EAI Endorsed Transactions
on e-Learning
Research Article EAI.EU

A Review of Deep Learning Approaches for Early Diagnosis of Alzheimer's Disease

Mengbo Xi[1, *]

[1] School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454000, PR China

Abstract

Alzheimer's disease (AD), one of the major neurodegenerative diseases, has become the most common cause of dementia problems. Up to now, there is a lack of effective targeted therapeutic drugs and effective treatment modalities to stop the progression of the disease. With the continuous development of computer technology, the use of computer-aided diagnostic technology tools for AD early classification studies will provide clinicians with important assistance. Deep learning-based Alzheimer's disease (AD) imaging classification has become a current research hotspot. In this paper, we first describe the commonly used publicly available datasets in the AD imaging classification task; then introduce the commonly used deep learning classification models for AD diagnosis; secondly, we compare the studies that target different biomarkers of the subjects and the use of unimodal or a combination of different modalities for the early classification of AD; and finally, The challenges of AD classification are summarized and future research directions are proposed.

Keywords: Deep learning, Alzheimer's Disease, Multimodal Image, Medical Image

Received on 03 January 2024, accepted on 15 January 2024, published on 16 January 2024

doi: 10.4108/eetel.4790

*Corresponding author. Email: xmb@home.hpu.edu.cn

1. Introduction

Alzheimer's disease (AD) is a neurodegenerative disease characterized by cognitive dysfunction [1]. Currently, the conventional diagnostic method is for doctors to utilize their professional knowledge and clinical experience to interpret brain neuroimages, and the diagnostic efficiency depends on the level of medical resources such as medical personnel and image acquisition equipment, which may lead to missed diagnosis and misdiagnosis due to the lack of timely feedback of patient information. Therefore, many scholars have developed advanced computer-aided diagnosis (CAD) systems [2, 3] to assist clinicians in improving the diagnostic efficiency and early prediction accuracy of AD.

In recent years, CAD systems based on deep learning methods have achieved remarkable results in the diagnosis of sleeve degenerative diseases such as Parkinson's disease, amyotrophic lateral sclerosis, and AD [4-6]. The deep learning method automatically extracts image abstract features by constructing a deep network, and improves the network architecture, global or local contextual information extraction, and multi-scale fusion of features, to realize the control of subjects with Controlled Normal (CN), progressive mild cognitive impairment (pMCI), stable mild cognitive impairment (sMCI), and Alzheimer's disease (AD) [7]. With the development of computer vision technology, deep learning methods are more and more widely used in the field of image processing, and many classical neural networks have emerged, especially convolutional neural network (CNN) [8], which is a class of feed-forward neural networks containing convolutional computation with deep structure, and mining deep features of images through end-to-end learning. CNN is a type of feed-forward neural network with convolutional computation and deep structure, which learns end-to-end to mine the deep features of an image without the need for complex manual feature extraction, CNN and its derivative models have shown great potential in the field of early diagnosis and prediction of AD patients' disease course [9-11].

In this paper, we firstly organize the commonly used datasets for Alzheimer's disease research, and secondly, we introduce the commonly used deep learning models for AD classification based on deep learning; subsequently, we introduce the research progress in AD diagnosis [12, 13] using different modalities; finally, we introduce the challenges in deep learning-assisted AD diagnosis, and finally, we provide an outlook on the future research direction.

2. Datasets

Datasets are one of the important factors for conducting deep learning research. In recent years, with the development of medical information technology, large-scale and standardized neuroimaging datasets provide support for deep learning in the field of AD-assisted diagnosis. Now the data samples of global public datasets are abundant, and researchers can choose single-modality data or joint multi-modality data from the corresponding datasets to construct assisted diagnosis models according to their work requirements. Table 1. summarizes the commonly used datasets in Alzheimer's disease research, which are organized and introduced through the paper reports of AD-related journals and related open-source dataset websites in recent years.

Table 1. Commonly used datasets for Alzheimer's disease

Datasets	Years	Stage	Types and numbers	Imaging type
ADNI	2004	ADNI-1 ADNI-GO/2 ADNI-3	483 CN 551 MCI 437 AD 300 sMCI 150 pMCI	MRI、PET
OASIS	2007	OASIS-1 OASIS-2 OASIS-3	701 CN 503 MCI 164 AD	MRI、PET
AIBL	2006	None	768 CN 133 MCI 211 AD	MRI、PET
MIRIAD	2013	None	23 CN 46 AD	MRI

2.1. ADNI

ADNI was founded in 2003 and is currently the most successful and widely used comprehensive research dataset for Alzheimer's disease. It has the characteristics of multicenter and cross disciplinary research, mainly studying human brain MRI neuroimaging, human brain PET imaging, other human biomarkers such as cerebrospinal fluid (CSF), blood biomarkers, as well as human genetic information, clinical data, neuropsychological evaluation, and other information. 800 adults from 59 countries and regions around the world have been recruited as participants, with an age range of 55 to 90 years old. Providing valuable clinical medical data for global AD research, ADNI has become a core data resource adopted by researchers [14].

2.2. OASIS

The OASIS dataset consists of 2 major categories, the cross-sectional dataset and the longitudinal dataset. The cross-sectional dataset covers MRI data resources of 416 subjects aged 18 to 96 years. The longitudinal dataset covers the MRI data resources of 150 subjects aged 60-96 years, and each subject generally has 2 or more scans with a full-year interval. Currently, OASIS is a core data resource second only to ADNI [15].

