Journal of Alzheimer's Disease 92 (2023) 209–228
DOI 10.3233/JAD-220519
IOS Press

Fully Connected Multi-Kernel Convolutional Neural Network Based on Alzheimer's Disease Diagnosis

Lan Deng and Yuanjun Wang*, Alzheimer's Disease Neuroimaging Initiative[1]
School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai, P.R. China

Accepted 28 December 2022
Pre-press 25 January 2023

Abstract.
Background: There is a shortage of clinicians with sufficient expertise in the diagnosis of Alzheimer's disease (AD), and cerebrospinal fluid biometric collection and positron emission tomography diagnosis are invasive. Therefore, it is of potential significance to obtain high-precision automatic diagnosis results from diffusion tensor imaging (DTI) through deep learning, and simultaneously output feature probability maps to provide clinical auxiliary diagnosis.
Objective: We proposed a factorization machine combined neural network (FMCNN) model combining a multi-function convolutional neural network (MCNN) with a fully convolutional network (FCN), while accurately diagnosing AD and mild cognitive impairment (MCI); corresponding fiber bundle visualization results are generated to describe their status.
Methods: First, the DTI data is preprocessed to eliminate the influence of external factors. The fiber bundles of the corpus callosum (CC), cingulum (CG), uncinate fasciculus (UNC), and white matter (WM) were then tracked based on deterministic fiber tracking. Then the streamlines are input into CNN, MCNN, and FMCNN as one-dimensional features for classification, and the models are evaluated by performance evaluation indicators. Finally, the fiber risk probability map is output through FMCNN.
Results: After comparing the model performance indicators of CNN, MCNN, and FMCNN, it was found that FMCNN showed the best performance in the indicators of accuracy, specificity, sensitivity, and area under the curve. By inputting the fiber bundles of the 10 regions of interest (UNC_L, UNC_R, UNC, CC, CG, CG+UNC, CG+CC, CC+UNC, CG+CC+UNC, and WM into CNN, MCNN, and FMCNN, respectively), WM shows the highest accuracy in CNN, MCNN, and FMCNN, which are 88.41%, 92.07%, and 96.95%, respectively.
Conclusion: The FMCNN proposed here can accurately diagnose AD and MCI, and the generated fiber probability map can represent the risk status of AD and MCI.

Keywords: Alzheimer's disease, deep learning, fiber tracking, fully convolutional networks

INTRODUCTION

Tens of millions of people worldwide still suffer from Alzheimer's disease (AD) [1] and attempts to develop effective disease-modifying therapies remain stagnant. Although researchers have made great strides in detecting AD pathology using cerebrospinal fluid biomarkers [2–4], positron emission tomography amyloid [5, 6], and tau imaging [7, 8], these models are usually limited to the research context. Current diagnostic criteria rely on highly skilled neurologists, who examine patients by taking a

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Correspondence to: Yuanjun Wang, School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, P.R. China. E-mail: yjusst@126.com.

patient history, objective cognitive assessments (e.g., Mini-Mental State Examination or other neuropsychological testing [9]), and diagnosis by structural magnetic resonance imaging (MRI) underlying AD. Clinic pathological studies have shown that the diagnostic sensitivity of clinicians is between 70.9% and 87.3%, and the specificity is between 44.3% and 70.8% [10]. Although MRI reveals characteristic brain changes noted in AD (e.g., hippocampal and temporal lobe atrophy [11]), these features lack specificity for the diagnosis of AD [12–14]. Current AD-based diagnosis faces the following challenges: Diagnosis of AD based on features such as hippocampal and temporal lobe atrophy is relatively imprecise; acquisition of cerebrospinal fluid biosignatures and positron emission tomography diagnosis are invasive; and there is a lack of adequate clinicians with relevant expertise. Based on this, high-precision automatic diagnosis results can be obtained from MRI through deep learning to provide auxiliary diagnosis for clinical use. Research such as that by Hinton [15] and Topol [16], based on deep learning [17], provides methods to derive high-accuracy predictions from MRI data collected within the scope of neurology practice. Qiu et al. [18] used convolutional neural networks (CNN) to classify the cognitive state of subjects based on MRI and multimodal data, which also proved the effectiveness of the current researchers' application of deep learning methods. Despite the superior results of these experiments, these models have not yet been fully integrated into clinical practice. Additionally, in the biomedical field, more and more people consider deep learning models to be "black box" algorithms [19]. In other words, while deep learning models have demonstrated high-accuracy classification across a broad range of diseases, they neither illuminate underlying diagnostic decisions nor identify input features relevant to output predictions, therefore, the clinical potential of deep learning is undermined by the increasing use of opaque decision frameworks. Finally, computerized individual-level representations of AD remain unresolved, because of the uncertainty of AD onset and the heterogeneity of symptoms. On the other hand, considering that the pathological features of AD are mainly senile plaques and fibrillary entanglement, it shows that from a clinical point of view, the neurological differences between AD patients and normal elderly are more significant. Therefore, it makes sense to use the connectivity or fibers between brain regions as input data for diagnosis, as Lella et al. [20] developed a machine learning framework based on whole-brain communicability for the classification and characterization of AD importance analysis. This study compared the performance of three state-of-the-art classification models [support vector machines (SVM), random forest (RF), and artificial neural networks (ANN)] on brain connectivity networks of normal control (NC) and AD patients in the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Furthermore, this experiment clinically validates the information content of the communication competence indicator by performing a feature importance analysis. Overcoming these challenges is not only critical to harnessing the potential of deep learning algorithms to improve patient care, but also paves the way for explainable evidence-based machine learning in the medical imaging community. Based on the above two points, we extract fiber information from DTI data and classify it by deep learning. We improve the traditional CNN, combine multi-function convolutional neural networks (MCNN) [21] with fully convolutional networks (FCN) [22], and propose a novel deep learning framework, factorization machine combined neural network (FMCNN) model, which can accurately diagnose AD while generating high-resolution fiber bundle visualization results of AD risk, using for accurate prediction of AD and mild cognitive impairment (MCI) status.

