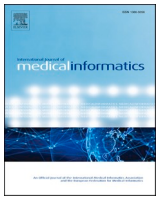
International Journal of Medical Informatics 191 (2024) 105551

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



A multi-scale feature selection module based architecture for the diagnosis of Alzheimer's disease on [^{18}F]FDG PET

Yuling Wang [a], Shijie Chen [a], Xin Tian [a], Yuan Lin [a], Dongqi Han [a], Ping Yao [b], Hang Xu [a], Yuanyuan Wang [c], Jie Zhao [a,*]

[a] School of Medical Imaging, Xuzhou Medical University, Xuzhou, China
[b] Xuzhou First People's Hosipital, Xuzhou, China
[c] Yancheng Institute of Technology, Yancheng, China

ARTICLE INFO

Keywords:
Alzheimer's disease
Diagnosis
Deep learning
Positron Emission Tomography

ABSTRACT

Objective: Alzheimer's disease (AD) is a prevalent form of dementia worldwide as a cryptic neurodegenerative disease. The symptoms of AD will last for several years, which brings great mental and economic burden to patients and their families. Unfortunately, the complete cure of AD still faces great challenges. Therefore, it is crucial to diagnose the disease in the early stage.

Materials and Methods: The Visual Geometry Group (VGG) network serves as the backbone for feature extraction, which could reduce the time cost of network training to a certain extent. In order to better extract image information and pay attention to the association information in the images, the group convolutional module and the multi-scale RNN-based feature selection module are proposed. The dataset employed in the study are drawn from [^{18}F]FDG-PET images within the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

Results: Comprehensive experimental results show that the proposed model outperforms several competing approaches in AD-related diagnostic tasks. In addition, the model reduces the number of parameters of the model compared to the backbone model, from 134.27 M to 17.36 M. Furthermore, the ablation reaserch is conducted to confirm the effectiveness of the proposed module.

Conclusions: The paper introduces a lightweight network architecture for the early diagnosis of AD. In contrast to analogous methodologies, the proposed method yields acceptable results.

1. Introduction

Alzheimer's Disease (AD) is a prevalent form of dementia characterized by prominent symptoms such as memory loss and cognitive decline, significantly disrupting the daily lives of those afflicted [1]. It is widely recognized that the pathological progression of AD commences many years prior to the manifestation of overt dementia symptoms, ushering in a pre-clinical stage characterized by the absence of discernible cognitive impairments [2–4]. While a definitive cure for this condition remains elusive, patients can receive medications aimed at enhancing cognitive function. To ensure more effective treatment, precise and early diagnosis of the disease is imperative [5].

Historically, AD diagnosis has relied on a triad of methods including brain histopathological analysis, neuropsychological testing, and neuroimaging. Among these, brain histopathological analysis has held the position of the gold standard for AD diagnosis [6]. However, the invasiveness of brain histopathological analysis procedure inflicts trauma upon patients, rendering it unsuitable for routine clinical application. In parallel, while neuropsychological testing provides valuable insights into cognitive function, it exhibits certain limitations in terms of diagnostic rigor and accuracy, thereby constraining its utility in the context of AD diagnosis. In light of these limitations, neuroimaging techniques have emerged as a promising avenue for enhancing the precision of AD diagnosis.

Mild cognitive impairment (MCI) is a prodromal form of dementia, defined as a cognitive impairment that does not interfere with activities of daily living. Although patients with MCI do not show obvious clinical features, it can lead to AD or other degenerative dementias in the future [7,8]. Diagnosis of this disease is challenging, even for experienced neurologists, and in some cases it can be difficult to decide on appropriate treatments. Therefore, physicians use diagnostic tests, such as neurofunctional imaging, to provide a more accurate clinical assessment

* Corresponding author.
E-mail address: zhaojie@xzhmu.edu.cn (J. Zhao).

https://doi.org/10.1016/j.ijmedinf.2024.105551
Received 5 February 2024; Received in revised form 6 June 2024; Accepted 14 July 2024
Available online 18 July 2024

[9]. Compared with Magnetic Resonance Imaging (MRI) imaging, Positron Emission Tomography (PET) imaging can observe abnormalities in relevant brain areas earlier [10], which is beneficial for early diagnosis of AD. [^{18}F]FDG PET scan measures brain glucose metabolism and has been reported as a useful biomarker for identifying the above-mentioned neurodegenerative diseases [11].

Machine learning, particularly deep learning techniques, has found widespread application across various domains within the field of medical image analysis, including but not limited to Computed Tomography (CT), Ultrasound (US), MRI, and PET imaging [12]. The deployment of deep learning in medical image analysis tasks is of immense importance, especially in the context of medical auxiliary diagnosis [13]. Notably, numerous deep learning methods have been introduced for AD diagnosis. However, these methods predominantly either rely on pre-existing domain knowledge or extract image feature information from individual image slices. For instance, Shi, et al. [14] proposed a network approach based on the division of Regions Of Interest (ROI) for AD diagnosis, whereas Rashid et al. [15] developed a technique to predict AD based on 2D image slices by extracting key features from anatomical images.

Despite the promising contributions of these methodologies, they exhibit certain limitations. Firstly, approaches that hinge on prior knowledge for feature extraction [14,16] necessitate a template of the anatomical region within the image, potentially introducing bias in the regional information when constructing ROI. Second, when employing 2D networks for image processing, most of the existing networks [15,17,18] mainly focus on the information of individual regions of the image while ignoring the correlations in the image. In addition, some researchers have opted for the use of 3D networks [19,20] to emphasize inter-image correlation information. However, it is essential to acknowledge that training 3D networks is associated with substantial time and cost investments.

In order to address the limitation of 2D networks to focus only on a single image region, this study proposes an automatic diagnostic architecture for diagnosing attention deficit disorder. Different from the existing approaches, the contributions are distinct in the following aspects:

1. Diagnosis for multiple stages of AD, the discovery of the MCI stage is more valuable and challenging.
2. Incorporation of a Group Convolution module (GConv) meticulously tailored to the network architecture, facilitating the extraction of deeper image features. This innovation effectively broaden the network, enhancing its overall expressive capability.
3. Introduction of a multi-scale RNN-based Feature Selection Module (RFSM), which comprehensively assesses both the global and regional information within individual images.

