning methods, respectively, for the CN versus AD and sMCI versus pMCI tasks. Compared to the three traditional methods [26, 55, 56], the deep learning methods [31, 35, 57–59] demonstrate significantly superior performance across the four evaluation metrics. Additionally, as shown in the results of the region-based approaches [26], it not only incurs substantial time consumption in extracting ROI but also fails to achieve satisfactory outcomes. In comparison to deep learning techniques, our proposed method achieves higher accuracy. Although Shi et



Fig. 6. Illustrations of occlusion sensitivity algorithms in AD cohort. The first row is an average plot of the sensitivities of all samples in the AD test cohort and mapped onto the AAL template. The next three rows are plots of the sensitivities of the individual sample sensitivities, also mapped onto the AAL template.

al. [58] achieves an impressive accuracy of 0.97 on the CN versus AD task, the limited scale of their utilized dataset restricts the generalizability of their results. The hierarchical-CNN algorithm [31] achieves the best specificity for the CN versus AD task, indicating the highest precision in correctly diagnosing CN cases. However, it exhibits the poorest sensitivity among all deep learning algorithms, implying a higher probability of misclassification in AD patients' diagnosis. Such misdiagnosis can potentially lead to AD patients being classified as CN, thus missing the optimal treatment window. The reasons for this phenomenon, however, might be due to the fact that the data distribution is not balanced. As shown in Table 4, the same dataset was used in hierarchical-CNN and attention-CNN, as well as voxel-based morphometry and ROI-based methods. However, all these methods exhibit high SPE and low SEN on CN versus AD, which leads to more false negatives. In comparison, we used resampling to ensure data balance and obtained the best sensitivity (0.96)

in addition to the highest accuracy in the CN versus AD task. This result suggested that it is necessary to accurately balance the dataset to obtain a better SEN, and then result in fewer misclassifications of AD patients.

Notably, DenseNet shows significantly higher ACC scores than the other two methods on both datasets. This is consistent with the conclusions drawn by Hazarika et al. [60], who performed a multi-class classification of CN, MCI, and AD using DenseNet on MRI 2D slices, and achieved better performance than the VGG, ResNet, and other networks. This consistency underscores the superiority of DenseNet's densely connected architecture to improve classification outcomes. Furthermore, it highlights the significance of interplay among features at multiple scales for enhancing AD classification performance.

Our MUSAN model distinguishes itself from other deep learning models. In comparison to the hierarchical-CNN, our MUSAN model is based



Fig. 7. Illustrations of occlusion sensitivity algorithms in pMCI cohort. The first row is an average plot of the sensitivities of all samples in the pMCI test cohort and mapped to the AAL template. The next three rows are plots of the sensitivities of the individual sample sensitivities, also mapped to the AAL template.

on whole-brain analysis, eliminating the need for patch-based partitioning of the original images, thus preserving more global information. Compared to the attention-CNN approach, our proposed method is more concise. The attention-CNN [35] requires pre-training to obtain attention maps, which served as prior knowledge for subsequent classification models. However, the classification performance of the attention-CNN is susceptible to limitations imposed by the pre-trained model. In contrast, the MUSAN model employs a self-attention framework for adaptive learning of attention maps. It utilizes the classification results to constrain the attention maps and continuously optimizes them through iterations. Compared to Zhang's method [57], we proposed a multi-scale computation approach and constructed a multi-scale information interaction model.

In summary, 1) the deep learning methods outperform the traditional methods in both the CN versus AD and sMCI versus pMCI classification tasks, indicating that deep learning's automatic feature extraction method is more effective in describing brain information compared to manually crafted feature extraction methods; 2) our proposed approach demonstrates superior performance as compared to other deep learning methods by showing higher accuracy and AUC in the sMCI versus pMCI classification task. This clearly illustrates that our model possesses better feature representations and generalization abilities in complex and challenging tasks; and 3) our method utilizes whole-brain data without patch segmentation and models the interrelationships between brain regions using a self-attention mechanism.