The flow of the classification algorithm proposed in this paper is shown in Fig. 1. First, the diffusion tensor imaging (DTI) data are preprocessed including eddy current correction, head motion correction, spatial standardization, and scalp removal to eliminate the influence of external factors on the data itself. The preprocessed DTI was then tracked by deterministic fiber tracking, and the seed points were set in the corpus callosum (CC), cingulum (CG), uncinate fasciculus (UNC), and white matter (WM) to track the fiber bundles corresponding to the region of interest. The fiber bundle streamlines are input into CNN, MCNN, and FMCNN as one-dimensional features for classification and the models are evaluated by performance evaluation indicators. Finally, the AD fiber risk probability map is output through FMCNN.

METHODS

Data

Data used in the preparation of this article were obtained from the ADNI database (https://adni.loni.usc.edu). The ADNI was launched

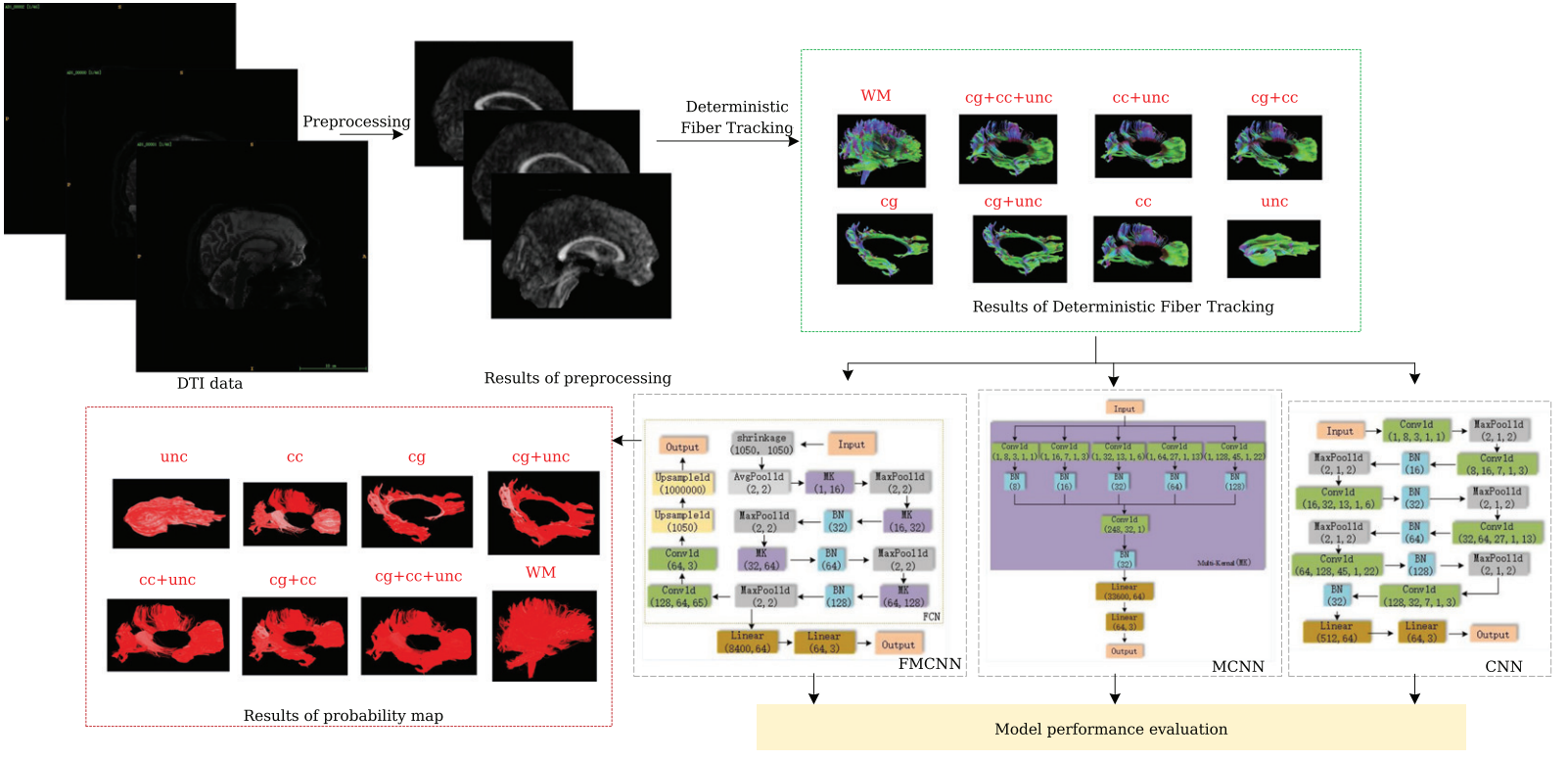


Fig. 1. The structural framework of this paper.

in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up -to-date information, see https://www.adni-info.org.

The ADNI is a longitudinal multicenter study to develop clinical, imaging, genetic, and biochemical biomarkers for early detection and tracking of AD [23]. Model training, internal validation, and testing were performed on the ADNI dataset. Selection criteria included individuals aged 55-75 years with a 1.5 T scan within 6 months of the date of clinical diagnosis of AD or NC. We excluded cases including AD with mixed dementia, non-AD dementia, history of severe traumatic brain injury, major depression, stroke and brain tumor, and occasional major systemic disease. These inclusion and exclusion criteria were adapted from the baseline recruitment protocol developed by the ADNI study [23]. Finally, we selected DTI data of 413 individuals from the ADNI cohort, including 121 AD, 130 MCI, and 162 NC. If the subject underwent multiple MRI scans within the time window, the experiment selected the scan closest to the date of clinical diagnosis.

Preprocessing