The structure of this article unfolds as follows: Section 2 elucidates pertinent prior research; Section 3 expounds upon the novel network model and its specific training strategies; Section 4 describes the diagnostic performance of the proposed model through experimental results; Section 5 discusses the performance and advantages and disadvantages of the model; finally, Section 6 summarizes the methodology proposed in the article and describes future research directions.

2. Related works

The field of AD diagnosis has garnered considerable global research interest in recent years. Notably, the fusion of AD diagnosis with artificial intelligence (AI) techniques has yielded promising outcomes. These AI methodologies encompass predictive models based on pre-trained networks, the development of 2D and 3D network architectures, and the utilization of machine learning algorithms such as support vector machines and random forest techniques.

Presently, the prevailing research approach predominantly revolves around the prediction and diagnosis of AD using 2D network architectures. This approach primarily relies on network models pre-trained on ImageNet and newly proposed network structures. The research endeavors by Janghel, Ding [18], and M. Ghazal [21] have demonstrated the effectiveness of fine-tuning pre-trained models, highlighting the potential of transfer learning in AD diagnostic tasks. Novel models have also exhibited commendable performance, as evidenced by Tuan [22], who introduced an Auto Encoder network designed to diagnose AD by effectively discerning regions of interest within the brain. Cui [16] introduced the BMnet network, focused on extracting interregional representation features and identifying challenging samples through the construction of embedding spaces. Chang [23] developed a CNN network model architecture for the diagnosis of AD and its distinction from other medical conditions. Collectively, these studies underscore the robust performance of 2D models in AD diagnosis.

Furthermore, to enhance the consideration of inter-image correlation and improve task accuracy, some researchers have proposed the use of 3D networks for AD diagnosis. For example, De Santi [19] proposed a 3D convolutional neural network approach for multi-classification AD research, while Etminani [20] developed and validated a 3D CNN model capable of predicting diseases such as AD and MCI. Although 3D networks can capture greater image-to-image correlation information, they do not consistently outperform their 2D counterparts and often entail higher time costs.

In conclusion, AI-based methods have yielded satisfactory results in the realm of AD diagnostics. The integration of image-based AD diagnosis into clinical practice has gradually evolved into a pivotal aid in enhancing clinical diagnoses.

3. Material and methods

3.1. Data collection

The data utilized in this study was sourced from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [24], accessible at 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, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of AD.

For model training and testing, we employed [^{18}F]FDG PET images that had undergone preprocessing by the ADNI team. The images subjected to the highest level of preprocessing were selected. The preprocessing steps encompassed several critical stages, including frame co-registration for dynamic acquisitions, frame averaging to generate a single PET image, reorientation into a standard 160 × 160 × 96 voxel image grid with isotropic voxels measuring 1.5 mm, intensity normalization utilizing subject-specific masks to achieve an average signal value of one within the mask, and lastly, the application of a smoothing filter to approximate the lowest scanner resolution used in ADNI. For comprehensive insights into acquisition protocols and preprocessing procedures, readers are encouraged to refer to the ADNI website (https://www.adni-info.org/).

Images in DICOM format of 304 different subjects obtained using different scanners were downloaded from the ADNI website. Acquisitions belong to three different classes and are labeled according to the clinical evaluation performed by ADNI centers: Cognitively Normal (CN) (106 scans), MCI (105 scans), and AD (93 scans). Each case consisted of 96 slices with dimensions of 160 × 160 and voxel sizes of 1.5 × 1.5 × 1.5 mm, ensuring consistency across the dataset. To provide visual context, a selection of representative samples utilized in our experiments is presented in Fig. 1.

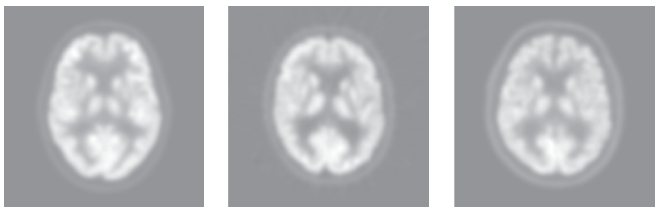


Fig. 1. (a) is the image of CN, (b) is the image of MCI, (c) is the image of AD.

3.2. Proposed methods

3.2.1. Convolutional neural network architecture

In diagnostic tasks, the VGG16 [25] network pre-trained on ImageNet is used as the backbone network for image feature extraction. The backbone network can be formulated as follows:

$$y = r(x) \tag{1}$$

where y is the output, $x \in R^{C\times N\times N}$ is the input of the model, where C is the number of channels, N is the size of images, and the $r(x)$ is the feature extraction part of VGG16 network. The backbone network consists of 13 convolutional layers with trainable weight parameters and pooling layers for data dimensionality reduction. The convolution kernels of the convolutional layers are all 3 × 3. In the features extracte part, only the features part of the backbone network are used.

To further capture deep image features, a new block based on group convolution is added after the backbone network. Drawing inspiration from the work of Liu et al. [26] on ConvNet, the GConv block is proposed, which is characterised by grouping the channels of the input image. GConv not only accomplishes what normal convolution does, but also reduces the likelihood of overfitting. Building upon findings from ConvNet, which demonstrated that reducing the use of activation layers can enhance accuracy. Therefore, only a ReLU activation function is set after the last convolutional layer of the GConv block. The GConv module includes 3 group convolutional layers and a ReLU activation layer. GConv module can be formulated as follows:

$$\mu = R[g(g(g(y^{\frac{C}{n}} \times N \times N)))] \tag{2}$$

where μ is the output, R is the activation layer, g is a standard group convolution layer, y is the input of GConv module, where n is the number of groups in the group convolution. This strategic reduction in activation layers streamlined the model while yielding improvements in experimental accuracy.