2.3. AIBL

The AIBL provides survey statistics on baseline demographics, diagnosis, cognitive function, health, and lifestyle for 1,000 subjects aged 60 years and older. Approximately 25% of the subjects participated in amyloid PET imaging scans with the Pittsburgh Compound (PiBPET) and MRI brain imaging. The AIBL's repeated assessment of subjects over 18-month intervals allows for more adequate identification of different biomarkers and strengthens the predictive criteria for AD in the context of the involvement of cognitive parameters and lifestyle, among other factors. The AIBL plays an important role.

2.4. MIRIAD

The MIRIAD dataset subjects consisted of 46 individuals with mild Alzheimer's disease and 23 normal controls and contained a series of longitudinal volume T1-weighted MRI scan medical imaging images of the above subjects. All of these images above consisted of the same sequence of 708 scans acquired by the same radiologic technologist using the same scanning equipment. Scan intervals ranging from 2 weeks, 6 weeks, 14 weeks, 26 weeks, 38 weeks, 52 weeks, 18 months, and 24 months from baseline were included in these images [16]. The dataset contains records of subjects' status regarding gender, age, and score on the Summary Mental State Examination.

3. Common Classification Models

3.1. Convolutional neural network (CNN)

CNN is the most widely used artificial neural network in deep learning, inspired by the principle of optic nerve in primates. It can preserve the image features while compressing the data volume.CNN is mainly composed of four parts: the convolutional layer, excitation layer, pooling layer, and fully connected layer [17]: firstly, the main function of the convolutional layer is to extract and analyze the image through convolutional kernel computation and unify the output size of the image by padding, and it can also be adjusted to achieve the height and width of the feature map by adjusting the parameter's step size to reduce The excitation layer uses ReLU, Sigmoid, and other excitation functions to introduce nonlinear factors, which makes the final expression of the convolutional neural network more generalized; the main function of the pooling layer is to extract the maximum features from the convolutional neural network under the guarantee of local invariance of the features, which prevents the data from overfitting and underfitting while compression of the data and enhances the robustness of the neural network. After that, the fully connected network built by the fully connected layer receives the extracted data features and performs classification to evaluate the final output. **Figure 1**. shows the common CNN model for input 2D images including the input layer, convolutional layer, pooling layer, fully connected layer, and output layer:

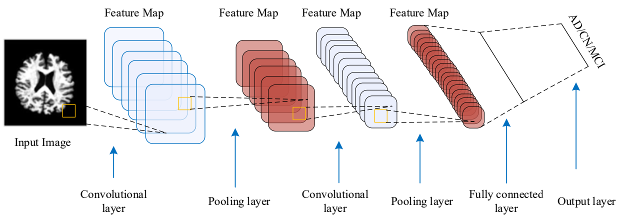

Figure 1. CNN model framework diagram

3.1.1. 2D CNN

2D CNNs for AD diagnosis are usually based on 2D slices of sagittal, cross-sectional, and coronal scans from 2D brain neuroimages such as MRI of the brain, which are mainly selected from structurally or functionally predefined brain regions, and representative features are extracted from each region.

Khagi et al. [18] used the OASIS dataset tuned AlexNet to extract sagittal and cross-sectional lesion features from 2D MRI slices. However, its recognition performance depends on the initial pre-training weights resulting in a weak ability to filter several residual features in MRI images. To remove the irrelevant features from the image, Lee et al. [19] proposed a feature selection method combining entropy slicing and outlier removal to extract the local information from the image, and the accuracy of this model in the binary classification of CN/AD in the test set reached 98.53%.

However, the above studies focus on screening more valuable multi-slices for training, which makes it difficult to capture the subtle lesion information on the image. For this reason, Nawaz et al. [20] proposed a Deep-CNN network. The convolutional layer uses 4~128 size filters to extract the feature representations for each stage of AD classification. However, the deep mesh structure of this algorithm takes up a lot of training time. To solve the above problem, Jain et al. [21] have fine-tuned the pre-trained network VGG-16 based on screened MRI coronal slices to reduce the cost of training time. Unlike the network framework improvement in the above literature, Saratxaga et al. [22] made an attempt at the training strategy and used an efficient CLR triangular learning strategy to construct BrainNet2D convolutional mesh for AD classification, which greatly accelerated the convergence speed of the diagnostic network.

In AD classification diagnosis, 2D CNNs often increase the depth and complexity of the mesh end to improve the nonlinear expression ability of the model, but this approach is accompanied by the proliferation of the number of network parameters and gradient dissipation. Therefore, Tufail et al. [23] used depth-separable convolution to construct a lightweight mesh end, which effectively reduces the number of convolutional parameters by separating the zone city information and channel convolution. To address the gradient problem of deep mesh endings, Puente-Castro et al. [24] introduced the idea of residual learning while deepening mesh endings, and connected feature vectors such as gender and age of subjects with the full connectivity layer of the model to improve the model extensibility and generalization, but the simple fusion of demographics ignores the heterogeneity of the pathogenesis of AD among different races. For this reason, Bae et al. [25] cross-trained the net terminals using two cross-racial datasets, Seoul National University Bundang Hospital (SNUBH) and ADNI, and the AD recognition accuracy reached more than 88% in both datasets.

The above analysis shows that the method of 2D convolutional neural net-termination has the following advantages: (1) The combination of feature selection algorithms such as outlier removal and entropy slicing can improve the feature utilization of 2D sliced images, which can help to improve the accuracy of AD classification. (2) The use of pre-trained models such as AlexNet as initialization or feature extractor saves the detailed and tedious hyper-parameter steps, which can promote the effectiveness of the end-of-network training and learning. (3) MRI imaging is rich in brain details, which is conducive to the construction of deeper 2D convolutional neural nets to be used for high-precision recognition and diagnosis [26, 27].