Comparisons between two tasks

Our proposed model was applied to both the CN versus AD and sMCI versus pMCI tasks. The results were consistent with those of previous studies [27, 28, 31, 35, 57, 59, 61–63]. Our model showed improved

the-	Table 4	the-art studies using baseline sMRI data of ADNI-1 for AD classification (AD versus NC) and MCI conversion prediction (pMCI versus
		art studies using baseline s

A brief description o	f the state-of-	the-art studies using baseline sMF	RI data of ADNI-1 for AL) classifica	tion (AD	ersus NC) and MCI	conversio	n predictio	on (pMCI v	ersussMCI)
References	Modality	Subjects	Methods		CN ver	sus AD			sMCI ve	ersus pMCI	
					classif	ication			class	ification	
				ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
[13]	MRI	358AD+670MCI+429CN	Region-based method	0.79	0.79	0.8	0.87	0.66	0.47	0.69	0.64
[44]	MRI	358AD+670MCI+429CN	Voxel-based	0.81	0.77	0.83	0.87	0.64	0.36	0.68	0.59
[43]	MRI	358AD+670MCI+429CN	morphometry Landmark-based	0.82	0.77	0.86	0.88	0.69	0.4	0.73	0.63
			morphometry								
[49]	MRI	41AD+99MCI+52CN	Deep polynomial network	0.97	0.95	0.98	0.97	0.78	0.68	0.86*	0.8
[18]	MRI	358AD+670MCI+429CN	Hierarchical-CNN	0.9	0.82	0.97	0.95	0.81	0.53	0.85	0.78
[45]	MRI	200AD+404MCI+231CN	3D-CNN	0.91	0.91	0.92	0.98	0.82	0.81	0.81	0.92*
[22]	MRI	358AD+670MCI+429CN	Attention-CNN	0.92	0.88	0.94	0.96*	0.82	0.57	0.86	0.79
MUSAN (ours)	MRI	189AD+332MCI+229CN	Multi-scale	0.93	96.0	0.82	0.95	0.85	0.74	0.84	0.86
			self-attention								

performance in the CN versus AD task but struggled to achieve satisfactory results in the sMCI versus pMCI task. This could be resulted from the smaller inter-class distance between sMCI and pMCI but larger inter-class distance between CN and AD. Liu et al. [62] employed voxel-level feature extraction methods and failed to achieve a recognition accuracy of 0.7 in the sMCI versus pMCI task. Similarly, both the voxel-level approaches of Zhang and our proposed method achieved good results in this task. Furthermore, Koikkalainen et al. utilized patch-based approaches for SVM or linear analyze classification models [27, 28], but their results were still inferior to those of patch-based deep learning models. This indicates that a deeper feature extractor is required for the classification of sMCI and pMCI, as the features extracted by traditional methods are insufficient to capture the differences between sMCI and pMCI.

Figure 5 illustrates the differences between the CN versus AD and sMCI versus pMCI tasks. The brain regions for the CN versus AD task primarily involved in the hippocampus, amygdala, and vermis. This finding aligns with those from Ossenkoppele et al. [64], which demonstrated that the medial temporal lobe plays a critical role in the development of AD. In their work, Poulin et al. [65] also proposed a relationship between the amygdala and AD and pointed out that amygdala atrophy plays a crucial role in the early diagnosis of AD. In a recent study, Uttam et al. [66] validated that hippocampal and amygdala volumes were helpful to differentiate AD patients from CN. Sjöbeck et al. [67] found that atrophy of Vermis can assist in determining the disease process in AD by studying the morphology of neuronal and glial changes. However, the performance of sMCI versus pMCI patients does not exhibit the same generality as the CN versus AD task. The sensitive regions for the sMCI versus pMCI task include the cingulum, pallidum, and inferior frontal gyrus. In a recent study, Liu et al. [68] suggested that diffusion kurtosis imaging techniques could be used to study the microscopic changes in the cingulum in patients with MCI and to assess the cognitive function of MCI patients. Lin et al. [69] also suggested that activity of the inferior frontal gyrus helps protect memory from AD in older adults. These discriminative regions align with those results mentioned in multimodal analyses [70]. While the CN versus AD task demonstrates generality, the classification task for sMCI and pMCI is more influenced by individual differences.