At the same time, in order to focus on the information of different areas in the image, a multi-scale RFSM is proposed. Complete feature information of the image is obtained by putting the feature map into the multi-scale RFSM. Finally, the complete image feature information is input to the classifier layer for classification. The comprehensive architecture of the network model is visually depicted in Fig. 2.

3.2.2. A multi-scale RNN-based feature selection Module(RFSM)

In this task, the complexity of the images is high due to the diagnosis of the human brain. It is important to focus not only on the overall information of the image, but also on the details of the image for accurate diagnosis. So the multi-scale RFSM is proposed to meet the requirement. Traditionally employed for processing sequential data, Recurrent Neural Networks (RNNs) have proven to be remarkably effective in tasks characterized by sequential data patterns. Therefore, the idea of combining an RNN with a classification task is proposed, and its specific structure is shown in Fig. 2.

The input to the RFSM consists of the pooled feature map and the

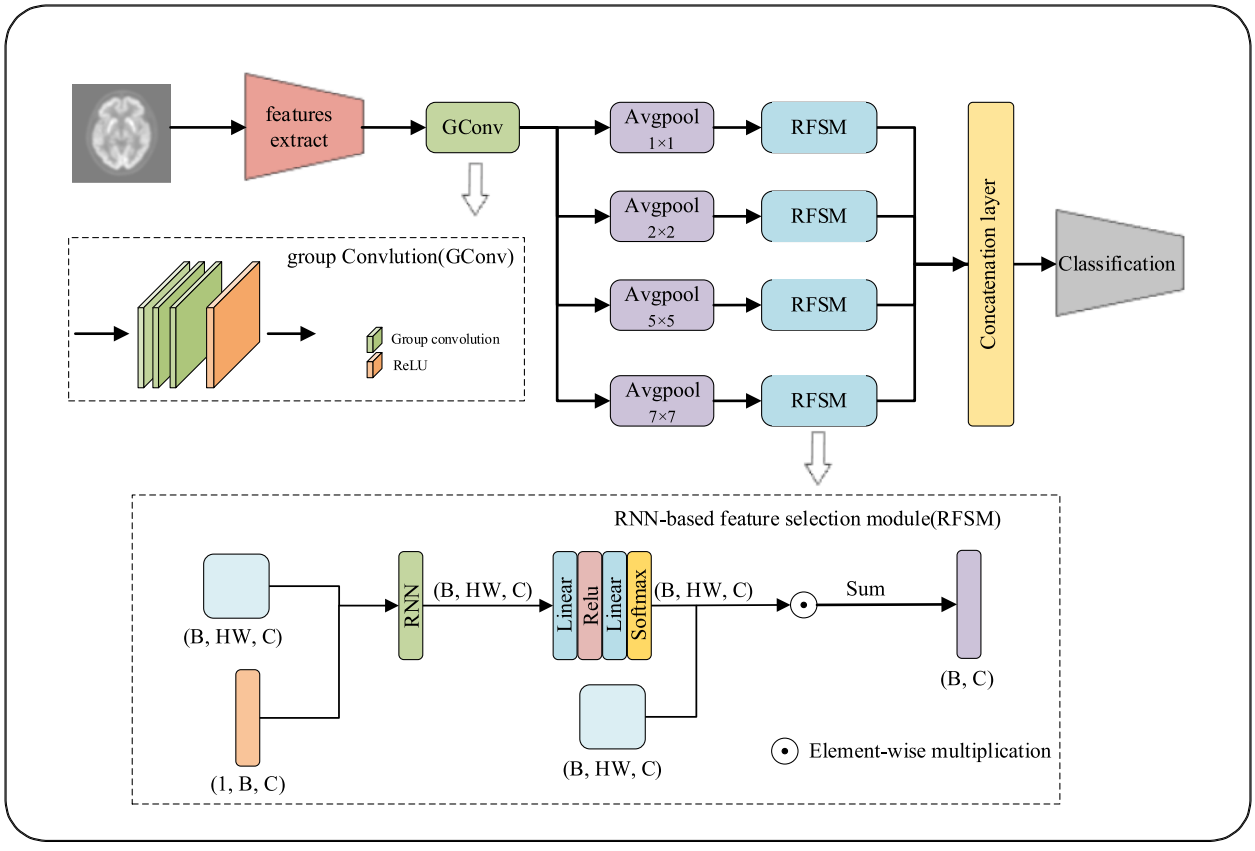


Fig. 2. Architecture of the proposed convolution neural network.

initial feature vector of the hidden layer. The feature map possesses a shape size of (B, H × W, C), while the initial feature vector is characterized by dimensions of (1, B, C). Through the application of RFSM, two additional layers are incorporated, namely a ReLU activation layer and a Softmax classifier. A pivotal step in this process involves the summation of the output vector derived from the RNN and the original input feature map along the first dimension. Consequently, this yields an output vector with dimensions of (B, C).

In addition, in order to concentrate the information from different regions of the image, the feature map of the image is converted into four sizes using the average pooling method: 1×1, 2×2, 5×5, 7×7. Then finally merge all the information in the first dimension. The fusion of overall information and regional information is achieved by the way. Finally the integrated image information is put into the classification layer for diagnosis. The multi-scale RFSM thus plays a vital role in capturing essential image features and contributes significantly to the classification model's efficacy. The calculation process of multi-scale RFSM is shown as Algorithm 1.

Algorithm 1. Calculations of the multi-scale RFSM

Input Preprogress images $\boldsymbol{x}$, Pre-trained model $\boldsymbol{C}$, GConv module $\boldsymbol{G}$, RFSM module $\boldsymbol{R}$
// Step 0: Obtain the feature maps $\boldsymbol{\mu}$
Use $\boldsymbol{C}$ and $\boldsymbol{G}$ to **progress** image $\boldsymbol{x}$cto obtain $\boldsymbol{\mu}$.
// Step 1: Obtain the feature information from multi-scale $\boldsymbol{R}$
for k in *avgpool kernels* $\boldsymbol{K}_{(1,\ 2,\ 5,\ 7)}$:
perform k on $\boldsymbol{\mu}$ to obtain $\boldsymbol{f}$
enter f into $\boldsymbol{R}$ to get $\boldsymbol{I_i}$
end for
feature information = torch.cat($\boldsymbol{I_i}$, dim = 1)
Output feature information

3.3. Experiments details

The utilization of the backbone network necessitates specific image dimensions, compelling us to adjust both the image size and the number of channels. In the study, individual PET image sizes were converted from $160 \times 160 \times 1$ to $224 \times 224 \times 3$ and then normalised.

