



Quantifying the impact of Pyramid Squeeze Attention mechanism and filtering approaches on Alzheimer's disease classification

Bin Yan ^a, Yang Li ^a, Lin Li ^a, Xiaocheng Yang ^a, Tie-qiang Li ^b, Guang Yang ^{c,d}, Mingfeng Jiang ^{a,*}

^a School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, 310018, China

^b Department of Clinical Science, Intervention and Technology, Karolinska Institutet, 171 77, Stockholm, Sweden

^c Cardiovascular Research Centre, Royal Brompton Hospital, London, SW3 6NP, UK

^d National Heart and Lung Institute, Imperial College London, London, SW7 2AZ, UK



ARTICLE INFO

Keywords:

Alzheimer's disease
Pyramid squeeze attention mechanism
Image filtering
Fully convolutional network
Image classification

ABSTRACT

Brain medical imaging and deep learning are important foundations for diagnosing and predicting Alzheimer's disease. In this study, we explored the impact of different image filtering approaches and Pyramid Squeeze Attention (PSA) mechanism on the image classification of Alzheimer's disease. First, during the image preprocessing, we register MRI images and remove skulls, then apply median filtering, Gaussian blur filtering, and anisotropic diffusion filtering to obtain different experimental images. After that, we add the Squeeze and Excitation (SE) mechanism and Pyramid Squeeze Attention (PSA) mechanism to the Fully Convolutional Network (FCN) model respectively, to obtain each MRI image's corresponding feature information of disease probability map. Besides, we also construct Multi-Layer Perceptron (MLP) model's framework, combining feature information of disease probability map with age, gender, and Mini-Mental State Examination (MMSE) of each sample, to get the final classification performance of model. Among them, the accuracy of the MLP-C model combining anisotropic diffusion filtering with the Pyramid Squeeze Attention mechanism can reach 98.85%. The corresponding quantitative experimental results show that different image filtering approaches and attention mechanisms provide effective assistance for the diagnosis and classification of Alzheimer's disease.

1. Introduction

Alzheimer's disease (AD) is an irreversible brain degenerative disease that causes the brain to shrink and neuronal death, making patients lose memory and even lose their life self-care ability [1]. At present, there is no effective way to cure AD patients, so it is of important practical significance for early diagnosis and nursing of suspected AD patients [2]. With the continuous development of medical imaging technology, the application of medical imaging can provide a promising way to implement early diagnosis [3].

Magnetic resonance imaging (MRI) provides good information for brain research, and it helps to identify tissue with higher spatial resolution [4]. However, the original MRI image has various problems such as interference noise and blurred edge, which are not conducive to subsequent research and analysis. Therefore, some image filtering approaches were used to smooth, sharpen, and enhance the edge of original images, which can strengthen or remove certain features [5]. For example, median filtering [6], Gaussian blur filtering [7], and

anisotropic diffusion filtering [8,9] have been fully applied for image processing.

Recently, the deep learning method, as an alternative to the traditional methods, has been used for auxiliary diagnosis of Alzheimer's disease [10,11]. The deep learning method has wide coverage, good adaptability, and good portability characteristics, which can be used to find the hidden associations between these features of images, to realize classification and recognition [12]. In addition, deep learning has made breakthrough results in the fields of object detection, image segmentation, and image classification [13,14]. Of course, classification algorithms based on deep learning mainly include two stages of feature extraction and classification [15,16]. Among these feature extraction methods, the Histogram of Oriented Gradient (HOG) method was to extract features by calculating gradient direction histogram of the local area of the image [17,18], and Principal Component Analysis (PCA) [19] was designed to remove redundant features of the image and reduce the influence of background noise.

In order to make full use of effective features of the image, El-

* Corresponding author. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, 310018, PR China.

E-mail addresses: tie-qiang.li@ki.se (T.-q. Li), g.yang@imperial.ac.uk (G. Yang), m.jiang@zstu.edu.cn (M. Jiang).

Table 1

The detailed information of the experimental dataset.

Dataset	ADNI	
Research group	AD	NC
Number of samples	309	241
age (Mean \pm Std)	76.38 \pm 7.27	79.59 \pm 4.87
Gender (Male/Female)	182/127	101/140
MMSE (Mean \pm Std)	22.61 \pm 1.98	28.79 \pm 0.91

dahshan et al. [20] applied PCA and two-dimensional discrete wavelet transform (2D-DWT) to extract image features and then adopted K-Nearest Neighbor (KNN) and Back-Propagation Neural Network (BPNN) to classify MRI images. In addition, deep learning can use data privilege to extract potential features from the Region of Interest (ROI) of different image types for the Alzheimer's disease classification task. Recently, Liu et al. [21] extracted a set of potential features from 83 ROIs of MRI and PET images, then trained a multi-layer neural network consisting of multiple autoencoders and combined it with multi-modal features for image classification. Suk et al. [22] used a Stacked Autoencoder (SAE) to learn high-level features from multi-modal ROI features and then applied multi-core SVM combined with these high-level features to improve classification performance. The above studies provide new ideas and methods for image classification of Alzheimer's disease. Moreover, Spiking Neural Networks (SNNs) are new artificial neural networks that can mimic the biological behavior of neural networks [23], and contribute to the diagnosis of Alzheimer's disease.

In the field of computer vision, visual attention mechanisms can be used to focus on the key regions of an image, and to obtain high-level feature information of the image [24,25]. In recent years, attention mechanisms have emerged in different types, such as Self-Attention (SA) [26], Squeeze and Excitation (SE) and Convolutional Block Attention Module (CBAM) [27]. Attention mechanisms can enable Convolutional Neural Network (CNN) model to enhance the learning of critical information in feature maps. However, due to the limited local receptive field of the convolutional layer, the CNN model can only learn the dependence of the local position of the feature map. Therefore, Jie et al. [28] proposed Squeeze and Excitation Networks (SE-Net) to obtain global information of each feature map through global average pooling and then adopted a fully connected layer to learn the global dependence between feature maps. However, SE-Net cannot utilize the spatial information of feature maps at different scales, and even only considers channel attention but ignores spatial attention. Zhang et al. [29] proposed the Pyramid Squeeze Attention (PSA) mechanism, which can handle spatial information of multi-scale input feature maps and effectively establish long-term dependence between multi-scale channel attention. This PSA method can enhance the learning efficiency of detailed features, to improve the classification performance of models.

In this paper, we investigate different image filtering approaches and the Pyramid Squeeze Attention mechanism to assist the diagnosis and analysis of Alzheimer's disease and quantify their impact on image classification of Alzheimer's disease. First, we register MRI images and remove skulls, then obtain different experimental images through median filtering, Gaussian blur filtering, and anisotropic diffusion filtering respectively. After that, we add the SE mechanism and the Pyramid Squeeze Attention mechanism to Fully Convolutional Network (FCN) model respectively, to improve the classification performance of model. The main contributions of this paper are as follows: (1) We propose a new method for image classification of Alzheimer's disease based on the

Pyramid Squeeze Attention mechanism and Fully Convolutional Network. (2) In image preprocessing, we quantify the impact of different image filtering approaches on classification model performance. (3) We also evaluate the feasibility and effectiveness of the SE mechanism on FCN model, to highlight the efficiency of the Pyramid Squeeze Attention mechanism.

2. Materials and methods

2.1. Dataset

In order to prove the effectiveness of image filtering approaches and the Pyramid Squeeze Attention mechanism on the FCN model, this paper uses T1-weighted MRI images of Alzheimer's disease in the ADNI dataset for experiments. The detailed information of the experimental dataset is shown in Table 1. The experimental dataset selects individual scan images over 55 years old, including a total of 550 1.5T three-dimensional MRI images, of which 309 images from 84 AD patients and 241 images from 55 normal cognitive (NC) people. The MMSE information in each MRI image can objectively reflect the mental state of Alzheimer's disease patients and normal cognitive persons. This index was obtained by clinical researchers through mini-mental state examination.

2.2. Image preprocessing

First, we use the FLIRT tool in the FSL software package [30] to align the MRI image with the MNI152 public template. We transform the input image into a standard form of images according to the affine transformation function, to generate the image with the size of $182 \times 218 \times 182$. FLIRT applies coordinate rotation, translation, scaling, and shearing to register two images together. In addition, the cost function $O(w)$ takes the intensity difference between the input image and public template as an optimization objective, as shown in Equation (1).

$$O(w) = \sum_{i=1}^N (g(x'_i, w)) - f(x_i))^2 \quad (1)$$

where $f(x_i)$ represents the intensity of the public template, $g(x'_i)$ represents the intensity of the input image, x'_i represents the value of x_i after affine transformation.

After image registration, we use the BET tool to remove the neck, eyeball, and skull of the image [31]. Then we denoise and smooth MRI images through different image filtering approaches, such as median filtering, Gaussian blur filtering, and anisotropic diffusion filtering. We also control the voxel values of all images' brain external regions within a certain range to avoid the interference of background information. The flowchart of image preprocessing and images before and after skull removal are shown in Figs. 1 and 2 respectively.

2.3. Anisotropic diffusion filtering approach

The quality of magnetic resonance imaging is easily influenced by the inherent limitations of imaging equipment and external noise. The main noises in MRI images are thermal noise and physiological noise, which have a significant impact on subsequent analysis and processing. Therefore, we hope to reduce the influence of these noises through image filtering approaches.

Although traditional filtering, such as median filtering [32] and Gaussian blur filtering [33], can remove noise and smooth image, it damages the strength of original signal and is not possible to maintain



Fig. 1. The flowchart for image preprocessing.

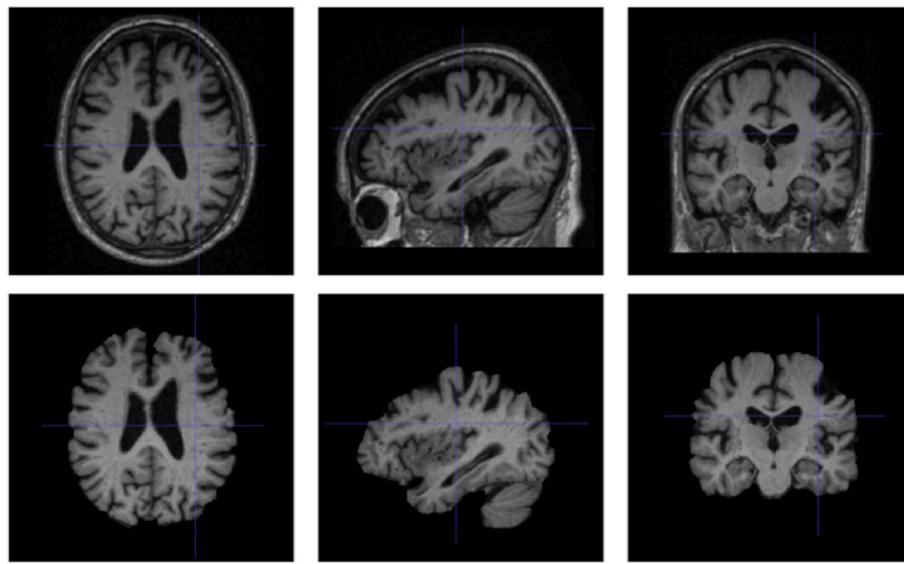


Fig. 2. Images before and after skull removal.

useful details of the image. Therefore, it is not feasible for further AD classification. As for anisotropic diffusion filtering, it can effectively retain feature information such as boundaries and lines of images [34]. It overcomes the defects of Gaussian blur filtering and decides whether to diffuse to the neighborhood according to the gray difference between each pixel and neighborhood pixels. Certainly, the greater the gray difference between the pixel and neighborhood pixels in a certain direction, the slower the diffusion speed in this direction. During the practical application, the rate of diffusion depends on the density gradient and anisotropic density gradient coefficient of the pixel.