However, the method also has the following problems that need to be solved: (1) For the evaluation of stereoscopic regional atrophy in MRI, the 2D convolution often needs to analyze multiple cuts, which results in very rough feature extraction and fails to capture the spatial information of the image completely. (2) 2D CNN diagnostic model initializes the net end by migration learning method, although it can reduce the difficulty of model training, in the practical application of AD classification and diagnosis, the stability of target type discrimination is low for complex neuroimages such as MRI. (3) The 2D CNN deep learning algorithm applied to MRI is often based on a single-scale feature extraction method to distinguish between the CN and AD populations, and this two-classification cannot provide effective information about early brain changes to achieve diagnosis and prediction of the MCI stage. (4) The deep neural net-end structure implies the need to deal with large-scale data and a large number of parameters, and it is difficult to avoid the problems of gradient dissipation and net-end performance degradation faced in the training phase.

3.1.2. 3D CNN

Two-dimensional convolutional neural networks (CNNs) have limited ability to express global features based on MRI 2D slices, which can easily lead to the loss of spatial and organizational information of the brain, while 3D CNNs can make better use of 3D features and extract high-resolution features from them, thus effectively improving the AD classification accuracy. In contrast, 3D convolutional neural networks can make better use of the 3D characteristics of images and extract high-resolution features from them, thus effectively improving the classification accuracy of AD. 3D CNNs for AD diagnosis are mainly based on two methods: (1) a morphological method that quantitatively analyses the differences in the local compositions of different brain tissues in the whole-brain MRI images in terms of voxels [28], to measure the occurrence of cerebral atrophy in the brain regions; (2) a morphological method that selects the regions of interest for AD, and then selects the regions of interest for AD, and then selects the areas of interest for AD and then selects the areas of interest for AD. (2) a predefined method to select the region of interest (ROI) of AD to form a 3D image block and extract subtle local lesion features from it in high-dimensional brain images.

Maqsood et al. [29] combined 3D voxels of brain grey matter, white matter, and cerebrospinal fluid into a single image, and used a migration learning method to combine abstract feature representations of MRI brain region structures and AD/CN obtained 89.6% classification accuracy. Considering the high-dimensionality of whole-brain voxel features, Basheera et al. [30] proposed a voxel-by-voxel 3D CNN network based on independent component analysis. The model shows good specificity and sensitivity for grey matter voxel features. To further refine the features at different semantic levels and overcome the limitation of the difficulty in obtaining high-standard samples, Mehmoo [31] and others fine-tuned the VGG-19 network to capture the spatial features of voxels of 3D MRI signals and optimized the training set by combining with data augmentation methods, but the final structure of the network would be accompanied by gradient problems when extracting high-dimensional features. To achieve gradient optimization, Karasawa et al. [32] proposed a ResNet-based 3D convolutional AD diagnostic mesh and removed 50% of the nodes of the network to further simplify the number of parameters. However, achieving a balance between model compression and maintaining performance is still challenging. The above 3D voxel-based methods can assess the global changes in the anatomical structure of a patient's brain but are weak in extracting local small-size features in high-dimensional brain images. To solve this problem, some studies have selected candidate frames of specific regions of interest (ROIs) in AD patient images as feature inputs to the end of convolutional neural nets. Zhu et al. [33] proposed a patch-net with spatial attention blocks to extract the discriminative features of ROI nuggets to improve the classification accuracy of the model.

The above analysis shows that the 3D convolutional neural network method has the following advantages: (1) The 3D convolutional neural network can make full use of the 3D spatial structure information between voxels in MRI images to extract more expressive and clinically meaningful semantic features, to achieve the accurate diagnosis of AD diseases. (2) The voxel-based method can quantitatively detect the density difference of brain tissue without the need of a priori assumption on the region of interest, which is objective and comprehensive. (3) The method based on ROI image blocks can make full use of the effective information of MRI images to extract the subtle local features of high-dimensional brain images.

However, the method also has the following problems that need to be solved: (1) The feature vectors extracted when analyzing whole-brain MRI images are high-dimensional, which results in a long training time for the neural nets, a large number of computational parameters, and a high cost of computational resources, so the future research direction needs to be explored towards the path of a lightweight model under the premise of maintaining the performance of the nets. (2) The training of 3D convolutional neural nets in the field of AD analysis requires a large amount of sample data, and there is a lack of large-scale standard databases, such as ImageNet, so deepening the structure of the nets to improve the performance of the nets may cause overfitting due to the lack of sample size. (3) Abnormalities of the brain lesions in all diseased populations do not always occur in the same brain regions of the selected ROIs, and fixing the brain regions of the same ROIs may result in the loss of critical information used to distinguish patients from others. critical information used to differentiate patients.

3.2. Recurrent neural network (RNN)

RNNs have memory, which CNNs do not have, and are suitable for learning sequential nonlinear features, so they are usually applied to the field of natural language processing. Although RNNs are not often applied to the field of computer vision, there are still a few researchers who have applied RNNs to AD imaging classification tasks because AD is a long-term developmental disease with obvious time series. For example, Abuhmed et al. [34] used RNN to learn the image features of a subject in the past time to predict the future condition of that subject.

It can be visualized from **Figure 2** that the RNN has an input x_t at each moment, and the output ht at the current moment is obtained after the state A_t of the network at moment t. The network state at moment t is jointly determined by the network state at the moment (t1) as well as the inputs, and such a design allows the network to have a memory function in the time series. The bidirectional recurrent neural network (BRNN) and bidirectional long short-term memory (BiLSTM) are both improved models based on RNN.