The dataset comprises a train-set, a validation-set, and a test-set. It is worth noting that all the image information for the entire case was used considering the integrity of the data. Considering that the 96 images in each case were similar, the dataset was not simply divided according to the corresponding proportions in order to avoid data leakage. Instead, place the training set from the test-set separately. Among the training set, 80 % is used as train-set, 20 % is used as validation-set.

The training regimen employs the cross-entropy loss function, the stochastic gradient descent (SGD) optimizer, and a minibatch strategy. Among the model training details, the cross-entropy loss function is the classical loss function for classification, and its specific definition is shown in Equation (3). In the SGD optimizer, Learning Rate (LR) is set to 1e-3, and batch size is set to 32. To avoid overfitting during network training, an early stop mechanism is set, in which the epoch is set to 500 and patience is set to 10.

$$L = -[y\log\hat{y} + (1-y)\log(1-\hat{y})] \tag{3}$$

where y is the real label value and $\hat{y}$ is the predicted label value.

All research in this paper is based on the Python 3.9 programming language, implemented using PyTorch 1.12.0. All experiments were conducted on a computer with 128 GB of memory, running the Windows 11 operating system. The NVIDIA GeForce RTX 3090 graphics card was utilized for the training, with the computer containing a 12th Gen Intel (R) Core(TM) i9-12900 K 3.20 GHz central processing unit.

3.4. Performance of evaluation

Based on the confusion matrix, the following parameters can be calculated to check the performance of the model:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$specificity = \frac{TN}{TN + FP} \tag{6}$$

where given by the classifier, *TP*, *TN*, *FP*, and *FN* are true positive, true negative, false positive, and false negative, respectively.

4. Experiments results

4.1. Performance comparison

To justify the performance of the proposed model for AD diagnosis, the experimental results are compared with several competing methods on the same experimental data in three classification tasks (i.e., CN vs. AD, CN vs. MCI, and AD vs. MCI), which are reported in Table 1. And in Table 1, ACC is accuracy, SEN is sensitivity and SPE is specificity.

It can be seen that the model proposed in the paper can achieve the best performance, i.e., the accuracy can reach 92 %, under the diagnostic task of CN vs. AD in Table 1. Meanwhile, under the diagnostic tasks of CN vs. MCI and AD vs. MCI, the proposed model achieves acceptable results in terms of accuracy, specificity and sensitivity. Since the proposed model is based on VGG16 model as the feature extraction part. Therefore, compared with the backbone model, the accuracy of the proposed model is improved in all three diagnostic tasks. This also shows the effectiveness of the proposed modules, i.e., the GConv module and the RFSM module.

In addition, in order to better validate the robustness of the proposed model, the Receiver Operating Characteristic (ROC) curves of all models were summarized, as shown in Fig. 3. It can be seen that the model has the best performance in the CN vs. AD diagnostic task, and the Area Under Curve (AUC) value can reach 0.96. The value of the AUC in the CN vs. MCI diagnostic task can reach 0.85, which is a slightly lower performance compared to the AD vs. MCI, but it has already reached an acceptable result in the comparison model.

4.2. Ablation experiments

The number of parameters of the backbone model is 134.27 M, and the model accuracy of 76 % in the CN vs. AD diagnostic task and 66 % in the CN vs. MCI diagnostic task and the AD vs. MCI diagnostic task. To improve the accuracy of the AD classification network model, the following strategic adjustments were made to the backbone model:

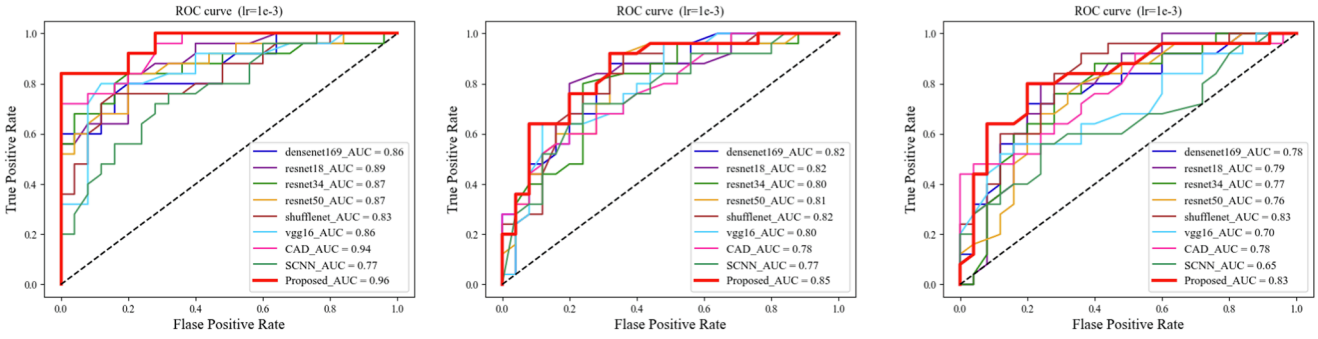
Proposed GConv block: The number of groups of the GConv block is consistent with the input channel, which effectively broadens the network and enhances the expressive ability.
Introduction to multi-scale RFSM: To investigate the impact of information from different regions of the feature map on diagnostic outcomes, a multi-scale concept was introduced. Experimental results show that focusing on information in different areas of the feature map has a great impact on the classification results. RFSM is then introduced to selectively retain salient image features in key blocks and regions that affect the classification results, while paying attention to the correlation information in the image. This optimization improves the efficiency and accuracy of classification models.