Suppose the image is I , the formula of anisotropic diffusion filtering is shown in Equation (2).

$$I_{t+1} = I_t + \lambda(cN_{x,y}\nabla_N(I_t) + cS_{x,y}\nabla_S(I_t) + cE_{x,y}\nabla_E(I_t) + cW_{x,y}\nabla_W(I_t)) \quad (2)$$

t represents the number of iterations, ∇ represents the gradient operator, c represents the diffusion coefficient and λ is a fixed value.

The gradient formulas in four directions are shown in Equations (3)–(6).

$$\nabla_N(I_{x,y}) = I_{x,y-1} - I_{x,y} \quad (3)$$

$$\nabla_S(I_{x,y}) = I_{x,y+1} - I_{x,y} \quad (4)$$

$$\nabla_E(I_{x,y}) = I_{x-1,y} - I_{x,y} \quad (5)$$

$$\nabla_W(I_{x,y}) = I_{x+1,y} - I_{x,y} \quad (6)$$

The diffusion formulas in four directions are shown in Equations (7)–(10).

$$cN_{x,y} = e^{\left(\frac{-\|\nabla_N(I)\|^2}{K^2}\right)} \quad (7)$$

$$cS_{x,y} = e^{\left(\frac{-\|\nabla_S(I)\|^2}{K^2}\right)} \quad (8)$$

$$cE_{x,y} = e^{\left(\frac{-\|\nabla_E(I)\|^2}{K^2}\right)} \quad (9)$$

$$cW_{x,y} = e^{\left(\frac{-\|\nabla_W(I)\|^2}{K^2}\right)} \quad (10)$$

K represents the thermal conductivity, which is used to control the sensitivity of image edges. When the thermal conductivity K and the number of iterations t are larger, the image is smoother. In our experiments, we set the thermal conductivity K to 0.9 and the number of iterations t to 5.

The MRI images processed using median filtering, Gaussian blur filtering, and anisotropic diffusion filtering are shown in Fig. 3 respectively.

2.4. AD classification methods

The overall workflow of our proposed Alzheimer's disease identification technique is shown in Fig. 4, which includes image preprocessing, model training, and classification. After the original image registration, we remove the irrelevant skull parts. Then, we obtain different experimental images through median filtering, Gaussian blur filtering, and anisotropic diffusion filtering respectively, as shown in Fig. 4(A). After

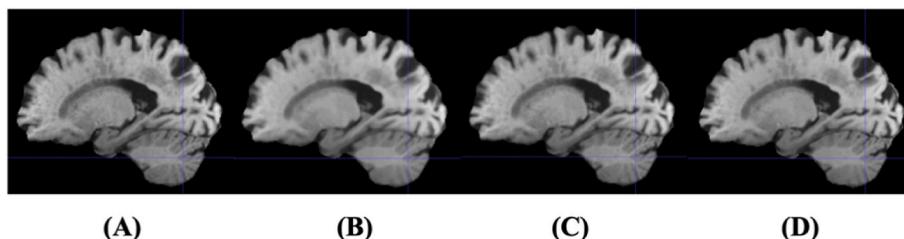


Fig. 3. (A) The image after skull removal. (B) The image after denoising by median filtering. (C) The image after denoising by Gaussian blur filtering. (D) The image after denoising by anisotropic diffusion filtering.

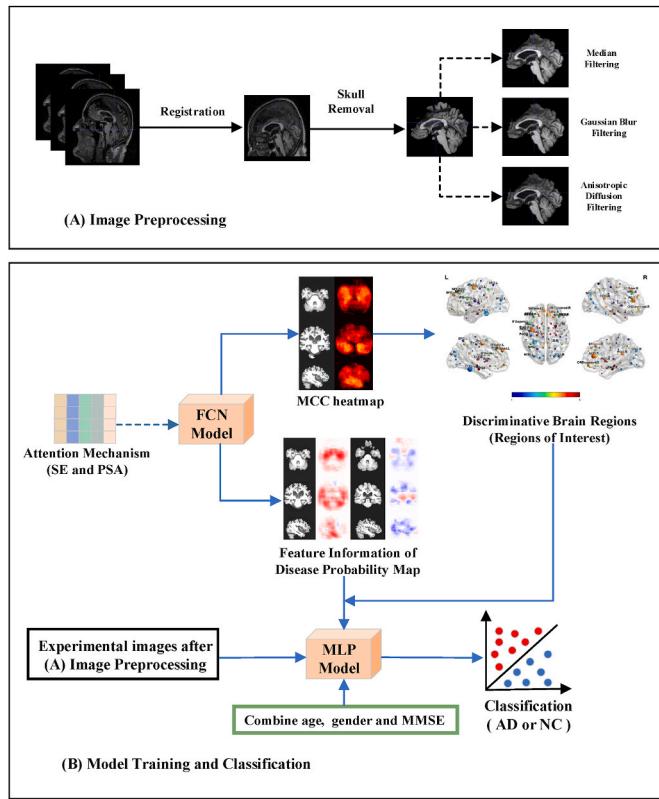


Fig. 4. The overall workflow of our proposed Alzheimer's disease identification technique. (A) Image Preprocessing. (B) Model Training and Classification.

FCN model training, we can get MCC heatmap and feature information of disease probability map. Moreover, we select the discriminative brain regions (regions of interest) for MLP models according to MCC heatmap and combine feature information of disease probability map with age, gender, and MMSE of samples to classify AD and NC, as shown in Fig. 4(B).

2.4.1. Squeeze and Excitation (SE) mechanism

The attention mechanism is an important research direction in the field of computer vision. The SE mechanism obtains global information for each feature map through global average pooling and then adopts a fully connected layer to learn global dependencies between feature maps. First, the SE mechanism performs the Squeeze operation on a feature map obtained by convolution to get channel-level global features. Then it performs the Excitation operation on the global feature to learn the relationship between each channel, and to obtain the weights of different channels. Finally, it multiplies the input feature map with the weights to get a new weighted feature map.

$$X \rightarrow U, X \in \mathbb{R}^{C \times D \times H \times W}, U \in \mathbb{R}^{C' \times D' \times H' \times W'} \quad (11)$$

Where X represents the input feature map, U represents the output feature map. The convolution kernel $V = [v_1, v_2, v_3, \dots, v_c]$ v_c represents the c -th convolution kernel. The output feature map $U = [u_1, u_2, u_3, \dots, u_c]$ and $u_c = v_c * X = \sum_{s=1}^C v_c^s * x^s$. Where $*$ represents a convolution operation and v_c^s represents the convolution kernel of s -th channel. By mixing spatial features of each channel with spatial relationships learned by the convolution kernel, models can directly learn the feature relationship between channels.

2.4.2. Pyramid Squeeze Attention (PSA) mechanism

As we know, attention mechanisms can significantly improve the performance of Convolutional Neural Networks (CNN) models. However, there are still two challenging problems that need to be solved. One is how to effectively obtain and utilize spatial information of feature maps at different scales. The other is that channel attention or spatial attention can only effectively capture local information but cannot establish long-term dependence. To solve these two problems, a novel and efficient Pyramid Squeeze Attention (PSA) came into being, as shown in Fig. 5. The PSA mechanism can handle spatial information of multi-scale input feature maps, and can effectively establish long-term dependence between multi-scale channel attention. It can learn more abundant multi-scale feature representations and adaptively recalibrate the weights of multi-dimensional channel attention to fully extract spatial information from multi-scale feature maps and realize feature interaction for cross-dimensional channel attention.

The PSA module mainly includes the Squeeze and Concat (SPC) module and the SE Weight module. C , D , H , and W respectively represent the channel number, depth, height, and width of the input feature map. Among them, the SPC module is the basic operator for realizing multi-scale feature extraction. The Squeeze operation represents compressing the number of input channels equally, and the Concat operation represents connecting the feature map of different channels. As shown in Fig. 6, we extract spatial information of input feature maps in a multi-branch manner, where the number of input channels for each branch is C and the number of output channels is $C' = C/S$. In pyramid structure, multi-scale convolution kernels are used to generate different spatial resolutions and depths. By compressing the number of input tensor's channels, spatial information of each channel's feature map at different scales can be effectively extracted.

The feature map F_i of each branch has the same output channel number $C' = C/S$, $i = 0, 1, \dots, S - 1$. For each branch, it can independently learn multi-scale spatial information and establish cross-channel interactions in a local manner. In order to handle input tensors of different kernel scales without increasing the amount of computation, the method of group convolution is introduced and applied to the convolution kernel.

$$G = 2^{\frac{K-1}{2}} \quad (12)$$

Here, K represents multi-scale kernel size and G represents group size.

$$F_i = \text{Conv}(k_i \times k_i \times k_i, G_i)(X) \quad (13)$$

$X \in \mathbb{R}^{C \times D \times H \times W}$ represents the input feature map, $F_i \in \mathbb{R}^{C' \times D \times H \times W}$

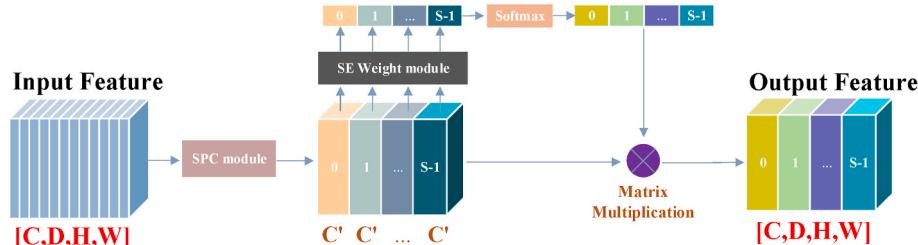


Fig. 5. The overall framework of the Pyramid Squeeze Attention (PSA) Module.