This step of image preprocessing is to eliminate differences in anatomical structures that exist between different subjects' brains and improve the signal-to-noise ratio of neuroimaging data. The key steps of preprocessing are as follows: 1) Non-uniformity correction: to eliminate external factors such as the non-uniformity of the radio frequency coil, the non-uniformity of the main magnetic field, and the nonlinear gradient field; 2) Head movement correction: to remove the error caused by the slight movement of the subject's head during data collection; 3) Tissue segmentation: to avoid background other than brain tissue influence on the results; 4) Data enhancement: firstly flip the data up and down to double the amount of data, and then flip left and right to quadruple the amount of data.

Deterministic fiber tracking

Deterministic fiber tracking algorithms incorporate local estimates of diffusion direction information. Once we know the orientation of the fibers at each voxel, we can connect these orientations to reconstruct the full trajectory, approximating the anatomical bundle. The simplest calculation flow is: start at the seed point position and follow the preferred direction until we reach a new voxel; then we can change to the reference direction for this voxel and continue until the entire fiber bundle is traced.

The two most well-known algorithms for fiber tracking are deterministic fiber tracking [24, 25] and probabilistic fiber tracking [26]. If the fiber model being used contains uncertainty, i.e., there is an error in estimating the fiber orientation at each voxel, then the trajectory propagation algorithm belongs to the probability domain. If it does not assume any uncertainty along the path of the trajectory, then it belongs to the deterministic domain. There are many classical trajectory propagation methods, such as global trajectory [8]. One of the simplest and earliest deterministic methods is Fibre Assignment by Continuous Tracking (FACT) [25]. The FACT algorithm starts with an input at any point in a voxel and propagates in two directions, forward and backward. The most important part of these tracking methods is the way they decide to stop tracking. FACT uses a single threshold variable R to decide:

$$R = \sum_{i}^{s} \sum_{j}^{s} \frac{|\rightharpoonup e_i \cdot \rightharpoonup e_j|}{s(s-1)} \tag{1}$$

where s is the number of adjacent voxels and e is the eigenvector corresponding to the highest eigenvalue in each voxel. The simplest case for defining a voxel neighborhood is to use all other 26 neighboring (contacting) voxels. When adjacent fibers are aligned, R will be close to 1 because the normalized vectors become more collinear so the absolute value of each dot product will be close to 1. When the adjacent fiber directions are not aligned, R becomes small in areas of discordance. In voxels where R is less than a predetermined threshold, for example, the threshold is set to 0.8, point tracking will stop and FACT will terminate. A significant problem with FACT is that it cannot track in areas with intersections. When there are fiber intersections in the voxel, it only traces one of the main paths in the intersection area. Deterministic algorithms track trajectories by making a series of discrete locally optimal decisions. We use streamlines to delineate them. The main disadvantage of deterministic algorithms is that they are susceptible to local noise.

The cingulate tract, uncinate, and corpus callosum are key components known to be clinically associated

with cognitive function and AD [27–31]. In particular, the cingulate tract includes important connections to the default mode network and the hippocampal memory network. The uncinate bundle is an important pathway in the frontolimbic network. The corpus callosum is responsible for communication between the cerebral hemispheres and plays an important role in overall brain integration. Therefore, this paper investigated the relationship between the white matter of these three regions and their combined regions and whole brain with AD, MCI, and NC.

In this experiment, we used DSI-Studio software to track the CC, CG, UNC, and WM by setting different seed points based on deterministic fiber tracking. And then we extracted features from these fiber bundles. One million points were tracked for each region of interest (ROI). The fiber bundles of the CC include the minor forceps of corpus callosum, the body of the corpus callosum, the tapetum of the corpus callosum, and the major forceps of corpus callosum; the fiber bundles of the CG include the left and right fiber bundles connecting the frontal and the parahippocampal gyrus, the left and right fiber bundles connecting the frontal and the parietal, the left and right fiber bundles connecting the parahippocampal gyrus and the parietal lobe, the left and right fibers surrounding the hippocampus, and the left and right fibers surrounding the parolfactory area; the uncinate fasciculus consists of left and right fiber bundles.

Classification model building

The experiment in this part compares CNN, MCNN, and FMCNN to verify the improvement of our constructed framework FMCNN on MCNN and CNN. The training set of the input classification model is 1000 blocks of size 1×1050.

The CNN model consists of 6 convolutional layers and 2 fully connected layers, and the five convolutional layers are followed by BN layers and maximum pooling layers. The epoch of CNN is set to 3000, the batch size is set to 10, and the optimal learning rate is 0.0053 through the global search algorithm. The specific frame diagram is shown in Fig. 2.

The MCNN model is an improvement on the CNN model which introduces the concept of multi-core, aiming to more comprehensively extract the abstract feature information in fiber bundles by connecting different convolution kernels in parallel to fuse abstract features at different levels. MCNN consists of 1 multi-kernel layer and 2 fully connected layers. The multi-kernel layer consists of 6 convolutional layers, of which 5 convolutional layers are connected in parallel, and the kernel sizes of these 5 convolutional layers are 3, 7, 13, 27, and 45, respectively. A batch normalization (BN) layer is connected after each convolutional layer for normalization. The epoch of the MCNN model is set to 3000, the batch size is set to 10, and the optimal learning rate is 0.0003 through the global search algorithm. The specific frame diagram is shown in Fig. 3.

FMCNN consists of 4 multi-kernel layers and 1 variable classification block. For classification tasks, the last 1 variable classification block acts as a dense layer to improve model efficiency [32]. The network is retrained with random initial weights. We used the Adam optimizer with a learning rate of 0.0001 and a batch size of 10, where the learning rate is the optimal learning rate calculated by the global search algorithm. During training, the model is saved when it reaches the minimum error on the ADNI validation dataset. After FCN training, fiber streamlines are forwarded to obtain a complete array of disease probabilities, which we refer to as disease probability maps.

The FMCNN is trained through the 17 layers network to obtain the model related parameters, and then tested through the 19 layers network model and output the corresponding probability map. The input data is a 1-dimensional vector and the channel is 1. The training is performed by repeatedly applying the data to the training set, and the final output shape is 3×1. These values can be transformed into respective AD, NC, and MCI probabilities by applying the softmax function, and then the larger of the two probabilities is used for the classification of disease states.