In addition to the above modifications to the network model, the hyperparameters in the model training process are also slightly adjusted, mainly to study the impact of LR on the model results, and the specific results are shown in Table 2.

Judging from the experimental data, the addition of the GConv block

Table 1
Comparison of classification results of different methods.

Classification	Metrics(%)	Densnet 169	Resnet 18	Resnet 34	Resnet 50	ShufflenetV2	VGG 16	CAD [27]	SCNN [28]	Proposed
CN vs. AD	ACC	78 %	78 %	78 %	80 %	78 %	76 %	82 %	68 %	**92 %**
	SEN	84 %	80 %	84 %	80 %	84 %	68 %	84 %	64 %	**100 %**
	SPE	72 %	76 %	72 %	80 %	68 %	84 %	80 %	72 %	**84 %**
CN vs. MCI	ACC	74 %	70 %	72 %	70 %	70 %	62 %	70 %	62 %	**78 %**
	SEN	72 %	56 %	74 %	60 %	68 %	68 %	60 %	60 %	**76 %**
	SPE	76 %	**84 %**	82 %	80 %	71 %	64 %	80 %	64 %	80 %
AD vs. MCI	ACC	74 %	70 %	72 %	70 %	74 %	66 %	64 %	62 %	**78 %**
	SEN	72 %	78 %	76 %	56 %	68 %	64 %	68 %	64 %	**80 %**
	SPE	76 %	62 %	68 %	**78 %**	70 %	68 %	60 %	60 %	76 %



(a) CN vs. AD (B) CN vs. MCI (C) AD vs. MCI

Fig. 3. ROC curves for each comparison model under the corresponding diagnostic task.

Table 2
Network model modification process and experimental results.

Classification		Parameters	ACC	SEN	SPE	AUC
CN vs. AD	Base	134.27 M	76 %	68 %	84 %	0.86
	Base + GConv	134.28 M	84 %	88 %	80 %	0.91
	Base + GConv + RFSM	17.36 M	92 %	100 %	84 %	0.96
CN vs. MCI	Base	134.27 M	62 %	60 %	64 %	0.80
	Base + GConv	134.28 M	72 %	68 %	76 %	0.82
	Base + GConv + RFSM	17.36 M	78 %	76 %	80 %	0.85
AD vs. MCI	Base	134.27 M	66 %	64 %	68 %	0.70
	Base + GConv	134.28 M	70 %	72 %	68 %	0.75
	Base + GConv + RFSM	17.36 M	78 %	80 %	76 %	0.83

only increased the number of parameters by 0.01 M, but compared with the backbone network model, the accuracy increased by 8 % in the task of CN vs. AD. As can be seen from Table 2, the improved model not only improves the accuracy of diagnosis, but reduces the number of parameters to a certain extent. The number of model parameters was reduced from 134.27 M to 17.36 M, while the accuracy increased from 76 % to 92 % in the task of CN vs. AD, an increase of about 16 %. Also in the task of CN vs. MCI and AD vs. MCI, the accuracy of the proposed model increased compared to the backbone model.

In order to verify the generality of the proposed modules, the GConv module and RFSM module are added to different backbone models, and the specific results are shown in Fig. 4. Here, "original" refers to the unchanged model, and "modified" indicates the model after integrating proposed modules. From the Fig. 4, it can be observed that for the majority of networks, the performance improves with the addition of GConv module and RFSM. However, there is a declining trend in the performance of some networks. Specifically, in terms of accuracy, the proposed model appears to be more suitable for the VGG and Resnet series models, while exhibiting a declining trend in ShufflenetV2. This indicates that the proposed module has certain limitations and may not be universally applicable to all networks.

5. Discussion

The study introduces a novel AD classification model based on backbone network, achieving a classification accuracy of 92 % in the task of CN vs. AD for PET images.

As can be seen from Table 3, the network model has achieved performance relevant to the relevant literature. The datasets reported in Table 3 are [^{18}F]FDG PET images of patients from ADNI, and the task is to classify AD. Although these experimental results cannot be directly compared with our experimental results, they can reflect the effectiveness of the proposed model to a certain extent.

The feature extraction part of the pre-trained model on ImageNet is used to extract image feature information in the task. In order to better adapt to the task of AD diagnosis, we subsequently fine-tuned the transfer learning model. It can extract useful information from images more accurately, accelerate the convergence speed of the model, and reduce the training time. In terms of model training time, the training time of ordinary models is twice or even more than the training time of transfer learning models. This also reflects an advantage of transfer learning in the task.

The RFSM is set to select image features in the work, which could get relatively good results. RFSM can effectively screen features that are effective for classification results, remove irrelevant information, and more importantly, capture information between images, thereby improving the accuracy of classification results. The application of the RFSM shows that the two major tasks of natural language processing and image processing are related to a certain extent, and they can be

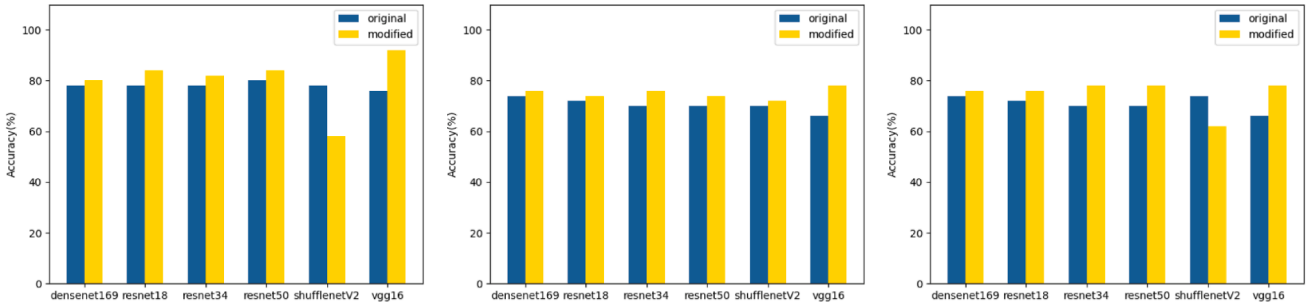


Fig. 4. The accuracy results of ablation experiments: (a) is CN vs. AD, (b) is CN vs. MCI, and (c) is AD vs. MCI.