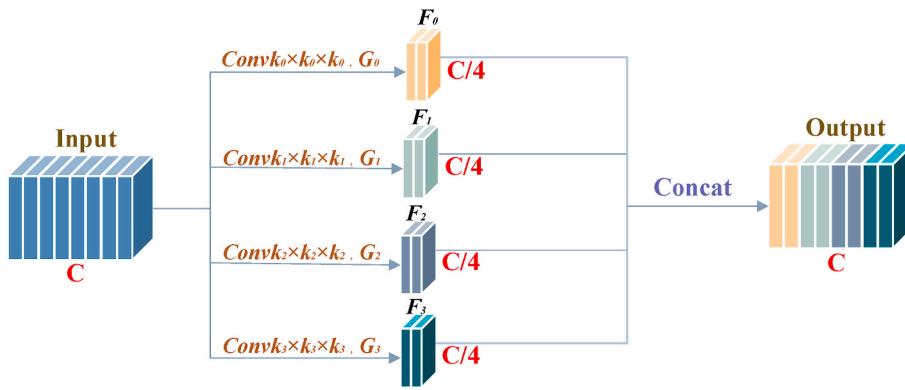


Fig. 6. The detailed information of the Squeeze and Concat (SPC) module with $S = 4$.

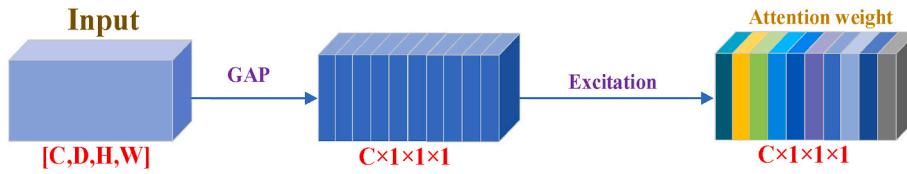


Fig. 7. The SE Weight module with Global Average Pooling (GAP) and Excitation operations.

represents the generator function for multi-scale feature maps of each branch.

$$k_i = 2 \times (i + 1) + 1 \quad (14)$$

$$G_i = 2^{\frac{k_i-1}{2}} \quad (15)$$

Where k_i represents the i -th kernel size, G_i represents the i -th group size. And when k_0 is 3, G_0 is the default value of 1.

In our experiments, the channel number of the input feature map in the PSA module is $C = 160$, and the channel number of the SPC module is $C' = C/4 = 40$. Because C' must be divisible by G_i , we set $k_3 = k_2 = 7$, $G_3 = G_2 = 8$.

$$F = \text{Cat}([F_0, F_1, \dots, F_{S-1}]) \quad (16)$$

$F \in \mathbb{R}^{C \times D \times H \times W}$ is the final multi-scale feature map, is the superposition of each branch feature map F_i . We can obtain the attention weight vector Z_i of different scales by extracting weight information of channel attention from a multi-scale processed feature map F_i .

$$Z_i = \text{SE Weight}(F_i) \quad (17)$$

Among them, $Z_i \in \mathbb{R}^{C' \times 1 \times 1 \times 1}$ is the attention weight vector. The SE Weight module is used to obtain the corresponding attention weights of input feature maps at different scales, as shown in Fig. 7.

$$g_c = \frac{1}{D \times H \times W} \sum_{i=1}^D \sum_{j=1}^H \sum_{n=1}^W x_c(i, j, n) \quad (18)$$

$$w_c = \sigma(W_1 \delta(W_0(g_c))) \quad (19)$$

Where g_c represents GAP operator, w_c represents the attention weight of c -th channel in the SE Weight module, δ represents ReLU operation, and σ represents excitation function. $W_0 \in \mathbb{R}^{C \times \frac{C}{2}}$, $W_1 \in \mathbb{R}^{\frac{C}{2} \times C}$ represent two fully connected layers. Through these two fully connected layers, the linear information between channels can be combined more effectively. Therefore, it is beneficial to channel information interaction between high-dimensional and low-dimensional. In addition, we can assign weight to channel after channel interaction by excitation function σ , so that information can be extracted more efficiently.

Table 2

The pseudo code of the Pyramid Squeeze Attention (PSA) mechanism.

```

Input: X with shape = [B, C, D, H, W]
1: Procedure PSA(X)
2: ∇ Step 1 → generating feature maps with different scales  $F_i$ 
3: for  $i = 0, 1, 2, 3$  do
4:  $F_i = \text{conv}_i(X)$  with shape = [B, C/4, D, H, W]
5: end for
6: ∇ Step 2 → concatenating different  $F_i$  to obtain multi-scale feature map  $F$ 
7:  $F = \text{torch.cat}([F_0, F_1, F_2, F_3], \text{dim} = 1)$  with shape = [B, C, D, H, W]
8: ∇ Step 3 → calculating and recalibrating the attention weight  $Z_i$ 
9: for  $i = 0, 1, 2, 3$  do
10:  $Z_i = \text{SE Weight}(F_i)$  with shape = [B, C/4, 1, 1, 1]
11:  $att_i = \text{Softmax}(Z_i, \text{dim} = 1)$  with shape = [B, C/4, 1, 1, 1]
12: end for
13: ∇ Step 4 → concatenating different  $att_i$  to obtain attention weight  $att$ 
14:  $att = \text{torch.cat}([att_0, att_1, att_2, att_3], \text{dim} = 1)$  with shape = [B, C, 1, 1, 1]
15: ∇ Step 5 → multiplying  $att$  with  $F$  to obtain the output  $Y$ 
16:  $Y = att * F$  with shape = [B, C, D, H, W]
17: end procedure
Output: Y with shape = [B, C, D, H, W]
```

In order to fuse cross-dimensional vectors and realize the interaction of attention information without destroying original channel attention vectors, we obtain entire multi-scale channel attention vectors in a concatenated manner.

$$att_i = \text{Softmax}(Z_i) = \frac{\exp(Z_i)}{\sum_{i=0}^{S-1} \exp(Z_i)} \quad (20)$$

$$att = att_0 \oplus att_1 \oplus \dots \oplus att_{S-1} \quad (21)$$

On the other hand, to make the sum of all channel attention weights equal to 1, we recalibrate the attention weights of multi-scale channels using Softmax. Z_i is the attention weight vector generated from the feature map F_i , and \oplus represents the connection operator. After that, we multiply the concatenated multi-scale channel attention weight att with the superimposed multi-scale feature map F to obtain the final feature map Y .

$$Y = F \odot att \quad (22)$$

From the detailed structural analysis of the PSA module, it fuses

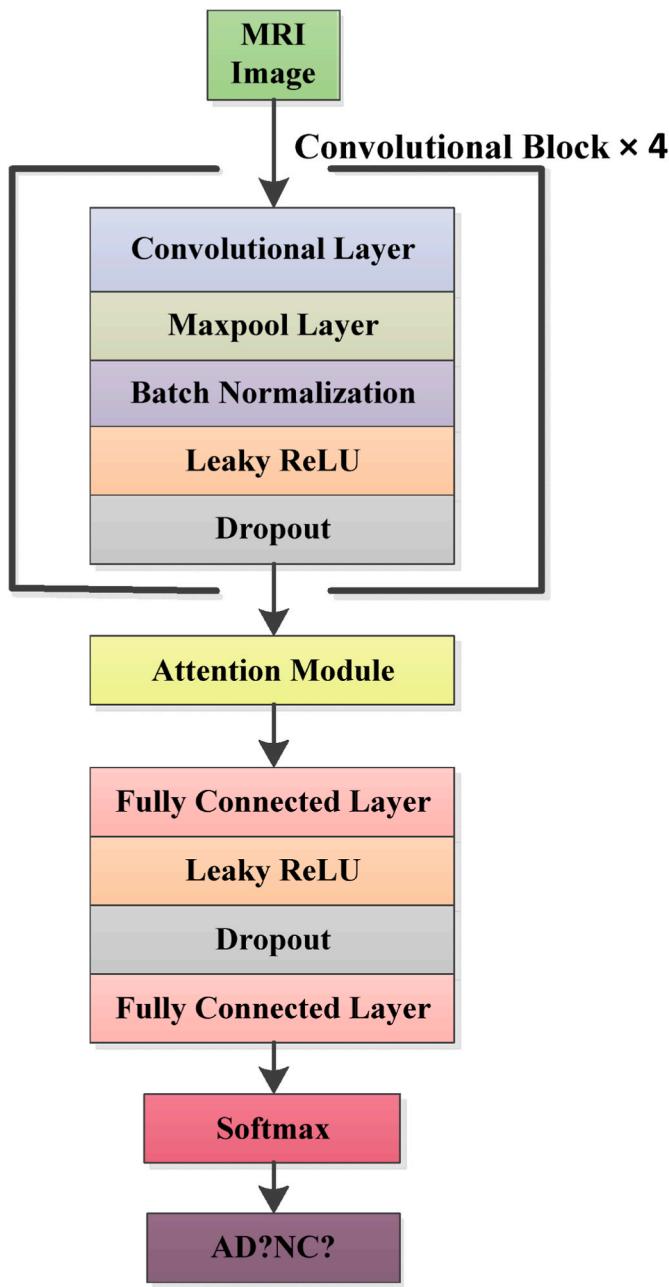


Fig. 8. The framework of the FCN model.

information at different scales and results in better pixel-level attention. Furthermore, it effectively extracts finer-grained multi-scale spatial information and establishes longer-distance channel dependence. In all, the pseudo code of the Pyramid Squeeze Attention mechanism is provided in Table 2, which can aid comprehension of the PSA method.

2.4.3. Framework of the model

FCN model consists of four convolutional blocks and two fully connected layers. Among them, the convolutional block includes a 3D convolutional layer, 3D maxpool layer, 3D batch normalization, Leaky ReLU, and Dropout, as shown in Fig. 8. The last two fully connected layers play an important role in improving the efficiency of the model, and the network is trained by randomly initializing weights. As shown in Fig. 9, we adopt a method of randomly sampling Patch from 3D-MRI images to train the FCN model, and the size of the random sampling Patch is $47 \times 47 \times 47$.

The specific parameter settings of the FCN model and the changes in output Patch size are shown in Table 3. Because convolution operation reduces the output size of the input layer, each Patch will generate two scalar values after being trained by the FCN model. Then they are converted into Alzheimer's disease probability and normal cognitive probability of the corresponding pixel by activation function Softmax. The risk probability values of Alzheimer's disease can show the disease state

Table 3
Specific parameter settings of the FCN model.