Therefore, due to the complexity of the sMCI versus pMCI task, we proposed MUSAN, a model

that can extract multi-scale information and integrate the correlations between different scales of information. Through ablation experiments, we comprehensively demonstrated the effectiveness of two proposed modules, SPC and self-attention, in classifying CN versus AD and sMCI versus pMCI. We observed that the impact of the SPC module surpasses that of self-attention. This result indicates that extracting multi-scale brain region information and integrating them is beneficial for improving the classification accuracy of both tasks. It also indicates that prediction of AD progression should focus more on features or brain regions at different scales. Additionally, we are impressed by the performance enhancement brought by self-attention, as it confirms our assumption that the correlation between brain regions can aid in diagnosing of AD. The success of self-attention also validates the potential of investigating inter-regional correlations, as it can only measure the correlation between two features rather than the correlation among multiple feature combinations.

Model interpretability

The occlusion sensitivity is used to address the interpretability problem in deep learning models. Compared to traditional region-based approaches, our proposed method offers a coarser level of interpretability. Sarasua et al. focused on AD classification based on features from the hippocampus to investigate the impact of hippocampal atrophy on AD [19, 71]. Similarly, Zhang et al. extracted features from specific brain regions to study AD [26]. These traditional methods determine the level of interpretability based on the granularity of the extracted brain regions. In fact, these methods [27, 28] even provide a weight index to represent the contribution of each brain region to AD. However, the classification accuracy of traditional algorithms significantly lags behind to that of deep learning methods. Furthermore, the relationship between brain regions and AD discovered by these methods may be influenced by model biases. Additionally, datasets used in methods are limited, and they do not differentiate between general patterns and individual differences.

Deep learning models can achieve significantly better results than traditional algorithms on largescale datasets. Therefore, if we can explain which brain regions contribute to the performance of deep learning models, it would provide more reliable reference information for doctors or patients. Compared

Table 5 Ordered significant regions from the occlusion sensitivity map related to the prediction of AD patients from normal individuals

Region	Occlusion Sensitivity
	Value
Vermis-10	0.1165
Left Amygdala	0.1031
Right Amygdala	0.0986
Vermis-9	0.0956
Left Hippocampus	0.0951
Right Hippocampus	0.0889

Table 6	
Ordered significant regions from the Occlusion sensitivity Ma	p
related to the prediction of pMCI normal individuals	

Region	Occlusion Sensitivity
e	Value
Right Posterior Cingulate cortex	0.0600
Right Anterior Cingulate cortex	0.0424
Left Pallidum	0.0418
Right Pallidum	0.0413
Left opercular part of	0.0364
Inferior Frontal Gyrus,	
Right Caudate	0.0328

to the studies [29, 35, 57], our proposed method can obtain more fine-grained brain region predictions. The regions of interest mentioned in the brain almost cover the entire brain, requiring the use of a threshold to determine the most important brain regions. However, the values represented by the obtained regions of interest indicate the model's prediction of AD risk in that region and the direct contribution of that region to the prediction score. In previous literature [31, 72], deep learning was combined with region-based approaches to improve model interpretability. This method offers a rapid way to enhance interpretability, but it cannot ignore the loss of global information during the training process. In our method, the shape of the block can be customized, enabling analysis of specific brain regions by adopting their specific shapes. Therefore, our method not only maintains high recognition accuracy but also improves the interpretability of deep learning models, thus expanding the scope of deep learning research in the medical field.

Limitations and feature work

In this study, we utilized an occlusion sensitivity method for interpreting the neural network. However, due to the difficulty in extracting brain region masks, we could only use square-shaped occlusion templates that do not specifically represent particular brain regions. In future work, we are planning to introduce a registration-based method to automatically obtain brain region masks, and then use them to perform occlusion in a partitioned manner, allowing for targeted investigation of the impact of a specific brain region associated with AD. This registrationbased segmentation will avoid the retraining problem of traditional segmentation models and the ability to use different templates for analysis after registration will be more effective in realizing the analysis of brain regions. Additionally, although sensitive regions have been successfully identified using occlusion sensitivity map, these regions did not provide useful information for the training steps of our model. Therefore, in future work, if we can use occlusion sensitivity to constrain the model, we may obtain more precise results. In addition, our method used only MRI images for the study and did not fully utilize information such as MMSE, which is readily available. Although information such as MMSE may be able to improve the model's classification of AD, we have not yet found a suitable method to balance the role of information such as MMSE with that of MRI images in the classification, and simply considering MMSE may leave the model lacking in the analysis of MRI images. Therefore, we will explore the incorporation of this additional information in the future work.