Table 3
Compare the experimental results in the CN vs. AD diagnostic task.

Method	Modality	Accuracy	Sensitivity	Specificity	AUC
Tufail et al.,2022 [29]	PET	55.4 %	42.9 %	67.8 %	——
Jiao et al.,2023 [30]	PET	81.9 %	83.8 %	78.6 %	——
Hao et al.,2020 [31]	PET	80.1 %	86.0 %	71.9 %	0.85
Janghel et al.,2021 [17]	PET	71.4 %	82.3 %	57.8 %	0.82
Tufail et al.,2020 [32]	PET	80 %	71 %	84 %	0.78
Pan et al.,2020 [33]	PET	88.9 %	83.9 %	93.8 %	0.96
Liu et al.,2018 [34]	PET	84.5 %	82.8 %	86 %	0.92
Ismail et al.,2023 [35]	PET	89.8 %	86.3 %	91.9 %	——
Proposed	PET	**92 %**	**100 %**	**84 %**	**0.96**

interconnected in some processing methods.

There are some limitations in this research. For example, the proposed module is not fully adapted to all backbone models, and the effect is more significant only on VGG and Resnet series models. The relevant module needs to be further modified. Second, the diagnostic effect of the model in the CN vs. MCI and AD vs. MCI diagnostic tasks needs to be further improved. Furthermore, we acknowledge the potential influence of the study sample composition on the model's generalizability, particularly given the significantly higher prevalence of AD (47 %) compared to the general/actual population (~10 %). Such disparity may lead to varied model performance across different populations. Moving forward, we aim to comprehensively address the dataset composition to enhance its representation of the overall population. Concurrently, efforts will be directed towards refining the model to optimize its performance.

6. Conclusion

In this research, a network model is presented that is designed for the diagnosis of [^{18}F]FDG PET images in AD, a crucial step in clinical staging diagnosis. The model not only reduces the number of parameters of the model from 134.27 M to 17.36 M, but achieves an accuracy from 76 % to 92 % in the CN vs. AD diagnostic task. Also achieves an accuracy from 62 % to 78 % in the CN vs. MCI diagnostic task, and the accuracy from 66 % to 78 % in the AD vs. MCI diagnostic task. In this study, GConv module and RFSM module are proposed. And the effectiveness of the GConv module and RFSM module is verified in the ablation experiments.

While previous studies have demonstrated the advantages of PET imaging in early AD diagnosis, it is noteworthy that PET images solely provide functional imaging information and lack anatomical imaging information, which may pose limitations, particularly in early AD diagnosis. Therefore, in subsequent investigations, we will continue to prioritize PET imaging as the primary modality, complemented by MRI imaging as auxiliary anatomical information, to further advance research on early AD diagnosis through multimodal imaging approaches. In conclusion, the article proposes a lightweight and effective network model for AD diagnosis.

7. Summary table

What was already known on the topic?

- Increasing numbers of individuals worldwide are afflicted by Alzheimer's disease (AD); however, there persist challenges in the manual diagnosis of Alzheimer's disease among physicians, particularly concerning early-stage AD diagnosis.
- The application of deep learning techniques enables physicians to identify Alzheimer's disease (AD) patients at an earlier stage utilizing positron emission tomography (PET) scans, thereby facilitating early intervention and mitigating the progression of AD pathology.
- Most of the previous studies focus on the information of a single region of the image, although the 3D network model can focus on the information between images, but its parameters are large, the training time is long, and there are no scientific studies to show that the 3D model has superior performance over the 2D model.

What this study added to our knowledge?

- This study proposes a novel method that combines RNN and CNN models to diagnose AD using a multi-scale RNN-based feature information extraction module.
- This methodology has demonstrated favorable outcomes in AD diagnosis and has been validated for its efficacy across various baseline models.
- This study presents a novel approach to advancing computer-aided diagnosis of AD.
- Future research endeavors should include reporting on the sensitivity, specificity, positive and negative predictive values, as well as overall accuracy of the models, thereby enhancing experimental rigor in validating and analyzing model performance.

8. Statements & declarations

Ethics approval: Since a public dataset was used, there is no need of Ethical Approval.

Author contributions

All authors contributed to the study conception and design. All authors read and approved the final manuscript.

CRediT authorship contribution statement

Yuling Wang: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shijie Chen:** Writing – review & editing, Software, Methodology, Conceptualization. **Xin Tian:** Investigation, Formal analysis, Conceptualization. **Yuan Lin:** Visualization, Formal analysis. **Dongqi Han:** Investigation. **Ping Yao:** Conceptualization. **Hang Xu:** Investigation. **Yuanyuan Wang:** Formal analysis. **Jie Zhao:** Writing – review & editing, Validation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author would like to express special thanks to the Medical Imaging Institute of Xuzhou Medical University and all the collaborators involved in this research for their assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2024.105551.

References

[1] M. Odusami, R. Maskeliūnas, R. Damaševičius, T. Krilavičius, 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 (2021) 1071, https://doi.org/10.3390/diagnostics11061071.

[2] R.K.J. Brown, N.I. Bohnen, K.K. Wong, S. Minoshima, K.A. Frey, Brain PET in suspected dementia: patterns of altered FDG metabolism, RadioGraphics 34 (2014) 684–701, https://doi.org/10.1148/rg.343135065.

[3] P.N.E. Young, M. Estarellas, E. Coomans, M. Srikrishna, H. Beaumont, A. Maass, A. V. Venkataraman, R. Lissaman, D. Jiménez, M.J. Betts, E. McGlinchey, D. Berron, A. O'Connor, N.C. Fox, J.B. Pereira, W. Jagust, S.F. Carter, R.W. Paterson, M. Schöll, Imaging biomarkers in neurodegeneration: current and future practices, Alzheimers Res Ther 12 (2020) 49, https://doi.org/10.1186/s13195-020-00612-7.