Input Layer	Detailed Description	Output Patch Size
Input		(1,47,47,47)
3D Convolutional Layer1	channel 20, kernel 4, stride 1, padding 0	(20,44,44,44)
3D Maxpool Layer1	kernel 2, stride 1, padding 0	(20,43,43,43)
3D Batch Normalization1	eps = 1e-05, momentum = 0.1, affine = True	
Leaky ReLU1; Dropout1	Negative slope = 0.01; P = 0.1	
3D Convolutional Layer2	channel 40, kernel 4, stride 1, padding 0	(40,40,40,40)
3D Maxpool Layer2	kernel 2, stride 2, padding 0	(40,20,20,20)
3D Batch Normalization2	eps = 1e-05, momentum = 0.1, affine = True	
Leaky ReLU2; Dropout2	Negative slope = 0.01; P = 0.1	
3D Convolutional Layer3	channel 80, kernel 3, stride 1, padding 0	(80,18,18,18)
3D Maxpool Layer3	kernel 2, stride 2, padding 0	(80,9,9,9)
3D Batch Normalization3	eps = 1e-05, momentum = 0.1, affine = True	
Leaky ReLU3; Dropout3	Negative slope = 0.01; P = 0.1	
3D Convolutional Layer4	channel 160, kernel 3, stride 1, padding 0	(160,7,7,7)
3D Maxpool Layer4	kernel 2, stride 1, padding 0	(160,6,6,6)
3D Batch Normalization4	eps = 1e-05, momentum = 0.1, affine = True	
Leaky ReLU4; Dropout4	Negative slope = 0.01; P = 0.1	
Fully Connected Layer1	channel 30, kernel 6, stride 1, padding 0	(30,1,1,1)
Leaky ReLU5; Dropout5	Negative slope = 0.01; P = 0.1	
Fully Connected Layer2	channel 2, kernel 1, stride 1, padding 0	(2,1,1,1)

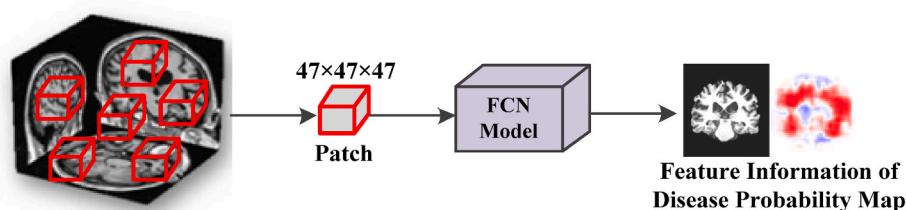


Fig. 9. Randomly sampling 3D-MRI image's Patch for training FCN model.

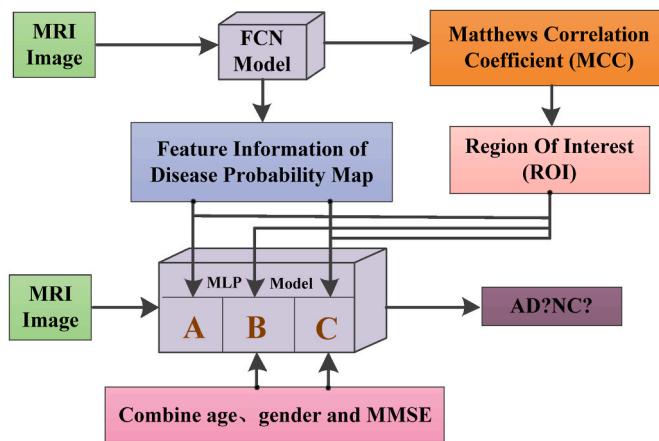


Fig. 10. The framework of MLP models.

of local brain structure, and the corresponding feature information of disease probability map will be used as auxiliary information for the MLP model.

Besides, we also build the MLP model's framework, as shown in Fig. 10. MLP model consists of two fully connected layers, batch normalization, Leaky ReLU, and Dropout. In the image classification experiment of the MLP model, we select the probability value of Alzheimer's disease from feature information of disease probability map, choose the region of interest (ROI) according to the MCC heatmap of the FCN model, and combine with age, gender, MMSE of samples.

Among them, MCC values can reflect the overall classification performance of FCN models, and MCC heatmap can observe that FCN models have a higher classification accuracy for certain pixel positions of 3D-MRI images. Therefore, we choose these high-accuracy regions as regions of interest (ROI) for all MLP models. As shown in Fig. 10, the MLP-A model indicates that only feature information of disease probability map of the FCN model is used to classify MRI images, the MLP-B model indicates that only the age, gender, and MMSE of samples are used to classify MRI images, and the MLP-C model indicates combining feature information of disease probability map with age, gender, MMSE of samples to classify MRI images.

2.5. Experimental settings and evaluation criteria

The experiments in this paper adopt Pytorch deep learning framework (<https://pytorch.org/>) and GeForce RTX 3090 GPU processor. The FCN model uses an Adam optimizer and cross-entropy loss function. In addition, the experimental parameters are set by parameter sweeping

method, and the best ones are selected as appropriate parameters of model, i.e. batch size is 10, the learning rate is 0.0001, and the number of training iterations is 3000. Once the validation set is verified for every 20 iterations, then the optimal model and weights are saved. Finally, the optimal model is tested by the test set to obtain the classification performance of the FCN model and feature information of disease probability map.

We use accuracy and Matthews correlation coefficient (MCC) [35] to evaluate the classification performance of the FCN model. In addition, we also record the age, gender, and MMSE of all Alzheimer's disease patients and normal cognitive persons, as input for the MLP model's classification experiments. For the MLP models, we use accuracy (marked as Accu), sensitivity (marked as Sens), specificity (marked as Spec), F1 score, and MCC to evaluate their classification performance. F1 score is an indicator used to measure two-category models in statistics. It takes into account the precision and recall of the classification model at the same time and can be regarded as the harmonic average of the model's precision and recall. And MCC comprehensively considers true positives, true negatives, false positives, and false negatives, and is a relatively balanced indicator in deep learning [36]. In order to ensure the accuracy and reliability of experiments, we use five-fold cross-validation for the FCN model. Among them, the training set includes 330 3D-MRI images, and both the validation set and the test set include 110 3D-MRI images. Of course, they are independent of each other. Based on each different feature information of disease probability map obtained by the FCN model, the MLP models repeat experiments three times. The final classification performance of the MLP models is represented by mean and standard deviation.

$$Accu = \frac{TP + TN}{TP + TN + FP + FN} \quad (23)$$

$$Sens = \frac{TP}{TP + FN} \quad (24)$$

$$Spec = \frac{TN}{TN + FP} \quad (25)$$

$$F1 = \frac{2TP}{2TP + FN + FP} \quad (26)$$

$$MCC = \frac{TP \times TN - FP \times FN}{[(TP + FP)(TP + FN)(TN + FP)(TN + FN)]^{0.5}} \quad (27)$$

Among them, True Positives (TP) represent correct predictions of positive samples, True Negatives (TN) represent correct predictions of negative samples, False Negatives (FN) represent false predictions of positive samples, and False Positives (FP) represent false predictions of negative samples.

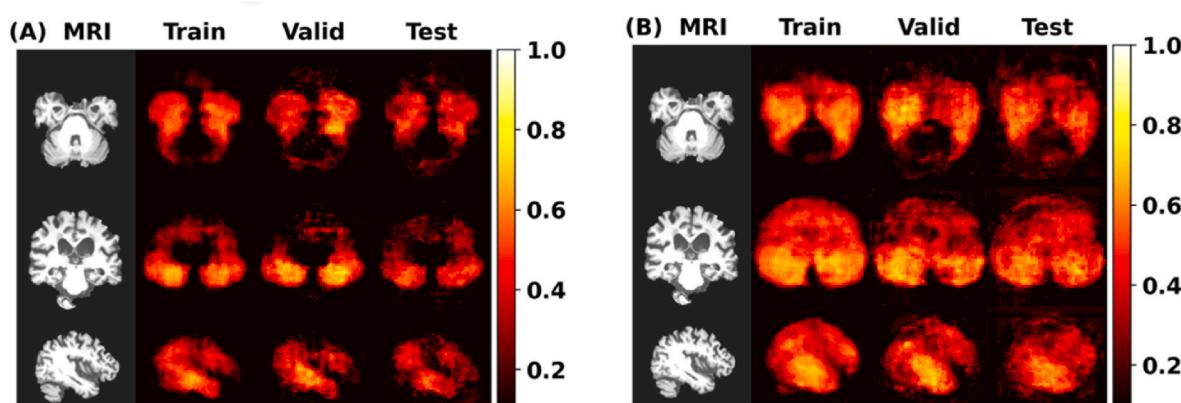


Fig. 11. (A) represents MCC heatmap of FCN model without the Pyramid Squeeze Attention mechanism. (B) represents MCC heatmap of FCN model with the Pyramid Squeeze Attention mechanism.

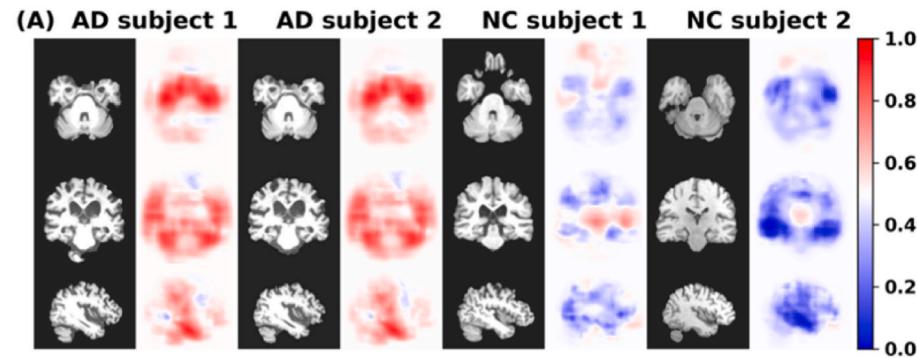
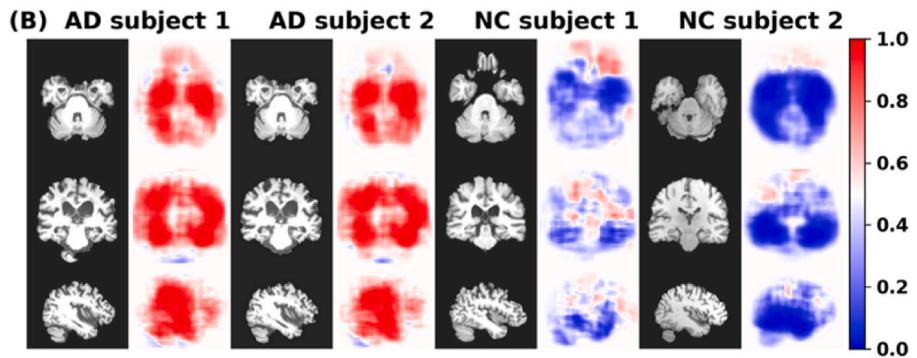


Fig. 12. (A) represents the feature information of disease probability map without the Pyramid Squeeze Attention mechanism. Among them, the first two samples were clinically diagnosed as Alzheimer's disease patients, and the last two samples were clinically confirmed as normal cognitive persons. Red indicates the risk probability of Alzheimer's disease is more than 0.5, and blue indicates less than 0.5. (B) represents the feature information of disease probability map with the Pyramid Squeeze Attention mechanism.