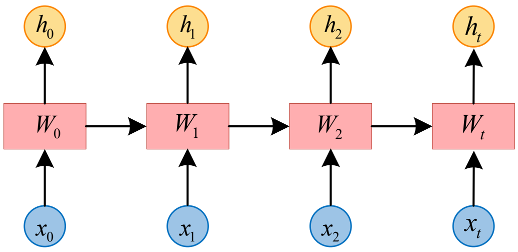

Figure 2. RNN model framework diagram

In terms of model selection, recurrent neural networks such as LSTMs and other models such as CNNs and graph attention layers are usually used together to better make dynamic temporal features useful for AD classification. For example, Haijing Sun et al. [35] combined convolutional neural network with LSTM for diagnosing AD. The classification based on this model extracts spatial features by convolutional neural network, and finally a three-layer LSTM model is constructed to extract the time-varying features of the spatial features, which achieves an accuracy of 93.5% for AD classification. Lu Zhang et al. [36] proposed a DCMAT network consisting of a recurrent neural network map attention layer, which uses the LSTM to deal with dynamic temporal features of MRI signals.

The above analysis shows that the method of recurrent neural network (RNN) has the following advantages:(1)It can fully extract the dynamic temporal features from the patient's time interval follow-up data, thus accelerating the process of early diagnosis of AD. (2) For irregularly collected clinical data, RNN can make the input data stable in dimensional changes and has the ability to store the data for a long period of time. However, the following problems exist in this method: (1) the application of long-time interval follow-up data will generate a very large amount of computational volume and computing time. (2) Although LSTM solves the gradient problem of the traditional RNN, the classification accuracy on small data sets is not ideal.

3.3. Graph neural network (GNN)

Compared with the fully connected layer in the basic network structure of the above neural network, the graph neural network has one more adjacency matrix.GNN transforms the attributes of the input graph such as node edges and global context under the guarantee of graph alignment invariance but does not change its connectivity, and the GNN can achieve information exchange and attribute updating of its critical points through message passing similar to convolutional operation. Based on multi-channel inputs or outputs, the conduction of information data can be better achieved. Compared with CNN, GNN can recognize more complex dynamic graphs and classify them. Currently, graph neural networks include graph convolutional networks, graph generative networks, graph attention networks, and so on, as shown in **Figure 3**. It is also more capable of recognizing the size and structure of graphs without a fixed order of nodes in the processing of graph parameters. The main purpose of using GNN is to extract and integrate features to the nodes, which is especially suitable for fast integration of irregular and unfixed data information. For most of the neural networks that can only deal with the analysis of data based on fixed input formats, GNN stands out for its advantages in processing complex information. Currently, GNNs are being used in chip design, traffic flow sensing and prediction, autonomous driving, drones, medical, and other applications [37, 38].

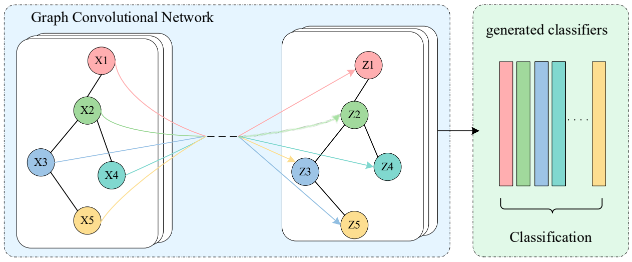

Figure 3. Framework of Graph Convolutional Network

Sarah Parisot et al. [39] first used graph convolutional neural networks to analyze fMRI images in ADNI in 2018. Classification based on the graph convolutional neural network approach begins by representing the population as a sparse graph whose nodes are associated with imaging-based feature vectors. The study evaluation explored the impact of the individual components of the framework on disease prediction, comparing it to an alternative baseline with an accuracy of 80.0% on the ADNI dataset.

Xiaoxiao Li et al. [40] proposed an interpretable graph neural network framework with a new regularised pooling layer, where the regions of interest that are more important for AD classification are calculated from the pooling scores of the nodes in the pooling layer. This method achieves higher accuracy than ordinary graph neural networks by regularising the pooling function. Xin Bi et al. [41] on the other hand, using the research on the final classification of the brain network, proposed a graph neural network with an ExtremeLearning Machine (ELM) aggregator, which has a very fast aggregation speed and powerful aggregation capability. ChundeYang et al. [42] proposed a method called PSGR to classify fMRI images by inputting the transformed brain maps into a graph attention network.

Graph Convolutional Networks can effectively process graph data but cannot handle high-dimensional brain networks and noise well. Lanting Li et al. [43] proposed a sparse brain network ensemble framework TE-HI-GCN model that combines hierarchical graph convolutional networks and migration learning to address this problem. This model proposes an integrated framework involving hierarchical graph convolutional neural networks and sparse brain network migration learning compared to the traditional graph convolutional network model, which can improve network embedding learning for disease diagnosis.

4. Classification of AD based on different modality imaging

4.1. A deep learning approach based on unimodal

A modality is a form in which something occurs or exists and can be information such as sounds, images, and words. For AD early diagnosis research, it can be AD biomarkers such as MRI, PET, and SPECT. The use of different modalities of biomarkers for AD early classification studies is one of the main tasks of deep learning methods applied to biomedical research. Next, the application of different unimodal neuroimaging in AD diagnosis will be presented.

4.1.1. sMRI images for AD classification

AD is a neurodegenerative disease that manifests itself as functional changes caused by structural changes in the patient's brain, which in turn cumulatively cause structural changes. Such structural changes will be well reflected in sMRI images, especially T1-weighted imaging, MRI does not produce ionizing radiation, which is harmless to the human body, and sMRI is also the modality with the largest amount of data in the publicly available dataset, so sMRI images are the most widely used in the imaging-based classification of AD.