[4] M.W. Weiner, P.S. Aisen, C.R. Jack, W.J. Jagust, J.Q. Trojanowski, L. Shaw, A. J. Saykin, J.C. Morris, N. Cairns, L.A. Beckett, A. Toga, R. Green, S. Walter, H. Soares, P. Snyder, E. Siemers, W. Potter, P.E. Cole, M. Schmidt, The Alzheimer's disease neuroimaging initiative: progress report and future plans, Alzheimer's & Dementia 6 (2010) 202, https://doi.org/10.1016/j.jalz.2010.03.007.

[5] J. Weller, A. Budson, Current understanding of Alzheimer's disease diagnosis and treatment, F1000Res 7, 2018, p. 1161.

[6] M.A. DeTure, D.W. Dickson, The neuropathological diagnosis of Alzheimer's disease, Mol Neurodegener 14 (2019) 32, https://doi.org/10.1186/s13024-019-0333-5.

[7] M.S. Albert, S.T. DeKosky, D. Dickson, B. Dubois, H.H. Feldman, N.C. Fox, A. Gamst, D.M. Holtzman, W.J. Jagust, R.C. Petersen, P.J. Snyder, M.C. Carrillo, B. Thies, C.H. Phelps, The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, Alzheimer's and Dementia 7 (2011) 270–279, https://doi.org/10.1016/j.jalz.2011.03.008.

[8] B. Dubois, D. Hospital, B. Dubois, H.H. Feldman, C. Jacova, J.L. Rey Cummings, S. T. DeKosky, P. Barberger-Gateau, A. Delacourte, G. Frisoni, N.C. Fox, D. Galasko, S. Gauthier, H. Hampel, G.A. Jicha, K. Meguro, F. Pasquier, P. Robert, M. Rossor, S. Salloway, M. Sarazin, L.C. de Souza, Y. Stern, P.J. Visser, P. Scheltens, Position Paper Revising the defi nition of Alzheimer's disease: a new lexicon, Www. Thelancet. Com/neurology 9 (2010), https://doi.org/10.1016/S1474.

[9] A.A. Lammertsma, F. Barkhof, P. Scheltens, R. Ossenkoppele, G. Chételat, J. Arbizu, H. Barthel, V. Garibotto, I. Law, S. Morbelli, E. van de Giessen, F. Agosta, F. Barkhof, D.J. Brooks, M.C. Carrillo, B. Dubois, A.M. Fjell, G.B. Frisoni, O. Hansson, K. Herholz, B.F. Hutton, C.R. Jack Jr, A.A. Lammertsma, S.M. Landau, S. Minoshima, F. Nobili, A. Nordberg, R. Ossenkoppele, W.J. G Oyen, D. Perani, G.D. Rabinovici, P. Scheltens, V.L. Villemagne, H. Zetterberg, A. Drzezga, Personal View Amyloid-PET and ^{18}F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias, 2020. www.thelancet.com/neurology.

[10] F.H. Bouwman, S.N.M. Schoonenboom, W.M. van der Flier, E.J. van Elk, A. Kok, F. Barkhof, M.A. Blankenstein, Ph. Scheltens, CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment, Neurobiol Aging 28 (2007) 1070–1074, https://doi.org/10.1016/j.neurobiolaging.2006.05.006.

[11] F. Nobili, J. Arbizu, F. Bouwman, A. Drzezga, F. Agosta, P. Nestor, Z. Walker, M. Boccardi, C. Festari, D. Altomare, F. Gandolfo, S. Orini, European Association of Nuclear Medicine and European Academy of Neurology recommendations for the use of brain 18 F-fluorodeoxyglucose positron emission tomography in neurodegenerative cognitive impairment and dementia: delphi consensus, Eur J Neurol 25 (2018) 1201–1217, https://doi.org/10.1111/ene.13728.

[12] X. Liu, L. Song, S. Liu, Y. Zhang, A review of deep-learning-based medical image segmentation methods, Sustainability 13 (2021) 1224, https://doi.org/10.3390/su13031224.

[13] P. Yifan, D. Bowen, in: An Efficient Deep Learning Model for Predicting Alzheimer's Disease Diagnosis by Using Pet, International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2020, pp. 366–372, https://doi.org/10.1109/ICCWAMTIP51612.2020.9317428.

[14] R. Shi, L. Wang, J. Jiang, in: An Unsupervised Region of Interest Extraction Model for Tau PET Images and Its Application in the Diagnosis of Alzheimer's Disease, 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) IEEE, 2022, pp. 2157–2160, https://doi.org/10.1109/EMBC48229.2022.9871269.

[15] A.H. Rashid, A. Gupta, J. Gupta, M. Tanveer, Biceph-net: A robust and lightweight framework for the diagnosis of Alzheimer's disease using 2D-MRI scans and deep similarity learning, IEEE J Biomed Health Inform 27 (2023) 1205–1213, https://doi.org/10.1109/JBHI.2022.3174033.

[16] W. Cui, C. Yan, Z. Yan, Y. Peng, Y. Leng, C. Liu, S. Chen, X. Jiang, J. Zheng, X. Yang, BMNet: A new region-based metric learning method for early Alzheimer's disease identification with FDG-PET Images, Front Neurosci 16 (2022), https://doi.org/10.3389/fnins.2022.831533.

[17] R.R. Janghel, Y.K. Rathore, Deep convolution neural network based system for early diagnosis of Alzheimer's disease, IRBM 42 (2021) 258–267, https://doi.org/10.1016/j.irbm.2020.06.006.

[18] Y. Ding, J.H. Sohn, M.G. Kawczynski, H. Trivedi, R. Harnish, N.W. Jenkins, D. Lituiev, T.P. Copeland, M.S. Aboian, C. Mari Aparici, S.C. Behr, R.R. Flavell, S.-Y. Huang, K.A. Zalocusky, L. Nardo, Y. Seo, R.A. Hawkins, M. Hernandez Pampaloni, D. Hadley, B.L. Franc, A deep learning model to predict a diagnosis of alzheimer disease by using ^{18} F-FDG PET of the brain, Radiology 290 (2019) 456–464, https://doi.org/10.1148/radiol.2018180958.