3. Experiments and results

3.1. Classification performance of FCN model

MCC values can reflect the overall classification performance of FCN model, and MCC heatmap can observe that some pixel positions of brain images have higher MCC values. The size of MCC heatmap obtained from the FCN model is $46 \times 55 \times 46$, while the size of MRI image is $182 \times 218 \times 182$. In order to facilitate observation, we magnify the MCC heatmap four times and visually display it with the MRI image in one picture at the same time, as shown in Fig. 11.

The MCC heatmap of FCN model without and with the Pyramid Squeeze Attention mechanism is presented in Fig. 11(A) and (B) respectively, in which brain images in axial, coronal, and sagittal directions are provided for comparison. It can be seen that the classification performance of the FCN model is improved after adding the

Pyramid Squeeze Attention mechanism, which is helpful for the diagnosis and classification of Alzheimer's disease.

We correspondingly show feature information of disease probability map of Alzheimer's disease patients and normal cognitive persons, as shown in Fig. 12. The (A) and (B) in Fig. 12 respectively represent feature information of disease probability map without and with the Pyramid Squeeze Attention mechanism. Of course, images of (A) and (B) in Fig. 12 correspond to the same Alzheimer's disease patient and the same normal cognitive person in turn. Red and blue indicate the risk probability of Alzheimer's disease in brain regions, with 0.5 as the dividing line.

Feature information of disease probability map can highlight brain regions of Alzheimer's disease at high risk. The results in Fig. 12 show that the difference in risk probability between Alzheimer's disease patients and normal cognitive persons becomes larger. The Pyramid Squeeze Attention mechanism can enhance the discrimination ability of

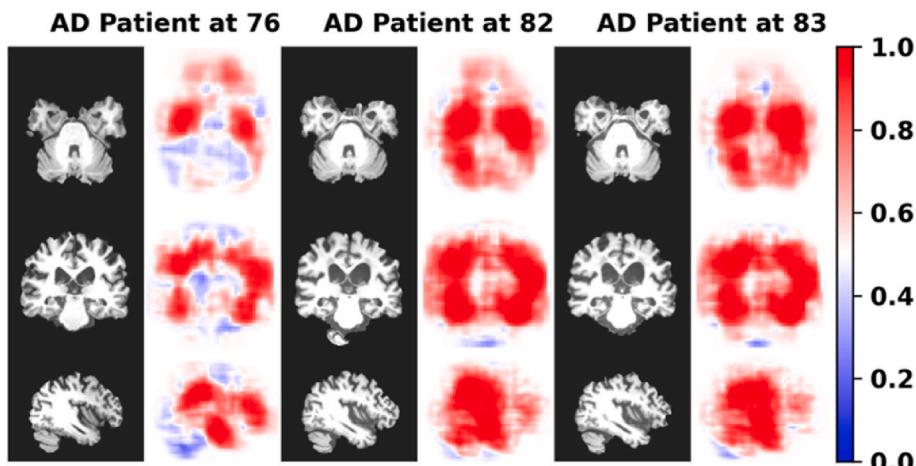


Fig. 13. Risk probability in the brain's local structure of Alzheimer's disease patient at different age stages.

Table 4

The impact of different image filtering approaches on FCN model.

FCN	Accu	MCC
Original Image	0.58575 ± 0.0062	0.18425 ± 0.0123
Skull Removal	0.58926 ± 0.0135	0.19332 ± 0.0192
Median Filtering	0.60315 ± 0.0144	0.20652 ± 0.0534
Gaussian Blur Filtering	0.61072 ± 0.0130	0.21048 ± 0.0475
Anisotropic Diffusion Filtering	0.61549 ± 0.0138	0.21507 ± 0.0278

Table 5

The impact of SE mechanism and PSA mechanism on FCN model.

FCN	Accu	MCC
Original Image + SE	0.61124 ± 0.0079	0.21325 ± 0.0152
Original Image + PSA	0.61578 ± 0.0118	0.21483 ± 0.0359
Anisotropic Diffusion Filtering + SE	0.62175 ± 0.0297	0.22185 ± 0.0479
Anisotropic Diffusion Filtering + PSA	0.62513 ± 0.0093	0.22309 ± 0.0223

the model and help distinguish Alzheimer's disease patients from normal cognitive people.

At the same time, in order to reflect the change in the brain's local structure of Alzheimer's disease patients with disease time. We choose brain images of the same patient at different age stages and visualize feature information of disease probability map, as shown in Fig. 13. The results in Fig. 13 show that, with the increase of disease time, the risk probability of the brain's local structure in Alzheimer's disease patients gradually increases. In addition, the hippocampus (HIP) and the parahippocampal gyrus (PHG) are always at high risk.

We use accuracy and MCC to evaluate the classification performance of FCN model. The classification index is calculated by taking each pixel of the 3D-MRI image as a separate sample. After FCN model training, each pixel of the 3D-MRI image will generate a predicted probability

value of Alzheimer's disease, then compare it with the input label and mark the corresponding pixel as TP, TN, FP, and FN. In addition, the impacts of different image filtering approaches and attention mechanisms on FCN model performance are shown in Tables 4 and 5.

The results in Table 4 show differences in different image filtering approaches. Among them, after removing skulls from MRI images, the accuracy and MCC of the FCN model can be improved in a small range. In addition, anisotropic diffusion filtering can retain detailed features of images better than median filtering and Gaussian blur filtering, which is helpful for the classification task of FCN model. The results in Table 5 show that attention mechanisms can greatly improve the classification performance of FCN model for different experimental images. The changes of FCN model's accuracy and MCC are shown in Figs. 14 and 15. Furthermore, the Pyramid Squeeze Attention mechanism can extract more useful feature information from images than the SE mechanism, produce better pixel-level attention and make an important contribution to the diagnosis of Alzheimer's disease.

Table 6

The impact of different image filtering approaches on MLP-A model.

MLP-A	Accu	Sens	Spec	F1	MCC
Original Image	0.8606 ± 0.0287	0.8280 ± 0.0240	0.9056 ± 0.0339	0.8529 ± 0.0359	0.7274 ± 0.0465
Skull Removal	0.8615 ± 0.0279	0.8397 ± 0.0472	0.9050 ± 0.0371	0.8622 ± 0.0230	0.7351 ± 0.0534
Median Filtering	0.8630 ± 0.0186	0.8349 ± 0.0508	0.9092 ± 0.0447	0.8680 ± 0.0139	0.7412 ± 0.0327
Gaussian Blur Filtering	0.8667 ± 0.0199	0.8343 ± 0.0510	0.9108 ± 0.0346	0.8721 ± 0.0173	0.7534 ± 0.0402
Anisotropic Diffusion Filtering	0.8722 ± 0.0168	0.8386 ± 0.0356	0.9134 ± 0.0387	0.8814 ± 0.0206	0.7581 ± 0.0446

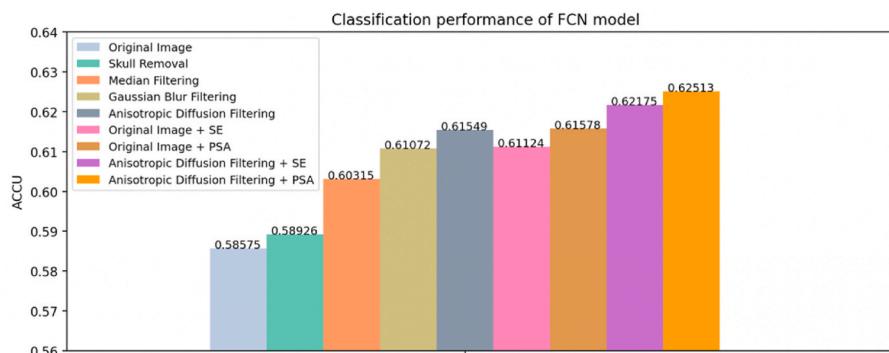


Fig. 14. The changes of FCN model's accuracy.

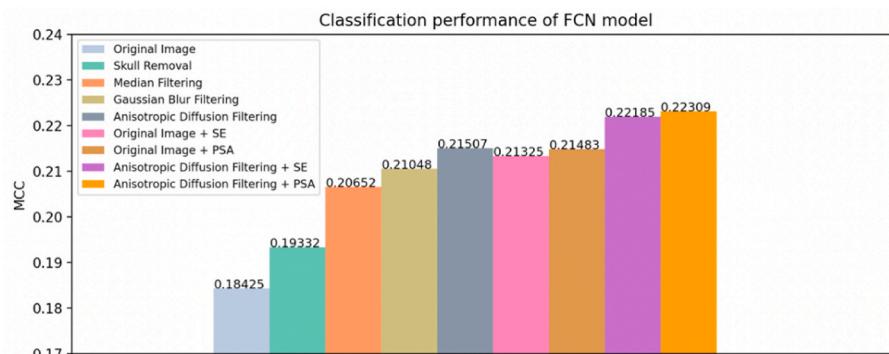


Fig. 15. The changes of FCN model's MCC.

Table 7

The impact of different image filtering approaches on MLP-B model.

MLP-B	Accu	Sens	Spec	F1	MCC
Original Image	0.9697 ± 0.0178	0.9645 ± 0.0313	0.9412 ± 0.0293	0.9786 ± 0.0186	0.9509 ± 0.0343
Skull Removal	0.9752 ± 0.0150	0.9678 ± 0.0304	0.9433 ± 0.0308	0.9765 ± 0.0134	0.9515 ± 0.0259
Median	0.9782 ± 0.0148	0.9660 ± 0.0170	0.9518 ± 0.0295	0.9814 ± 0.0125	0.9532 ± 0.0296
Gaussian Blur	0.9796 ± 0.0143	0.9683 ± 0.0203	0.9589 ± 0.0384	0.9805 ± 0.0162	0.9541 ± 0.0386
Anisotropic	0.9815 0.9687	0.9594	0.9816	0.9566	
Diffusion	± 0.0175	± 0.0198	± 0.0417	± 0.0144	± 0.0394
Filtering					

Table 8

The impact of different image filtering approaches on MLP-C model.

MLP-C	Accu	Sens	Spec	F1	MCC
Original Image	0.9715 ± 0.0131	0.9702 ± 0.0138	0.9597 ± 0.0289	0.9790 ± 0.0115	0.9529 ± 0.0265
Skull Removal	0.9776 ± 0.0080	0.9757 ± 0.0146	0.9603 ± 0.0187	0.9795 ± 0.0141	0.9538 ± 0.0159
Median	0.9785 ± 0.0056	0.9732 ± 0.0040	0.9608 ± 0.0203	0.9820 ± 0.0112	0.9541 ± 0.0092
Gaussian Blur	0.9802 ± 0.0093	0.9741 ± 0.0091	0.9615 ± 0.0164	0.9823 ± 0.0080	0.9550 ± 0.0106
Anisotropic	0.9832 ± 0.0171	0.9751 ± 0.0098	0.9613 ± 0.0130	0.9858 ± 0.0054	0.9572 ± 0.0128
Diffusion					
Filtering					

Table 9

The impact of SE mechanism and PSA mechanism on MLP-A model.