Since the amount of collectable data remains limited, utilizing more extensive data sets can strengthen the models. Generative AI's current prevalence offers a solution for generating images in a controlled manner. In future research, we aim to employ the Diffusion model [73, 74] to generate separate AD and CN sample data and utilize distinct prompts to regulate the process.

Conclusions

In this work, we proposed a multi-scale selfattention deep learning model that significantly improves the identification accuracy of sMCI versus pMCI and CN versus AD task. Through occlusion sensitivity, we predicted the brain regions associated with individuals of AD. Moreover, we evaluated our method on three datasets, and these experimental results demonstrated that our method outperforms several other advanced methods in terms of performance.

ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through 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.; ElanPharmaceuticals, Inc.; Eli Lilly and Company; Euro Immun; F. Hoffmann-La Roche Ltd andits 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 is providing funds to support ADNI clinical sites in Canada. Private sector contributions arefacilitated by the Foundation for the National Institutes of Health (https://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

FUNDING

The research was funded by the National Natural Science Foundation of China (No. 62006220, and No.62001462), Ningxia Key Research Program (No. 2022BEG03158), Shenzhen Science and Technology Research Program (No. JCYJ20180507182441903, No. JCYJ20210324120804012), and the Special Innovation Project of the Guangdong Provincial Department of Education (2021WQNCX068).

CONFLICT OF INTEREST

Dr Jinping Xu is an Editorial Board Member of this journal but was not involved in the peer-review process nor had access any information regarding its peer-review. Other authors have no conflict of interest to report.

DATA AVAILABILITY

The data supporting the findings of this study are openly available in the Alzheimer's Disease Neuroimaging Initiative (https://adni.loni.usc.edu/).

REFERENCES

- Chiba T, Yamada M, Sasabe J, Terashita K, Shimoda M, Matsuoka M, Aiso S (2009) Amyloid-β causes memory impairment by disturbing the JAK2/STAT3 axis in hippocampal neurons. *Mol Psychiatry* 14, 206-222.
- [2] Ahmed OB, Mizotin MM, Benois-Pineau J, Allard M, Catheline G, Amar CB (2015) Alzheimer's disease diagnosis on structural MR images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. *Comput Med Imaging Graph* 44, 13-25.
- [3] Salvatore C, Cerasa A, Battista P, Gilardi MC, Quattrone A, Castiglioni I (2015) Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach. *Front Neurosci* 9, 307.
- [4] Poulakis K, Pereira JB, Muehlboeck JS, Wahlund L-O, Smedby Ö, Volpe G, Masters CL, Ames D, Niimi Y, Iwatsubo T, Ferreira D, Westman E (2022) Multi-cohort and longitudinal Bayesian clustering study of stage and subtype in Alzheimer's disease. *Nat Commun* 13, 4566.
- [5] Sheng J, Xin Y, Zhang Q, Wang L, Yang Z, Yin J (2022) Predictive classification of Alzheimer's disease using brain imaging and genetic data. *Sci Rep* 12, 2405.
- [6] Wu H, Luo J, Lu X, Zeng Y (2022) 3D transfer learning network for classification of Alzheimer's disease with MRI. *Int J Mach Learn Cybern* 13, 1997–2011.
- [7] Ashtari-Majlan M, Seifi A, Dehshibi MM (2022) A multistream convolutional neural network for classification of progressive MCI in Alzheimer's disease using structural MRI images. *IEEE J Biomed Health Inform* 26, 3918-3926.
- [8] Anitha M, Karpagam V, Selvy PT (2021) Diagnostic framework for automatic classification and visualization of Alzheimer's disease with feature extraction using wavelet transform. *Neuroquantology* 19, 84-95.
- [9] Li C, Song L, Zhu G, Hu B-l, Liu X, Wang Q (2022) Alzheimer's level classification by 3D PMNet using PET/MRI multi-modal images. 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), pp. 1068-1073.
- [10] Park S-W, Yeo N, Kim J-K, Yum S, Jang J-W (2022) The image and multimodal classification for Alzheimer's disease by convolutional neural network and combination with long short-term memory using tau-PET images. *Alzheimers Dement* 18, 1552-5260.
- [11] Yin C, Li S, Zhao W, Feng J (2013) Brain imaging of mild cognitive impairment and Alzheimer's disease. *Neural Regen Res* 8, 435-444.
- [12] Zhao W, Wang X, Yin C, He M, Li S, Han Y (2019) Trajectories of the hippocampal subfields atrophy in the Alzheimer's disease: A structural imaging study. *Front Neuroinform* 13, 13.
- [13] Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, Cai M, Yang J, Li Y, Meng X, Xu B (2020) A deep learning algo-

rithm using CT images to screen for Corona virus disease (COVID-19). *EurRadiol* **31**, 6096-6104.