YiGiT et al. [44] used axial, sagittal, and coronal planes of sMRI images as inputs to CNNs respectively, and found that the classification accuracy using axial plane projection data reached up to 83% for AD and CN classification tasks: for MCI and CN classification tasks, the classification accuracy using sagittal projection data reached up to 82%. Deep learning-based image classification tasks require a large amount of data as a training set, and CNN, as a supervised learning model, relies on labeled images, and labeled medical images are difficult to obtain. To solve such a problem, Bi et al. [45] proposed an unsupervised learning-based classification model for AD, MCI, and CN, where the training set uses unlabelled SMRI images, combines principal component analysis (PCA) and CNN for feature extraction, and uses k-means clustering algorithm (k-eans) for feature extraction. clustering algorithm (k-means) for classification, the final classification accuracy is 97.01% for AD and MCI, and 91.25% for AD, MCI, and CN.

4.1.2. fMRI images for AD classification

Many researchers have used fMRI to study the AD classification problem. MRI estimates brain activity by detecting blood oxygen level dependence, sacrificing spatial resolution but improving temporal resolution. fMRI can be used to study task-specific relevant brain regions and is often used in psychology and cognitive science. fMRI, among other things, can be used in resting-state MRI to enable the functional regions of the human brain and functional networks to be constructed, and also provides important relationships on the temporal order of regions [46], therefore, for diseases affecting cognitive functions, fMRI can provide information on the function of brain regions compared to sMRI. Parmar et al.. [47] selected 53 3D MRIs with continuous time series from 4DfMRI as the input of 3D CNN, and after five convolutional layers and three fully connected layers, finally obtained the classification model with AD and CN classification accuracy of 94.58%. Bi et al. [48] constructed a brain spectrum based on an AAL board, then constructed a brain network, learned neighbouring positional features by RNN, and finally used an extreme learning machine (ELM) as a classifier, which resulted in an AD and CN classification AUC of 91.3%.

4.1.3. PET images for AD classification

PET is a relatively advanced imaging technique in nuclear medicine, and since one of the distinguishing features of AD is the accumulation of amyloid plaques in the brain, PET imaging of amyloid allows physicians to detect brain plaques in patients with AD, and some studies [49] have shown that the introduction of PET images of amyloid has had a significant impact on the diagnosis of AD disease and that 82% of patients with MCI and 91% of patients with dementia were recommended by clinicians to take medication for AD, respectively, as a result of the significant amyloid deposits in their brains as shown by PET brain scans. For patients with significant amyloid deposits on PET brain scans, 82% of patients with MCI and 91% of patients with dementia were advised by clinicians to take AD-specific medications, whereas only 40% of patients with MCI and 63% of patients with dementia were taking AD-specific medications before PET scanning [50], showing that amyloid positivity is highly correlated with AD, which can be evidenced by the use of deep learning AD image classification based on PET data. which can also be corroborated in deep learning AD image classification based on PET data.

Punjabi et al. [51] compared in detail the effect of SMRI images and AV-45 amyloid PET images on AD and CN classification efficacy. In the paper, 1299 sMRI data and 585 AV-45 amyloid PET data were used, and the CNN was trained using all SMRI data, all PET data, and the same amount of SMRI data as PET data, respectively, and the results showed that only about half the amount of PET data from sMRI was used to achieve 85.15% accuracy rate in AD and CN classification accuracy.

4.2. A deep learning approach based on multimodality

Multimodal deep learning is relative to unimodal and refers to the use of information from multiple unimodal modalities in deep learning methods to achieve information fusion between different modalities. Many researchers have used multimodal data for AD early classification studies[52, 53].

Forouzannzhad et al. [54] combined positron emission tomography (PET) and magnetic resonance imaging (MRI) multimodal imaging techniques and standard neuropsychological test scores to conduct a classification study for the early diagnosis of AD using a deep neural network DNN (Deep Neural Networks). The classification accuracy of CN with early mild cognitive impairment EMCI was as high as 84.0% for normal controls, and 84.1%,96.8%,69.5%,90.3%, and 80.2% for CN with late mild cognitive impairment LMCI, CN with AD, EMCI with LMCI, EMCI with AD, and LMCI with AD, respectively. In contrast, the classification accuracy of CN with EMCI on MRI images alone was only 68.0%. The study demonstrated that the multimodal approach was superior to unimodal image analysis.

Kang et al. [55] constructed a migration learning method for the VGG16 model based on SMRI and Diffusion Tensor Imaging DTI (Diffusion Tensor Imaging) bimodal data for classification study of EMCI and CN. The data were obtained from the ADNI dataset, and the experiments used the multimodal fusion strategy to merge the slices with the same index into RGB slices to form a slice dataset for inputting into the model for training, and LASSO (Least Absolute Shrinkage and Selection Operator) algorithm was used to extract the part of the features related to the disease of EMCI, and 94.2% of the features were obtained in the experiment. The experiment obtained a classification accuracy of 94.2% and a sensitivity of 97.3%. The experimental results show that multimodal data can provide more and more useful information for distinguishing EMCI and CN and validate that DTI images can be used as an important biomarker for EMCI from a clinical point of view.

Khvostikov et al. [56] compared the unimodal data used in experiments by fusing SMRI and DTI imaging modalities in the hippocampal region of interest city and the AD classification algorithm based on the 3D-CNN model and obtained an accuracy of 93.3% in the multimodal case in the classification of AD and MCI, which is a big advantage compared to the 65.8% in the unimodal case of sMRI. This is an advantage over the 65.8% accuracy in the sMRI unimodal case. To balance the classes with different sizes of data in the experiments, the data augmentation method is used to eliminate the influence of different sizes of data on the net final training process.