MLP-A	Accu	Sens	Spec	F1	MCC
Original Image	0.8758 ± 0.0326	0.8357 ± 0.0293	0.9142 ± 0.0493	0.8767 ± 0.0282	0.7518 ± 0.0312
Original Image	0.8794 ± 0.0489	0.8397 ± 0.0452	0.9167 ± 0.0227	0.8795 ± 0.0131	0.7553 ± 0.0116
Anisotropic	0.8800 ± 0.0238	0.8416 ± 0.0484	0.9206 ± 0.0580	0.8901 ± 0.0212	0.7612 ± 0.0309
Diffusion					
Filtering + SE					
Anisotropic	0.8897 ± 0.0149	0.8445 ± 0.0551	0.9248 ± 0.0349	0.8939 ± 0.0121	0.7628 ± 0.0270
Diffusion					
Filtering + PSA					

Table 10

The impact of SE mechanism and PSA mechanism on MLP-B model.

MLP-B	Accu	Sens	Spec	F1	MCC
Original Image	0.9701 ± 0.0106	0.9652 ± 0.0224	0.9511 ± 0.0283	0.9784 ± 0.0101	0.9542 ± 0.0243
Original Image	0.9727 ± 0.0073	0.9685 ± 0.0080	0.9583 ± 0.0171	0.9793 ± 0.0066	0.9545 ± 0.0084
Anisotropic	0.9818 ± 0.0213	0.9691 ± 0.0125	0.9589 ± 0.0469	0.9818 ± 0.0167	0.9571 ± 0.0349
Diffusion					
Filtering + SE					
Anisotropic	0.9836 ± 0.0165	0.9704 ± 0.0091	0.9617 ± 0.0366	0.9821 ± 0.0182	0.9575 ± 0.0433
Diffusion					
Filtering + PSA					

3.2. Classification performance of MLP models

We conduct multiple experiments on MLP models, then express the final classification performance in terms of mean and standard deviation. In Tables 6–8, we summarize differences in classification performance of all MLP models (including the MLP-A model, the MLP-B model, and the MLP-C model) with different image filtering approaches.

Table 11

The impact of SE mechanism and PSA mechanism on MLP-C model.

MLP-C	Accu	Sens	Spec	F1	MCC
Original Image	0.9812 ± 0.0221	0.9706 ± 0.0079	0.9589 ± 0.0361	0.9802 ± 0.0181	0.9556 ± 0.0374
Original Image	0.9817 ± 0.0104	0.9712 ± 0.0119	0.9608 ± 0.0136	0.9819 ± 0.0073	0.9561 ± 0.0177
Anisotropic	0.9842 ± 0.0157	0.9743 ± 0.0086	0.9617 ± 0.0303	0.9851 ± 0.0166	0.9581 ± 0.0285
Diffusion					
Filtering + SE					
Anisotropic	0.9885 ± 0.0122	0.9765 ± 0.0078	0.9644 ± 0.0417	0.9866 ± 0.0143	0.9607 ± 0.0395
Diffusion					
Filtering + PSA					

The results in Tables 6–8 show the impact of different image filtering approaches and skull removal on MLP models. Among them, image filtering approaches and skull removal can improve the classification performance of MLP models in a small range. In addition, the fusion model MLP-C combines feature information of disease probability map with age, gender, and MMSE of samples, which has better classification performance and keen discrimination ability of images. We also summarize differences in classification performance of all MLP models after adding the SE mechanism and the Pyramid Squeeze Attention (PSA) mechanism respectively, as shown in Tables 9–11.

The results in Tables 9–11 show the impact of SE mechanism and Pyramid Squeeze Attention mechanism on MLP models. The changes of MLP models' accuracy and MCC are shown in Figs. 16 and 17, attention mechanisms can effectively improve the classification performance of MLP models for different experimental images. Moreover, the Pyramid Squeeze Attention mechanism can obtain detailed features of the image better than the SE mechanism and extract spatial information from multi-scale feature maps more fully.

And we summarize ROC-AUC values of MLP models in Table 12, to objectively judge the influence of anisotropic diffusion filtering and Pyramid Squeeze Attention mechanism.

The difference in ROC-AUC values of MLP-B and MLP-C models is small, we only show the ROC curves of MLP-A model in different conditions. From Table 12 and Fig. 18, it can be seen that anisotropic diffusion filtering and the Pyramid Squeeze Attention mechanism have a great effect on ROC-AUC values of MLP-A model.

In addition, sensitivity can evaluate the classification performance of the model and measure the discriminative ability of the model to positive samples, which represents the proportion of all positive samples being paired. It can be seen from the experimental results in Tables 6–11 that, anisotropic diffusion filtering, the SE mechanism, and the Pyramid Squeeze Attention mechanism can effectively improve the sensitivity of MLP models. This shows that our proposed methods have a high discriminatory ability for images of Alzheimer's disease patients.

4. Discussion

From Table 4 and Tables 6–8, it can be found that the anisotropic diffusion filtering method is an effective way to denoise original MRI images. The impact of different image filtering approaches on the classification performance of FCN and MLP models was also investigated, which can effectively improve model performance in a small range. In order to compare the effectiveness and differences of different attention mechanisms, we add the SE mechanism and the Pyramid Squeeze Attention mechanism to Fully Convolutional Network model respectively. It can be seen from the results in Tables 9–11 that, the Pyramid Squeeze Attention mechanism produces better multi-scale feature information than the SE mechanism and can more effectively extract multi-scale spatial information of images. The Pyramid Squeeze Attention mechanism can provide more abundant feature information of disease probability map for MLP models, to improve the classification

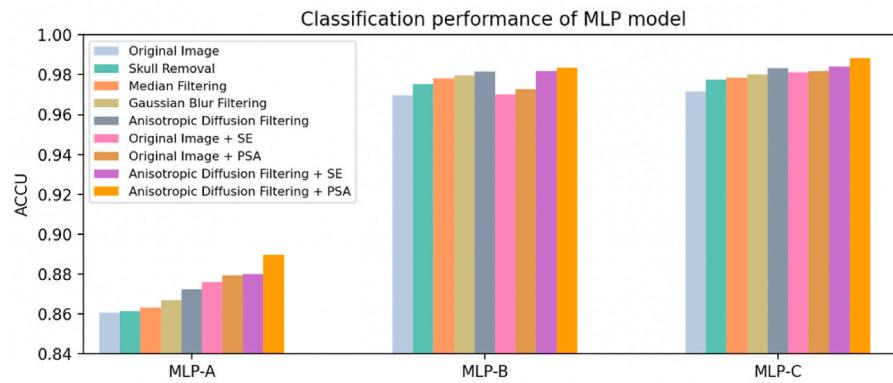


Fig. 16. The changes of MLP models' accuracy.

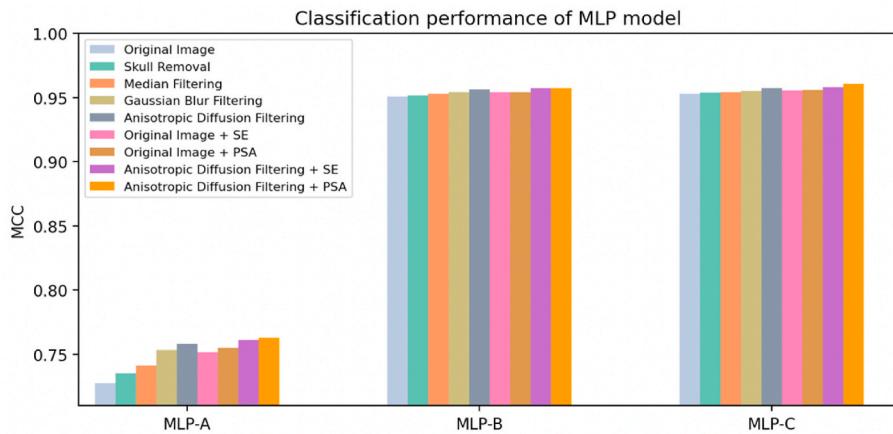


Fig. 17. The changes of MLP models' MCC.

Table 12
ROC-AUC values of MLP models.

Method	AUC (mean ± std)		
	MLP-A	MLP-B	MLP-C
Original Image	0.946 ± 0.014	0.994 ± 0.002	0.997 ± 0.001
Original Image + PSA	0.955 ± 0.013	0.995 ± 0.001	0.997 ± 0.001
Anisotropic Diffusion Filtering	0.953 ± 0.008	0.995 ± 0.001	0.998 ± 0.001
Anisotropic Diffusion Filtering + PSA	0.960 ± 0.010	0.997 ± 0.001	0.999 ± 0.001

performance of MLP models. Moreover, the MLP-C model combines feature information of disease probability map with age, gender, and MMSE of samples, which is more conducive to accurately judging images of Alzheimer's disease than the other two MLP models.

As we know, with the increase of disease time, AD patients' hippocampus shows obvious atrophy. However, the hippocampus of normal cognitive (NC) persons with getting older does not show obvious atrophy. The atrophy of the hippocampus will cause symptoms such as memory loss, so AD patients often experience dementia, confusion, and memory loss. Therefore, the hippocampus (HIP) and the parahippocampal gyrus (PHG) are one of the important areas to distinguish Alzheimer's disease. After adding the Pyramid Squeeze Attention mechanism, the MCC heatmap of FCN model can provide more favorable brain regions for MLP models to identify Alzheimer's disease. These brain regions for classifying AD and NC are also known as discriminative

brain regions [37]. And the BrainNet viewer [38] toolbox is used to visually display the selected discriminative brain regions, as shown in Fig. 19. Regions of the temporal lobe are closely related to language and memory. It can be seen from Fig. 19, the temporal lobe regions mainly choose the following discriminative brain regions: the left hippocampus (HIP.L), the left parahippocampal gyrus (PHG.L), the right hippocampus (HIP.R), the right parahippocampal gyrus (PHG.R) and the right temporal pole superior temporal gyrus (TPOsup.R).

We analyze the advantages and disadvantages of different classifiers and classification techniques [39,40]. For example, the Long Short-Term Memory (LSTM) network [41] is a typical Recurrent Neural Network (RNN) that can solve some vanishing gradient problems. However, it has disadvantages in parallel processing and a large amount of calculation. And the basic model of Support Vector Machines (SVM) [42] is a linear classifier with the largest interval defined in feature space. For a nonlinear classification problem in the input space, it can be transformed into a linear classification problem in a certain dimensional feature space through nonlinear transformation, and then a linear SVM is learned in a high-dimensional feature space [43]. However, SVM is suitable for small data samples and high-dimensional classification problems. In addition, a Stacked Auto-Encoder (SAE) [44] is a commonly used model in the field of deep learning, which consists of multiple autoencoders stacked in series. It can extract high-order features of input data layer by layer, convert complex input data into a simple high-order feature, and then apply these high-order features to a classifier [45]. As shown in Table 13, our proposed methods are compared with classification methods of Alzheimer's disease in relevant references, which include related AD studies using fMRI and sMRI data.