[19] L.A. De Santi, E. Pasini, M.F. Santarelli, D. Genovesi, V. Positano, An explainable convolutional neural network for the early diagnosis of alzheimer's disease from 18F-FDG PET, J Digit Imaging 36 (2022) 189–203, https://doi.org/10.1007/s10278-022-00719-3.

[20] K. Etminani, A. Soliman, A. Davidsson, J.R. Chang, B. Martínez-Sanchis, S. Byttner, V. Camacho, M. Bauckneht, R. Stegeran, M. Ressner, M. Agudelo-Cifuentes, A. Chincarini, M. Brendel, A. Rominger, R. Bruffaerts, R. Vandenberghe, M. G. Kramberger, M. Trost, N. Nicastro, G.B. Frisoni, A.W. Lemstra, B.N.M. van Berckel, A. Pilotto, A. Padovani, S. Morbelli, D. Aarsland, F. Nobili, V. Garibotto, M. Ochoa-Figueroa, A 3D deep learning model to predict the diagnosis of dementia with Lewy bodies, Alzheimer's disease, and mild cognitive impairment using brain 18F-FDG PET, Eur J Nucl Med Mol Imaging 49 (2022) 563–584, https://doi.org/10.1007/s00259-021-05483-0.

[21] T.M. Ghazal, S. Abbas, S. Munir, M.A. Khan, M. Ahmad, G.F. Issa, S. Binish Zahra, M. Adnan Khan, M. Kamrul Hasan, Alzheimer disease detection empowered with transfer learning, computers, Materials & Continua 70 (2022) 5005–5019, https://doi.org/10.32604/cmc.2022.020866.

[22] P.M. Tuan, T.-L. Phan, M. Adel, E. Guedj, N.L. Trung, AutoEncoder-based feature ranking for Alzheimer Disease classification using PET image, Machine Learning with Applications 6 (2021) 100184, https://doi.org/10.1016/j.mlwa.2021.100184.

[23] A.J. Chang, R. Roth, E. Bougioukli, T. Ruber, S.S. Keller, D.L. Drane, R.E. Gross, J. Welsh, A. Abrol, V. Calhoun, I. Karakis, E. Kaestner, B. Weber, C. McDonald, E. Gleichgerrcht, L. Bonilha, MRI-based deep learning can discriminate between temporal lobe epilepsy, Alzheimer's Disease, and Healthy Controls, Communications Medicine 3 (2023) 33, https://doi.org/10.1038/s43856-023-00262-4.

[24] R.C. Petersen, P.S. Aisen, L.A. Beckett, M.C. Donohue, A.C. Gamst, D.J. Harvey, C. R. Jack, W.J. Jagust, L.M. Shaw, A.W. Toga, J.Q. Trojanowski, M.W. Weiner, Alzheimer's disease neuroimaging initiative (ADNI), Neurology 74 (2010) 201–209, https://doi.org/10.1212/WNL.0b013e3181cb3e25.

[25] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, (2014). http://arxiv.org/abs/1409.1556.

[26] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, S. Xie, A ConvNet for the 2020s, (2022). http://arxiv.org/abs/2201.03545.

[27] A. Ahila, M. Poongodi, M. Hamdi, S. Bourouis, K. Rastislav, F. Mohmed, Evaluation of neuro images for the diagnosis of Alzheimer's disease using deep learning neural network, Front Public Health 10 (2022) 834032, https://doi.org/10.3389/fpubh.2022.834032.

[28] M. El-Geneedy, H.E.D. Moustafa, F. Khalifa, H. Khater, E. AbdElhalim, An MRI-based deep learning approach for accurate detection of Alzheimer's disease, Alexandria Engineering Journal 63 (2023) 211–221, https://doi.org/10.1016/j.aej.2022.07.062.

[29] A. Bin Tufail, N. Anwar, M.T. Ben Othman, I. Ullah, R.A. Khan, Y.-K. Ma, D. Adhikari, A.U. Rehman, M. Shafiq, H. Hamam, Early-stage Alzheimer's disease categorization using PET neuroimaging modality and convolutional neural networks in the 2D and 3D domains, Sensors 22 (2022) 4609, https://doi.org/10.3390/s22124609.

[30] F. Jiao, M. Wang, X. Sun, Z. Ju, J. Lu, L. Wang, J. Jiang, C. Zuo, Based on Tau PET radiomics analysis for the classification of Alzheimer's disease and mild cognitive impairment, Brain Sci 13 (2023) 367, https://doi.org/10.3390/brainsci13020367.

[31] X. Hao, Y. Bao, Y. Guo, M. Yu, D. Zhang, S.L. Risacher, A.J. Saykin, X. Yao, L. Shen, Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease, Med Image Anal 60 (2020) 101625, https://doi.org/10.1016/j.media.2019.101625.

[32] A. Bin Tufail, Y. Ma, Q.-N. Zhang, in: Multiclass Classification of Initial Stages of Alzheimer's Disease through Neuroimaging Modalities and Convolutional Neural Networks, 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC) IEEE, 2020, pp. 51–56, https://doi.org/10.1109/ITOEC49072.2020.9141553.

[33] X. Pan, T.-L. Phan, M. Adel, C. Fossati, T. Gaidon, J. Wojak, E. Guedj, Multi-view separable pyramid network for AD prediction at MCI stage by ^{18}F-FDG brain PET imaging, IEEE Trans Med Imaging 40 (2021) 81–92, https://doi.org/10.1109/TMI.2020.3022591.

[34] M. Liu, D. Cheng, W. Yan, Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks Using FDG-PET images, Front Neuroinform 12 (2018), https://doi.org/10.3389/fninf.2018.00035.

[35] W.N. Ismail F.R. P. P., M.A.S. Ali A Meta-Heuristic Multi-Objective Optimization Method for Alzheimer's Disease Detection Based on Multi-Modal Data Mathematics 11 2023 957 10.3390/math11040957.