- [14] Sanjay VM, Swarnalatha P (2022) Deep learning techniques for early detection of Alzheimer's disease: A review. Int J Electric Electronics Res 10, 899-905.
- [15] Thayumanasamy I, Ramamurthy KN (2022) Performance analysis of machine learning and deep learning models for classification of Alzheimer's disease from brain MRI. *Traitement Signal* **39**, 1961-1970.
- [16] Qiu S, Miller MI, Joshi P, Lee J, Xue C, Ni Y, Wang Y, De Anda-Duran I, Hwang PH, Cramer JA, Dwyer B, Hao H-I, Kaku MC, Kedar S, Lee PH, Mian A, Murman DL, O'Shea S, Paul AB, Saint-Hilaire M, Alton Sartor E, Saxena A, Shih LC, Small JE, Smith MJ, Swaminathan A, Takahashi CE, Taraschenko O, You H, Yuan J, Zhou Y, Zhu S, Alosco ML, Mez J, Stein TD, Poston KL, Au R, Kolachalama VB (2022) Multimodal deep learning for Alzheimer's disease dementia assessment. Nat Commun 13, 3404.
- [17] Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ (2021) Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors* (*Basel*) 21, 2852.
- [18] Venugopalan J, Tong L, Hassanzadeh HR, Wang MD (2021) Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci Rep* 11, 3254.
- [19] Gutman B, Wang Y, Morra JH, Toga AW, Thompson PM (2009) Disease classification with hippocampal shape invariants. *Hippocampus* 19, 572-578.
- [20] Wang L, Beg F, Ratnanather T, Ceritoglu C, Younes L, Morris JC, Csernansky JG, Miller MI (2007) Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. *IEEE Trans Med Imaging* 26, 462-470.
- [21] Cui W, Yan C, Yan Z, Peng Y, Leng Y, Liu C, Chen S, Jiang X, Zheng J, Yang X (2022) BMNet: A new regionbased metric learning method for early Alzheimer's disease identification with FDG-PET images. *Front Neurosci* 16, 831533.
- [22] Prokopenko D, Lee S, Hecker J, Mullin K, Morgan S, Katsumata Y, Alzheimer's Disease Neuroimaging Initiative, Weiner MW, Fardo DW, Laird NM, Bertram L, Hide WA, Lange C, Tanzi RE (2021) Region-based analysis of rare genomic variants in whole-genome sequencing datasets reveal two novel Alzheimer's disease-associated genes: DTNB and DLG2. *Mol Psychiatry* 27, 1963–1969.
- [23] Mueller K, Meyer-Baese A, Erlebacher G (2022) Multimodal region-based transformer for the classification and prediction of Alzheimer's disease. *Medical Imaging 2022: Biomedical Applications In Molecular, Structural, And Functional Imaging* **12036**, 366-375.
- [24] Sørensen L, Igel C, Hansen NL, Osler M, Lauritzen MJ, Rostrup E, Nielsen M (2016) Early detection of Alzheimer's disease using MRI hippocampal texture. *Hum Brain Mapp* 37, 1148-1161.
- [25] Planche V, Ruet A, Coupé P, Lamargue-Hamel D, Deloire MSA, Pereira B, Manjón JV, Munsch F, Moscufo N, Meier DS, Guttmann CRG, Dousset V, Brochet B, Tourdias T (2017) Hippocampal microstructural damage correlates with memory impairment in clinically isolated syndrome suggestive of multiple sclerosis. *Mult Scler* 23, 1214-1224.
- [26] Zhang D, Wang Y, Zhou L, Yuan H, Shen D (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 55, 856-867.
- [27] Koikkalainen J, Lötjönen J, Thurfjell L, Rueckert D, Waldemar G, Soininen H (2011) Multi-template tensor-based

morphometry: Application to analysis of Alzheimer's disease. *Neuroimage* **56**, 1134-1144.