However, there are many unknown factors in the deep learning

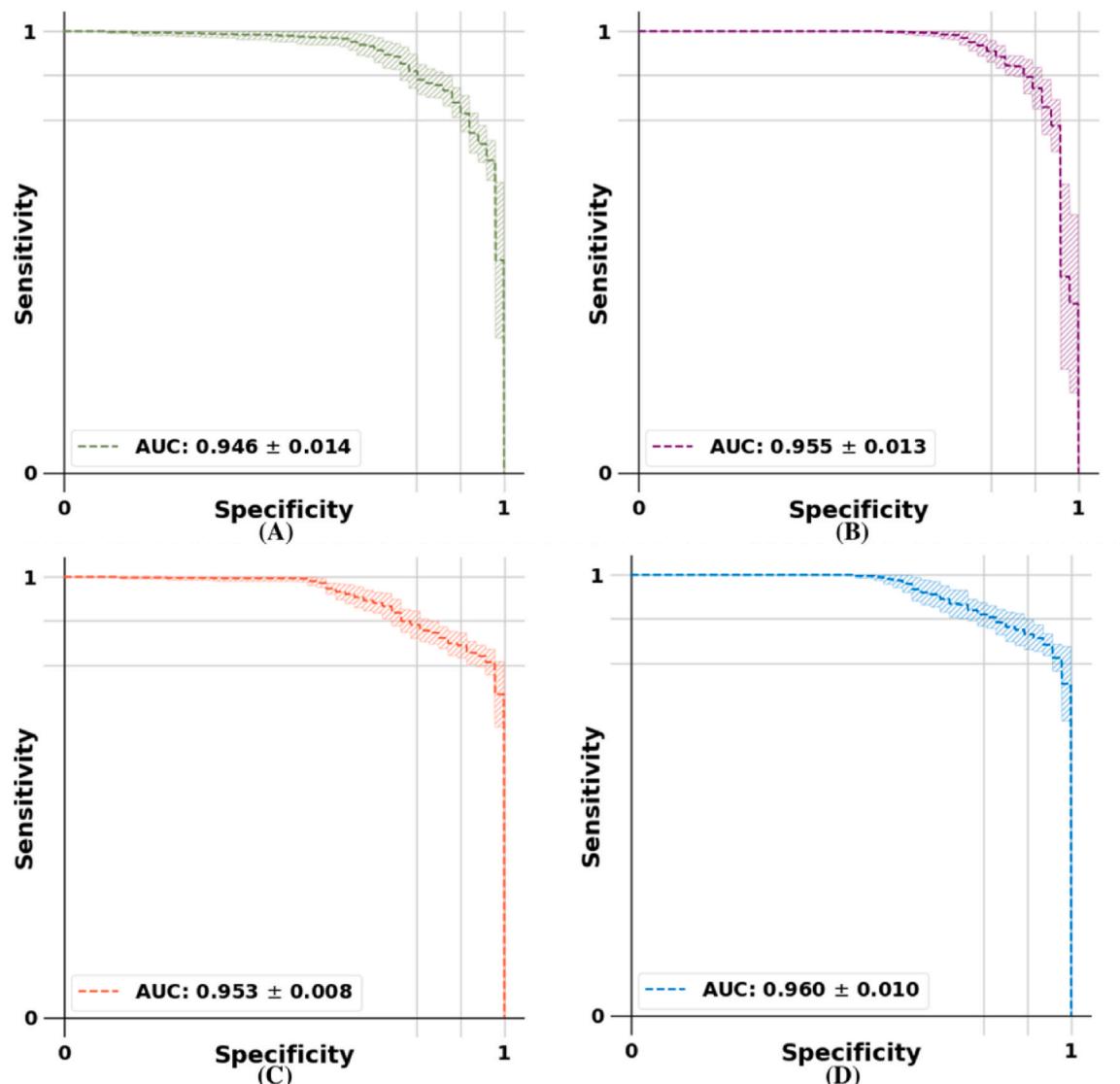


Fig. 18. (A) the ROC curve of MLP-A model under the original image. (B) the ROC curve of MLP-A model under the original image and PSA mechanism. (C) the ROC curve of MLP-A model under anisotropic diffusion filtering. (D) the ROC curve of MLP-A model under anisotropic diffusion filtering and PSA mechanism.

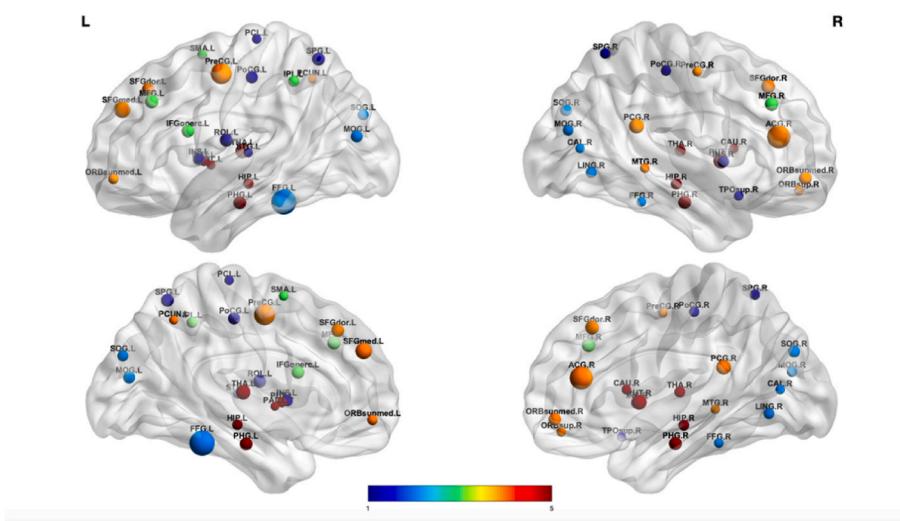


Fig. 19. Visual display model-selected discriminative brain regions.

Table 13

Classification methods of Alzheimer's disease in relevant references.

Author	Type of Dataset	Methods	Number of Samples	Accu	Sens	Spec
Oh K et al. [46]	MRI	Inception autoencoder based CNN architecture	198AD/230NC	0.8660	0.8855	0.8454
Prajapati R et al. [47]	MRI	Deep neural network with FC layers	58AD/73NC	0.8519	/	/
Korolev S et al. [48]	MRI	3D-CNN and Residual	50AD/61NC	0.79	/	/
Xia Z et al. [49]	MRI	3D-CNN + ConvLSTM	198AD/229NC	0.9419	0.9375	0.9457
Gupta A et al. [50]	MRI	Sparse autoencoder + Cross-domain features	200AD/232NC	0.9474	0.9524	0.9426
Hinrichs C et al. [51]	MRI	SVM + Linear Program boost (LP) + Voxel-level engineered features	89AD/94NC	0.82	0.85	0.80
Ieracitano C et al. [52]	MRI	Electroencephalographic (EEG) Signals + CNN	63AD/63NC	0.8578	0.6967	/
Zhao J et al. [53]	fMRI	Functional connectivity of static (sWGFC) and dynamic (dWGFC) between WM and GM + SVM	45AD/45NC	0.8111	0.8444	0.7778
Bi X et al. [54]	fMRI	A random support vector machine cluster	25AD/35NC	0.9444	/	/
Zhu W et al. [55]	sMRI	A dual attention multi-instance deep learning network (DA-MIDL)	389AD/400NC	0.9240	0.9100	0.9380
Feng J et al. [56]	sMRI	Nonsubsampled contourlet subband-based individual networks	200AD/200NC	0.9421	0.9658	0.9244
Gao X et al. [57]	MRI + PET	Integrate TPA-GAN with PT-DCN	196AD/227NC	0.927	0.917	0.935
Odusami M et al. [58]	fMRI	ResNet18 with Dropout	25AD/25NC	0.8080	0.9183	0.8387
Zhang J et al. [59]	MRI	CNN with connection-wise attention mechanism	280AD/275NC	0.9735	0.9710	0.9795
Our proposed methods	MRI	Anisotropic diffusion filtering Original image + PSA	309AD/241NC	0.9832 0.9817	0.9751 0.9712	0.9613 0.9608
		Anisotropic diffusion filtering + PSA		0.9885	0.9765	0.9644

model, which make it difficult to intuitively analyze the corresponding relationship between the classification model and results, so it is difficult to apply in clinical practice. Therefore, we expect to further study the interpretability of the deep learning model to improve the credibility of classification results.

5. Conclusions and future directions

In this paper, we proposed a new method for image classification of Alzheimer's disease based on the Pyramid Squeeze Attention mechanism and Fully Convolutional Network. During image preprocessing, we explore different image filtering approaches and quantify their impact on model performance. The experimental results show that anisotropic diffusion filtering can better retain detailed features of images, and the Pyramid Squeeze Attention mechanism can more fully extract spatial information from a multi-scale feature map. Compared with classification methods of relevant references, our proposed method can efficiently and stably distinguish images of Alzheimer's disease patients, which has important practical significance for diagnosis and early intervention of Alzheimer's disease.

We only classify AD and NC without considering mild cognitive impairment (MCI). In the future, we will consider fusing spatial information of multi-modal MRI and PET images to obtain more detailed features [60]. Moreover, multi-modal feature selection using feature correlation and feature structure fusion can make better use of strongly correlated image features, which provides a new idea for the diagnosis and classification of MCI and NC.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgement

This work is supported in part by the National Natural Science Foundation of China (62011530130, 61672466 and 62101497), Joint Fund of Zhejiang Provincial Natural Science Foundation

(LSZ19F010001), the Key Research and Development Program of Zhejiang Province (2020C03060), the 521 Talents project of Zhejiang Sci-Tech University, the ERC IMI (101005122), the H2020 (952172), the MRC (MC/PC/21013), the Royal Society (IEC\ NSFC\211235), and the UKRI Future Leaders Fellowship (MR/V023799/1).