5. Challenges and prospects

5.1. Challenges

With population aging becoming a global trend, AD has become one of the leading causes of death among people in developed countries. As a neurodegenerative brain disease, AD seriously affects the quality of life of patients and families. Therefore, early diagnosis and detection of AD are of great importance to patients and their families. In this paper, we systematically review the current status of the application of deep learning on AD classification [57] and summarise the latest research progress of the model, which shows that deep learning has a key role in the early diagnosis of AD, but some problems need to be solved at present:

(1) Model performance is limited by data sources. Existing studies based on MRI, PET, and other imaging histology in AD are mostly retrospective analyses of different scanning equipment, different imaging parameters and scanning modalities in different medical centers will vary. Therefore, it will affect the classification effect of the neural network model on diseases, and there will be a situation that the model is trained better on a certain dataset, but performs poorly on other datasets, which makes it difficult to popularise the application in the actual clinic.

(2) The diagnostic performance of the multimodal model needs to be improved. The fusion algorithms for data in different formats (image and laboratory data) need to be improved. Considering the practical clinical application of multimodal technology, the model may be affected by the loss of its modal data, and if only the existing complete modal data are used for training. It will further exacerbate the small sample problem and lead to the degradation of the model performance.

(3) The "black box" nature of deep learning methods leads to poor model interpretability. The deep neural network includes multiple hidden layers, which leads to great uncertainty in the feature selection and decision-making process. Deep learning-based AD classification of three-dimensional, multimodal medical images involves nonlinear convolution and pooling of different dimensions from the source data, making it difficult to interpret the importance of feature recognition in the original data.

5.2. Prospects

Combining the above issues, the future development direction of deep learning in the classification application of AD should focus on improving the algorithmic accuracy, increasing the model's ability to learn from small samples, and optimising the existing deep learning model to meet the needs of classifying AD at multiple stages. In the future, further research can be carried out in the following aspects:

(1) AD classification study based on incomplete multimodal images

A common problem in AD assisted diagnosis studies based on multimodal images is the problem of missing images. In the clinic, most subjects refuse to undergo PET scanning due to the high cost of PET scanning or radiological hazards. In previous multimodal fusion studies, most of the subjects with missing modal data were directly abandoned for adoption. In recent years, adversarial generative networks (GANs) have been successfully applied to learn bi-directional mappings between related images. Since MRI and PET images of the same subject are potentially correlated, the bi-directional mapping between MRI images and PET images can be learnt with the help of GAN to generate the missing modal data of the subject, saving a lot of human and financial resources.

(2) AD classification research based on different biomarkers

Improve the accuracy of early diagnosis of AD by fusing data from different biomarkers. The etiology of AD is complex and heterogeneous, and fusing multimodal is better than unimodal methods for classification. Adding other biomarkers based on neuroimaging, such as clinical diagnostic data and genetic data, is conducive to further understanding the underlying physiological mechanisms of Alzheimer's disease and improves the classification accuracy for MCI transformation prediction.

(3) AD classification research based on interpretability

Combining visualisation techniques to improve deep learning model interpretability. Deep neural networks combined with visual analysis methods such as class activation maps to understand the distribution characteristics of focal brain regions, improve model transparency while ensuring decision-making accuracy, further explain the relationship between imaging features and diagnostic results, and assist doctors in clinical decision-making. In addition, the interpretability of the model can be further enhanced by inputting ablation experiments, usually by removing or retaining certain brain regions in the input image to investigate which brain regions contribute more to AD image classification. If the removal of certain brain regions does not have a negative effect on the classification result, these brain regions are considered to have little contribution; if the addition of certain brain regions has a positive effect on the classification result, these brain regions are considered to be favourable for AD classification.

References

[1] F. Li, M. Liu, and A. s. D. N. Initiative, "A hybrid convolutional and recurrent neural network for hippocampus analysis in Alzheimer's disease," *Journal of neuroscience methods,* vol. 323, pp. 108-118, 2019.

[2] Y. Zhang, "Classification of Alzheimer Disease based on structural magnetic resonance imaging by kernel support vector machine decision tree," *Progress in Electromagnetics Research,* vol. 144, pp. 185-191, 2014.

[3] Y. Zhang, "Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC," *Biomedical Signal Processing and Control,* vol. 21, pp. 58-73, 2015.

[4] M. Dünnwald, P. Ernst, E. Düzel, K. Tönnies, M. J. Betts, and S. Oeltze-Jafra, "Fully automated deep learning-based localization and segmentation of the locus coeruleus in aging and Parkinson's disease using neuromelanin-sensitive MRI," *International Journal of Computer Assisted Radiology and Surgery,* vol. 16, pp. 2129-2135, 2021.

[5] K. Imamura, Y. Yada, Y. Izumi, M. Morita, A. Kawata, T. Arisato*, et al.*, "Prediction model of amyotrophic lateral sclerosis by deep learning with patient induced pluripotent stem cells," *Annals of neurology,* vol. 89, pp. 1226-1233, 2021.

[6] E. Yee, D. Ma, K. Popuri, L. Wang, M. F. Beg, and A. s. D. N. Initiative, "Construction of MRI-based Alzheimer's disease score based on efficient 3D convolutional neural network: Comprehensive validation on 7,902 images from a multi-center dataset," *Journal of Alzheimer's Disease,* vol. 79, pp. 47-58, 2021.

[7] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and A. s. D. N. Initiative, "Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network," *Frontiers in neuroscience,* vol. 13, p. 509, 2019.