- [28] Liu M, Zhang D, Shen D (2016) Relationship induced multitemplate learning for diagnosis of Alzheimer's disease and mild cognitive impairment. *IEEE Trans Med Imaging* 35, 1463-1474.
- [29] Qiu S, Joshi P, Miller MI, Xue C, Zhou X, Karjadi C, Chang GH, Joshi AS, Dwyer B, Zhu S, Kaku MC, Zhou Y, Alderazi YJ, Swaminathan A, Kedar S, Saint-Hilaire M, Auerbach SH, Yuan J, Sartor E, Au R, Kolachalama VB (2020) Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain* 143, 1920-1933.
- [30] Leela M, Helenprabha K, Sharmila L (2023) Prediction and classification of Alzheimer disease categories using integrated deep transfer learning approach. *Measur Sens* 27, 100749.
- [31] Lian C, Liu M, Zhang J, Shen D (2020) Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE Trans Pattern Anal Mach Intell* 42, 880-893.
- [32] Lim B, Lai KW, Haiskin K, Kulathilake S, Chao O, Hum Y, Samiappan D, Wu X, Zuo X (2022) Deep learning model for prediction of progressive mild cognitive impairment to Alzheimer's disease using structural MRI. *Front Aging Neurosci* 14, 876202.
- [33] Li H, Habes M, Fan Y (2017) Deep ordinal ranking for multi-category diagnosis of Alzheimer's disease using hippocampal MRI data. arXivPreprarXiv 1709.01599.
- [34] Hoang GM, Kim U-H, Kim J (2023) Vision transformers for the prediction of mild cognitive impairment to Alzheimer's disease progression using mid-sagittal sMRI. *Front Aging Neurosci* 15, 1102869.
- [35] Lian C, Liu M, Pan Y, Shen D (2022) Attention-guided hybrid network for dementia diagnosis with structural MR images. *IEEE Trans Cybern* 52, 1992-2003.
- [36] Zhou B, Khosla A, Lapedriza À, Oliva A, Torralba A (2015) Learning deep features for discriminative localization. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2921-2929.
- [37] Selvaraju RR, Das A, Vedantam R, Cogswell M, Parikh D, Batra D (2016) Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int J Comput Vis* 128, 336-359.
- [38] Zeiler MD, Fergus R (2013) Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, pp. 818-833.
- [39] Chen Y, Xia Y (2021) Iterative sparse and deep learning for accurate diagnosis of Alzheimer's disease. *Pattern Recognit* 116, 107944.
- [40] Jack CR, Jr., Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, L Whitwell J, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DLG, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW (2008) The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* 27, 685–691.
- [41] Beare R, Lowekamp B, Yaniv Z (2018) Image segmentation, registration and characterization in R with SimpleITK. J Stat Softw 86, 8.
- [42] Buda M, Saha A, Mazurowski MA (2019) Association of genomic subtypes of lower-grade gliomas with shape fea-

tures automatically extracted by a deep learning algorithm. *Comput Biol Med* **109**, 218-225.

- [43] Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17, 143-155.
- [44] Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17, 87-97.
- [45] Retico A, Bosco P, Cerello P, Fiorina E, Chincarini A, Fantacci ME (2015) Predictive Models based on support vector machines: Whole-brain versus regional analysis of structural MRI in the Alzheimer's disease. *J Neuroimaging* 25, 552-563.
- [46] He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778.
- [47] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* **521**, 436-444.
- [48] He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 37, 1904-1916.
- [49] Wang X, Girshick R, Gupta A, He K (2018) Non-local neural networks. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7794-7803.
- [50] Liu M, Li F, Yan H, Wang K, Ma Y, Shen L, Xu M (2019) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage* 208, 116459.
- [51] Bae J, Stocks J, Heywood A, Jung Y, Jenkins L, Hill V, Katsaggelos A, Popuri K, Rosen H, Beg MF, Wang L (2021) Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network. *Neurobiol Aging* **99**, 53-64.
- [52] Ji H, Liu Z, Yan W, Klette R (2020) Early diagnosis of Alzheimer's disease based on selective kernel network with spatial attention. *Pattern Recognit* **12047**, 503-515.
- [53] Huang G, Liu Z, Weinberger KQ (2016) Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269.
- [54] He G, Ping A, Wang X, Zhu Y (2019) Alzheimer's disease diagnosis model based on three-dimensional full convolutional DenseNet. 2019 10th International Conference on Information Technology in Medicine and Education (ITME), pp. 13-17.
- [55] Zhang J, Gao Y, Gao Y, Munsell BC, Shen D (2016) Detecting anatomical landmarks for fast Alzheimer's disease diagnosis. *IEEE Trans Med Imaging* 35, 2524-2533.
- [56] Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. *Neuroimage* 11, 805-821.
- [57] Zhang X, Han L, Zhu W, Sun L, Zhang D (2022) An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE J Biomed Health Inform* 26, 5289-5297.
- [58] Shi J, Zheng X, Li Y, Zhang Q, Ying S (2018) Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE J Biomed Health Inform* 22, 173–183.
- [59] Zhang D, Wang Y, Zhou L, Yuan H, Shen D (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 55, 856-867.
- [60] Hazarika RA, Kandar D, Maji AK (2022) An experimental analysis of different deep learning based models for Alzheimer's disease classification using brain magnetic res-