References

- [1] T. Tong, Q. Gao, R. Guerrero, A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 64 (2017) 155–165.
- [2] H.P. Liew, Healthcare satisfaction among the older adults with Alzheimer's disease or dementia, *SN Social Sciences* 1 (2021) 1–13.
- [3] H. Wu, J. Luo, X. Lu, 3D transfer learning network for classification of Alzheimer's disease with MRI, *Int. J. Mach. Learning Cybernetics* 13 (7) (2022) 1997–2011.
- [4] S. Alinsaif, J. Lang, 3D shearlet-based descriptors combined with deep features for the classification of Alzheimer's disease based on MRI data, *Comput. Biol. Med.* 138 (2021), 104879.
- [5] V.N. Varghees, M.S. Manikandan, R. Gini, Adaptive MRI image denoising using total-variation and local noise estimation, in: *IEEE-international Conference on Advances in Engineering, Science and Management (ICAESM)*, Tamil Nadu, Nagapattinam, 2012, pp. 506–511. March 30-31.
- [6] A. Roy, J. Singha, L. Manam, Combination of adaptive vector median filter and weighted mean filter for removal of high-density impulse noise from colour images, *IET Image Process.* 11 (6) (2017) 352–361.
- [7] P. Singh, A. Verma, A. Garg, A study in finding effectiveness of Gaussian blur filter over bilateral filter in natural scenes for graph based image segmentation, in: *4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2017, pp. 1–6. January 6-7.
- [8] X. Ma, H. Shen, L. Zhang, Adaptive anisotropic diffusion method for polarimetric SAR speckle filtering, *IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens.* 8 (3) (2017) 1041–1050.
- [9] L.E. Rossovskii, Image filtering with the use of anisotropic diffusion, *Comput. Math. Math. Phys.* 57 (3) (2017) 401–408.
- [10] E. Kaplan, S. Dogan, T. Tuncer, Feed-forward LPQNet based automatic Alzheimer's disease detection model, *Comput. Biol. Med.* 137 (2021), 104828.
- [11] M. Odusami, R. Maskeliunas, R. Damaševičius, An intelligent system for early recognition of alzheimer's disease using neuroimaging, *Sensors* 22 (3) (2022) 740.
- [12] A. Loddo, S. Butta, Deep learning based pipelines for Alzheimer's disease diagnosis: a comparative study and a novel deep-ensemble method, *Comput. Biol. Med.* 141 (2022), 105032.
- [13] N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, Y. Li, A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease, *Neurocomputing* 320 (2018) 195–202.
- [14] I. Beheshti, H. Demirel, Feature-ranking-based Alzheimer's disease classification from structural MRI, *Magn. Reson. Imaging* 34 (3) (2016) 252–263.

- [15] W. Zhao, S. Du, Spectral-spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach, *IEEE Trans. Geosci. Rem. Sens.* 54 (8) (2016) 4544–4554.
- [16] Y. Luo, Y. Wen, D. Tao, Large margin multi-modal multi-task feature extraction for image classification, *IEEE Trans. Image Process.* 25 (1) (2015) 414–427.
- [17] C.Q. Lai, S.S. Teoh, A review on pedestrian detection techniques based on Histogram of Oriented gradient feature, in: 2014 IEEE Student Conference on Research and Development, Penang, Malaysia, 2014, pp. 1–6, December 16–17.
- [18] R.M. Mueid, C. Ahmed, M.A.R. Ahad, Pedestrian activity classification using patterns of motion and histogram of oriented gradient, *J. Multimodal User Interfaces* 10 (2016) 299–305.
- [19] F. Ahmad, W.M. Dar, Classification of alzheimer's disease stages: an approach using PCA-based algorithm, *Am. J. Alzheimer's Dis. Other Dementias* 33 (2018) 433–439.
- [20] E.S.A. El-Dahshan, T. Hosny, A.B.M. Salem, Hybrid intelligent techniques for MRI brain images classification, *Digit. Signal Process.* 20 (2010) 433–441.
- [21] S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease, *IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng.* 62 (4) (2015) 1132–1140.
- [22] H.I. Suk, S.W. Lee, D. Shen, Latent feature representation with stacked auto-encoder for AD/MCI diagnosis, *Brain Struct. Funct.* 220 (2) (2015) 841–859.
- [23] M. Yaqoob, B. Wróbel, Very small spiking neural networks evolved to recognize a pattern in a continuous input stream, in: 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 2017, pp. 1–8, November 27–30.
- [24] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, H. Lu, Dual attention network for scene segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019), Long Beach, CA, USA, 2019, pp. 3146–3154, June 16–20.
- [25] R. Jin, X. Han, T. Yu, A real-time image semantic segmentation method based on multilabel classification, *Math. Probl. Eng.* 2021 (1) (2021) 1–13.
- [26] H. Zhang, I. Goodfellow, D. Metaxas, Self-Attention Generative Adversarial Networks, 2018 arXiv:1805.08318vol. 1.
- [27] S. Woo, J. Park, J.Y. Lee, CBAM: Convolutional Block Attention Module, 2018 arXiv:1807.06521vol. 2.
- [28] H. Jie, S. Li, S. Gang, Squeeze-and-Excitation Networks, 2019 arXiv: 1709.01507vol. 4.
- [29] H. Zhang, K. Zu, J. Lu, EPSANet: an Efficient Pyramid Squeeze Attention Block on Convolutional Neural Network, 2021 arXiv:2105.14447vol. 2.
- [30] V. Rajagopalan, E.P. Pioro, Disparate voxel based morphometry (VBM) results between SPM and FSL softwares in ALS patients with frontotemporal dementia: which VBM results to consider, *BMC Neurol.* 15 (1) (2015) 1–7.
- [31] J. Barbosa, T. Nunes, Semi-automatic active contour-based segmentation to remove eyes, meninges, and skull from MRI, *Res. Biomed. Eng.* 36 (2020) 369–377.
- [32] A. Ramadhan, F. Mahmood, A. Elci, Image denoising by median filter in wavelet domain, *Int. J. Multimed. Appl.* 9 (1) (2017) 31–40.
- [33] T. Dillon, A Combinatorial Interpretation of Gaussian Blur, 2018 arXiv: 1812.03569vol. 2.
- [34] A.G. Khoei, K.M. Mohammadi, M. Jani, A Least Squares Support Vector Regression for Anisotropic Diffusion Filtering, 2022 arXiv:2202.00595vol. 1.
- [35] D. Chicco, N. Tsch, G. Jurman, The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation, *BioData Min.* 14 (1) (2021) 1–22.
- [36] D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, *BMC Genom.* 21 (1) (2020) 1–13.
- [37] B. Ng, A.C. Milazzo, A. Altmann, Node-based Gaussian graphical model for identifying discriminative brain regions from connectivity graphs, *Machine Learning in Medical Imaging* 9352 (2015) 44–51.
- [38] M. Xia, J. Wang, Y. He, BrainNet Viewer: a network visualization tool for human brain connectomics, *PLoS One* 8 (7) (2013), e68910.
- [39] A. Ashraf, S. Naz, S.H. Shirazi, Deep transfer learning for alzheimer neurological disorder detection, *Multimed. Tool. Appl.* 80 (20) (2021) 30117–30142.
- [40] M. Odusami, R. Maskeliūnas, R. Damaševičius, An intelligent system for early recognition of alzheimer's disease using neuroimaging, *Sensors* 22 (3) (2022) 740.
- [41] F. Oliver, S. Alex, K. Murtadha, Automated detection of atrial fibrillation using long short-term memory network with RR interval signals, *Comput. Biol. Med.* 102 (2018) 327–335.
- [42] L. Shen, H. Chen, Z. Yu, Evolving support vector machines using fruit fly optimization for medical data classification, *Knowl. Base Syst.* 96 (2016) 61–75.
- [43] E.H. Houssein, D.S. Abdelminaam, H.N. Hassan, A hybrid barnacles mating optimizer algorithm with support vector machines for gene selection of microarray cancer classification, *IEEE Access* 9 (2021) 64895–64905.
- [44] X. Zhang, H. Dou, T. Ju, Fusing heterogeneous features from stacked sparse autoencoder for histopathological image analysis, *IEEE J. Biomed. Health Inform.* 20 (5) (2016) 1377–1383.
- [45] L.L. Cao, W.B. Huang, F.C. Sun, Building feature space of extreme learning machine with sparse denoising stacked-autoencoder, *Neurocomputing* 174 (2016) 60–71.
- [46] K. Oh, Y.C. Chung, K.W. Kim, Classification and visualization of alzheimer's disease using volumetric convolutional neural network and transfer learning, *Sci. Rep.* 9 (1) (2019) 1–16.
- [47] R. Prajapati, U. Khatri, G.R. Kwon, An efficient deep neural network binary classifier for alzheimer's disease classification, in: 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Jeju Island, Korea, 2021, pp. 231–234, April 13–16.
- [48] S. Korolev, A. Safiullin, M. Belyaev, Y. Dodonova, Residual and Plain Convolutional Neural Networks for 3D Brain MRI Classification, 2017 arXiv: 1701.06643vol. 1.
- [49] Z. Xia, G. Yue, Y. Xu, C. Feng, M. Yang, T. Wang, B. Lei, A novel end-to-end hybrid network for Alzheimer's disease detection using 3D CNN and 3D CLSTM, in: 2020 IEEE 17th International Symposium on Biomedical Imaging, Iowa City, IA, USA, 2020, pp. 1–4, April 3–7.
- [50] A. Gupta, M. Ayhan, A. Maida, Natural image bases to represent neuroimaging data, in: 30th International Conference on Machine Learning (ICML-13), Atlanta, Georgia, USA, 2013, pp. 987–994, June 16–21.
- [51] C. Hinrichs, V. Singh, L. Mukherjee, Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset, *Neuroimage* 48 (1) (2009) 138–149.
- [52] C. Ieracitano, N. Mammone, A. Hussain, A Convolutional Neural Network based self-learning approach for classifying neurodegenerative states from EEG signals in dementia, in: 2020 International Joint Conference on Neural Networks, IJCNN, Glasgow, UK, 2020, pp. 1–8, July 19–24.
- [53] J. Zhao, X. Ding, Y. Du, Functional connectivity between white matter and gray matter based on fMRI for Alzheimer's disease classification, *Brain behav.* 9 (10) (2019), e01407.
- [54] X. Bi, Q. Shu, Q. Sun, Random support vector machine cluster analysis of resting-state fMRI in Alzheimer's disease, *PLoS One* 13 (3) (2018), e0194479.
- [55] W. Zhu, L. Sun, J. Huang, Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI, *IEEE Trans. Med. Imag.* 40 (9) (2021) 2354–2366.
- [56] J. Feng, S.W. Zhang, L. Chen, Alzheimer's disease classification using features extracted from nonsubsampled contourlet subband-based individual networks, *Neurocomputing* 421 (2021) 260–272.
- [57] X. Gao, F. Shi, D. Shen, Task-induced pyramid and attention GAN for multimodal brain image imputation and classification in Alzheimers disease, *IEEE J. Biomed. Health Inform.* 26 (1) (2021) 36–43.
- [58] M. Odusami, R. Maskeliūnas, R. Damaševičius, Analysis of features of alzheimer's disease: detection of early stage from functional brain changes in magnetic resonance images using a finetuned ResNet18 network, *Diagnostics* 11 (6) (2021) 1071.
- [59] J. Zhang, B. Zheng, A. Gao, A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification, *Magn. Reson. Imag.* 78 (2021) 119–126.
- [60] J. Zhang, X.H. He, L.B. Qing, BPGAN: brain PET synthesis from MRI using generative adversarial network for multi-modal Alzheimer's disease diagnosis, *Comput. Methods Progr. Biomed.* 217 (2022), 106676.