[8] M. Liu, D. Zhang, D. Shen, and A. s. D. N. Initiative, "Ensemble sparse classification of Alzheimer's disease," *NeuroImage,* vol. 60, pp. 1106-1116, 2012.

[9] W. Feng, N. V. Halm-Lutterodt, H. Tang, A. Mecum, M. K. Mesregah, Y. Ma*, et al.*, "Automated MRI-based deep learning model for detection of Alzheimer's disease process," *International Journal of Neural Systems,* vol. 30, p. 2050032, 2020.

[10] M. Kavitha, N. Yudistira, and T. Kurita, "Multi instance learning via deep CNN for multi-class recognition of Alzheimer's disease," in *2019 IEEE 11th international workshop on computational intelligence and applications (IWCIA)*, 2019, pp. 89-94.

[11] L. Nanni, M. Interlenghi, S. Brahnam, C. Salvatore, S. Papa, R. Nemni*, et al.*, "Comparison of transfer learning and conventional machine learning applied to structural brain MRI for the early diagnosis and prognosis of Alzheimer's disease," *Frontiers in neurology,* vol. 11, p. 576194, 2020.

[12] Y. Zhang, "Detection of Alzheimer's disease by displacement field and machine learning," *PeerJ,* vol. 3, Article ID: e1251, 2015.

[13] S. Wang, "Detection of Alzheimer's Disease by Three-Dimensional Displacement Field Estimation in Structural Magnetic Resonance Imaging," *Journal of Alzheimer's Disease,* vol. 50, pp. 233-248, 2016.

[14] C. Hinrichs, V. Singh, G. Xu, and S. Johnson, "MKL for robust multi-modality AD classification," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009: 12th International Conference, London, UK, September 20-24, 2009, Proceedings, Part II 12*, 2009, pp. 786-794.

[15] A. Chaddad, C. Desrosiers, and M. Toews, "Local discriminative characterization of MRI for Alzheimer's disease," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 1-5.

[16] I. B. Malone, D. Cash, G. R. Ridgway, D. G. MacManus, S. Ourselin, N. C. Fox*, et al.*, "MIRIAD—Public release of a multiple time point Alzheimer's MR imaging dataset," *NeuroImage,* vol. 70, pp. 33-36, 2013.

[17] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, 2021.

[18] B. Khagi, G. R. Kwon, and R. Lama, "Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques," *International Journal of Imaging Systems and Technology,* vol. 29, pp. 297-310, 2019.
[19] B. Lee, W. Ellahi, and J. Y. Choi, "Using deep CNN with data permutation scheme for classification of Alzheimer's disease in structural magnetic resonance imaging (sMRI)," *IEICE TRANSACTIONS on Information and Systems,* vol. 102, pp. 1384-1395, 2019.
[20] A. Nawaz, S. M. Anwar, R. Liaqat, J. Iqbal, U. Bagci, and M. Majid, "Deep convolutional neural network based classification of Alzheimer's disease using MRI data," in *2020 IEEE 23rd International Multitopic Conference (INMIC)*, 2020, pp. 1-6.
[21] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research,* vol. 57, pp. 147-159, 2019.
[22] C. L. Saratxaga, I. Moya, A. Picón, M. Acosta, A. Moreno-Fernandez-de-Leceta, E. Garrote*, et al.*, "MRI deep learning-based solution for Alzheimer's disease prediction," *Journal of personalized medicine,* vol. 11, p. 902, 2021.
[23] A. B. Tufail, Y.-K. Ma, and Q.-N. Zhang, "Binary classification of Alzheimer's disease using sMRI imaging modality and deep learning," *Journal of digital imaging,* vol. 33, pp. 1073-1090, 2020.
[24] A. Puente-Castro, E. Fernandez-Blanco, A. Pazos, and C. R. Munteanu, "Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques," *Computers in biology and medicine,* vol. 120, p. 103764, 2020.
[25] J. B. Bae, S. Lee, W. Jung, S. Park, W. Kim, H. Oh*, et al.*, "Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging," *Scientific reports,* vol. 10, p. 22252, 2020.
[26] S. H. Wang and Y. D. Lv, "Alcoholism Detection by Data Augmentation and Convolutional Neural Network with Stochastic Pooling," *Journal of Medical Systems,* vol. 42, Article ID: 2, 2018.
[27] Y.-D. Zhang, "Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform," *Multimedia Tools and Applications,* vol. 77, pp. 22821-22839, 2018.
[28] A. Mechelli, C. J. Price, K. J. Friston, and J. Ashburner, "Voxel-based morphometry of the human brain: methods and applications," *Current Medical Imaging,* vol. 1, pp. 105-113, 2005.
[29] M. Maqsood, F. Nazir, U. Khan, F. Aadil, H. Jamal, I. Mehmood*, et al.*, "Transfer learning assisted classification and detection of Alzheimer's disease stages using 3D MRI scans," *Sensors,* vol. 19, p. 2645, 2019.
[30] S. Basheera and M. S. S. Ram, "Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation," *Alzheimer's & Dementia: Translational Research & Clinical Interventions,* vol. 5, pp. 974-986, 2019.
[31] A. Mehmood, S. Yang, Z. Feng, M. Wang, A. S. Ahmad, R. Khan*, et al.*, "A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images," *Neuroscience,* vol. 460, pp. 43-52, 2021.
[32] H. Karasawa, C.-L. Liu, and H. Ohwada, "Deep 3d convolutional neural network architectures for alzheimer's disease diagnosis," in *Intelligent Information and Database Systems: 10th Asian Conference, ACIIDS 2018, Dong Hoi City, Vietnam, March 19-21, 2018, Proceedings, Part I 10*, 2018, pp. 287-296.
[33] W. Zhu, L. Sun, J. Huang, L. Han, and D. Zhang, "Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI," *IEEE Transactions on Medical Imaging,* vol. 40, pp. 2354-2366, 2021.
[34] T. Abuhmed, S. El-Sappagh, and J. M. Alonso, "Robust hybrid deep learning models for Alzheimer's progression detection," *Knowledge-Based Systems,* vol. 213, p. 106688, 2021.
[35] H. Sun, A. Wang, and S. He, "Temporal and spatial analysis of alzheimer's disease based on an improved convolutional neural network and a resting-state FMRI brain functional network," *International Journal of Environmental Research and Public Health,* vol. 19, p. 4508, 2022.
[36] L. Zhang, L. Wang, and D. Zhu, "Jointly Analyzing Alzheimer's Disease Related Structure-Function Using Deep Cross-Model Attention Network," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 563-567.
[37] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, "EEG-GNN: Graph neural networks for classification of electroencephalogram (EEG) signals," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021, pp. 1061-1067.
[38] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu*, et al.*, "Graph neural networks: A review of methods and applications," *AI open,* vol. 1, pp. 57-81, 2020.
[39] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker*, et al.*, "Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease," *Medical image analysis,* vol. 48, pp. 117-130, 2018.
[40] X. Li, Y. Zhou, N. C. Dvornek, M. Zhang, J. Zhuang, P. Ventola*, et al.*, "Pooling regularized graph neural network for fmri biomarker analysis," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23*, 2020, pp. 625-635.
[41] X. Bi, Z. Liu, Y. He, X. Zhao, Y. Sun, and H. Liu, "GNEA: a graph neural network with ELM aggregator for brain network classification," *Complexity,* vol. 2020, pp. 1-11, 2020.
[42] C. Yang, P. Wang, J. Tan, Q. Liu, and X. Li, "Autism spectrum disorder diagnosis using graph attention network based on spatial-constrained sparse functional brain networks," *Computers in Biology and Medicine,* vol. 139, p. 104963, 2021.
[43] L. Li, H. Jiang, G. Wen, P. Cao, M. Xu, X. Liu*, et al.*, "TE-HI-GCN: An ensemble of transfer hierarchical graph convolutional networks for disorder diagnosis," *Neuroinformatics*, pp. 1-23, 2021.
[44] A. YİĞİT and Z. Işik, "Applying deep learning models to structural MRI for stage prediction of Alzheimer's disease," *Turkish Journal of Electrical Engineering and Computer Sciences,* vol. 28, pp. 196-210, 2020.
[45] X. Bi, S. Li, B. Xiao, Y. Li, G. Wang, and X. Ma, "Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology," *Neurocomputing,* vol. 392, pp. 296-304, 2020.
[46] H. Guo and Y. Zhang, "Resting state fMRI and improved deep learning algorithm for earlier detection of Alzheimer's disease," *IEEE Access,* vol. 8, pp. 115383-115392, 2020.