onance images. J King Saud Univ Comput Inf Sci 34, 8576-8598.

- [61] Liu M, Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative (2014) Hierarchical fusion of features and classifier decisions for Alzheimer's disease diagnosis. *Hum Brain Mapp* 35, 1305–1319.
- [62] Liu X, Tosun D, Weiner MW, Schuff N, Alzheimer's Disease Neuroimaging Initiative (2013) Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. *Neuroimage* 83, 148–157.
- [63] Tong T, Wolz R, Gao Q, Guerrero R, Hajnal JV, Rueckert D, Alzheimer's Disease Neuroimaging Initiative (2014) Multiple instance learning for classification of dementia in brain MRI. *Med Image Anal* 18, 808–818.
- [64] Ossenkoppele R, van der Kant R, Hansson O (2022) Tau biomarkers in Alzheimer's disease: Towards implementation in clinical practice and trials. *Lancet Neurol* 21, 726–734.
- [65] Poulin SP, Dautoff R, Morris JC, Barrett LF, Dickerson BC (2011) Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry Res Neuroimaging* 194, 7-13.
- [66] Khatri U, Kwon GR (2022) Alzheimer's disease diagnosis and biomarker analysis using resting-state functional MRI functional brain network with multi-measures features and hippocampal subfield and amygdala volume of structural MRI. Front Aging Neurosci 14, 818-871.
- [67] Sjöbeck M, Englund E (2001) Alzheimer's disease and the cerebellum: A morphologic study on neuronal and glial changes. *Dement Geriatr Cogn Disord* 12, 211-218.
- [68] Liu Y, Liu D, Liu M, Li K, Shi Q, Wang C, Pan Z, Zhou L (2022) The microstructural abnormalities of cingulum was related to patients with mild cognitive impairment: A diffusion kurtosis imaging study. *Neurol Sci* 44, 171-180.

- [69] Lin FV, Ren P, Lo RY, Chapman BP, Jacobs A, Baran TM, Porsteinsson AP, Foxe JJ (2016) Insula and inferior frontal gyrus' activities protect memory performance against Alzheimer's disease pathology in old age. J Alzheimers Dis 55, 669-678.
- [70] Xu L, Wu X, Li R, Chen K, Long Z, Zhang J, Guo X, Yao L, Alzheimer's Disease Neuroimaging Initiative (2016) Prediction of progressive mild cognitive impairment by multi-modal neuroimaging biomarkers. J Alzheimers Dis 51, 1045–1056.
- [71] Sarasua I, Pölsterl S, Wachinger C (2022) Hippocampal representations for deep learning on Alzheimer's disease. *Sci Rep* 12, 8619.
- [72] Wang C, Li Y, Tsuboshita Y, Sakurai T, Goto T, Yamaguchi H, Yamashita Y, Sekiguchi A, Tachimori H, Alzheimer's Disease Neuroimaging Initiative (2022) A high-generalizability machine learning framework for predicting the progression of Alzheimer's disease using limited data. NPJ Digit Med 5, 43.
- [73] Nichol A, Dhariwal P (2021) Improved denoising diffusion probabilistic models. arXivPreprarXiv 2102.09672.
- [74] Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. arXivPreprarXiv 2006.11239.