[47] H. S. Parmar, B. Nutter, R. Long, S. Antani, and S. Mitra, "Deep learning of volumetric 3D CNN for fMRI in Alzheimer's disease classification," in *Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2020, pp. 66-71.

[48] X. Bi, X. Zhao, H. Huang, D. Chen, and Y. Ma, "Functional brain network classification for Alzheimer's disease detection with deep features and extreme learning machine," *Cognitive Computation,* vol. 12, pp. 513-527, 2020.

[49] F. Gao, "Integrated positron emission tomography/magnetic resonance imaging in clinical diagnosis of Alzheimer's disease," *European Journal of Radiology,* vol. 145, p. 110017, 2021.

[50] G. D. Rabinovici, C. Gatsonis, C. Apgar, K. Chaudhary, I. Gareen, L. Hanna*, et al.*, "Association of amyloid positron emission tomography with subsequent change in clinical management among medicare beneficiaries with mild cognitive impairment or dementia," *Jama,* vol. 321, pp. 1286-1294, 2019.

[51] A. Punjabi, A. Martersteck, Y. Wang, T. B. Parrish, A. K. Katsaggelos, and A. s. D. N. Initiative, "Neuroimaging modality fusion in Alzheimer's classification using convolutional neural networks," *PloS one,* vol. 14, p. e0225759, 2019.

[52] J. Zhang, X. He, Y. Liu, Q. Cai, H. Chen, and L. Qing, "Multi-modal cross-attention network for Alzheimer's disease diagnosis with multi-modality data," *Computers in Biology and Medicine,* vol. 162, p. 107050, 2023.

[53] Y. Zhang, X. He, Y. H. Chan, Q. Teng, and J. C. Rajapakse, "Multi-modal graph neural network for early diagnosis of Alzheimer's disease from sMRI and PET scans," *Computers in Biology and Medicine,* vol. 164, p. 107328, 2023.

[54] P. Forouzannezhad, A. Abbaspour, C. Li, M. Cabrerizo, and M. Adjouadi, "A deep neural network approach for early diagnosis of mild cognitive impairment using multiple features," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, 2018, pp. 1341-1346.

[55] L. Kang, J. Jiang, J. Huang, and T. Zhang, "Identifying early mild cognitive impairment by multi-modality MRI-based deep learning," *Frontiers in aging neuroscience,* vol. 12, p. 206, 2020.

[56] A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, "3D CNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies," *arXiv preprint arXiv:1801.05968*, 2018.

[57] Y. Zhang, "Three-Dimensional Eigenbrain for the Detection of Subjects and Brain Regions Related with Alzheimer's Disease," *Journal of Alzheimer's Disease,* vol. 50, pp. 1163-1179, 2016.