After evaluating the classification stability of the model, the variable classification block is transformed into a disease probability map of the subjects generated by 2 convolutional layers and 2 upsampling layers. AD status is predicted by selecting an AD probability value from a disease probability map. This choice is based on observations of the overall performance of the FCN classifier, estimated using area under the curve (AUC), accuracy (ACC), specificity (SPE), and sensitivity (SEN) on ADNI training data. The FMCNN framework is shown in Fig. 4.

Evaluation metrics for classification models

The model in this paper is constructed based on ADNI data, which is randomly divided into three groups with a ratio of 8 : 1:1 for training, validation, and testing. Models are built on a per-train and val-

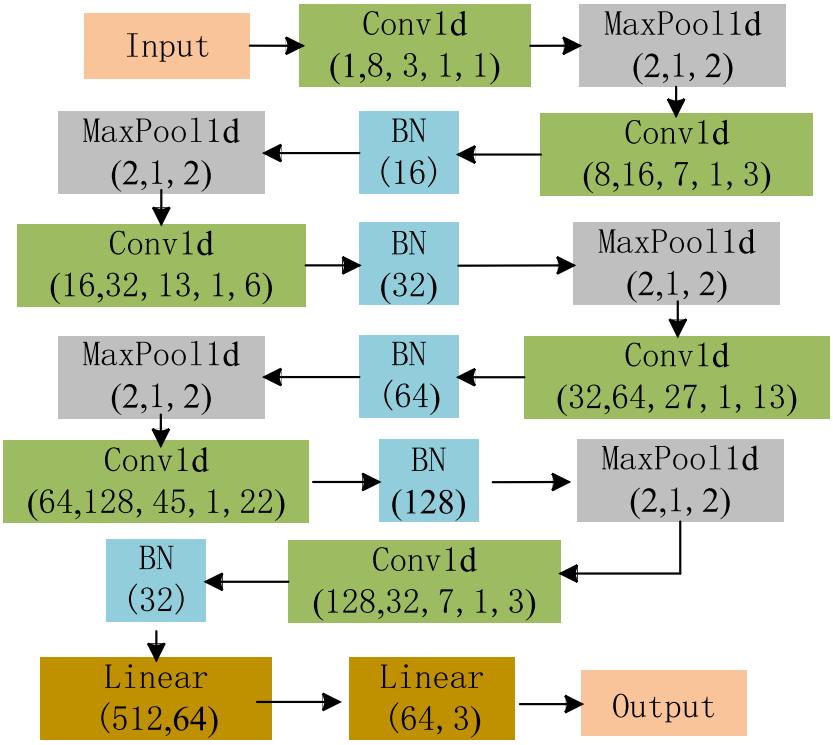


Fig. 2. CNN framework diagram.

idation basis and evaluated on performance on the test dataset. The performance measures of the diagnostic model in this paper will use SEN, SPE, ACC, precision-recall (P-R) curve, average precision (AP) value, receiver operating characteristic (ROC) curve, and AUC were quantified.

SEN, SPE, and ACC can be calculated from the four classification test result values of true positive (TP), true negative (TN), false negative (FN), and false positive (FP). In this paper, AD was taken as Positive and MCI and NC as negative. When both the label and the sample prediction classification are AD, the classification result is considered as TP. Similarly, when the label and sample prediction classification are the same, MCI or NC, the classification result is considered as TN. When labeled AD and predicted as MCI or NC, the classification result is considered to be FP. When labeled MCI or NC and predicted as AD, the classification result is considered to be FN.

1) SEN

SEN is defined as the proportion of true positives that the classification model can correctly identify, and it describes the ability of the classification model to predict a positive sample as a positive sample. The numerical value of SEN represents the probability that the classification model will classify a sample with the actual label of 1 as 1. The higher the SEN number, the less likely the classification model will return FP results. Its formula is as follows:

$$SEN = TPR = \frac{TP}{TP + FN} \tag{2}$$

2) SPE

SPE is the proportion of TN that the classification model can correctly identify, and it describes the ability of the classification model to predict negative samples as negative samples. The numerical value of specificity represents the probability of classifying a sample with an actual label of 0 as 0 in a test of the classification model without giving a FP result. The SPE of the classification model is 99% means when we perform a classification test on a sample whose actual label is 0, the probability that this sample will be classified as 0 by the classification model is 99%. Its formula is as follows:

$$SPE = \frac{TN}{TN + FP} \tag{3}$$

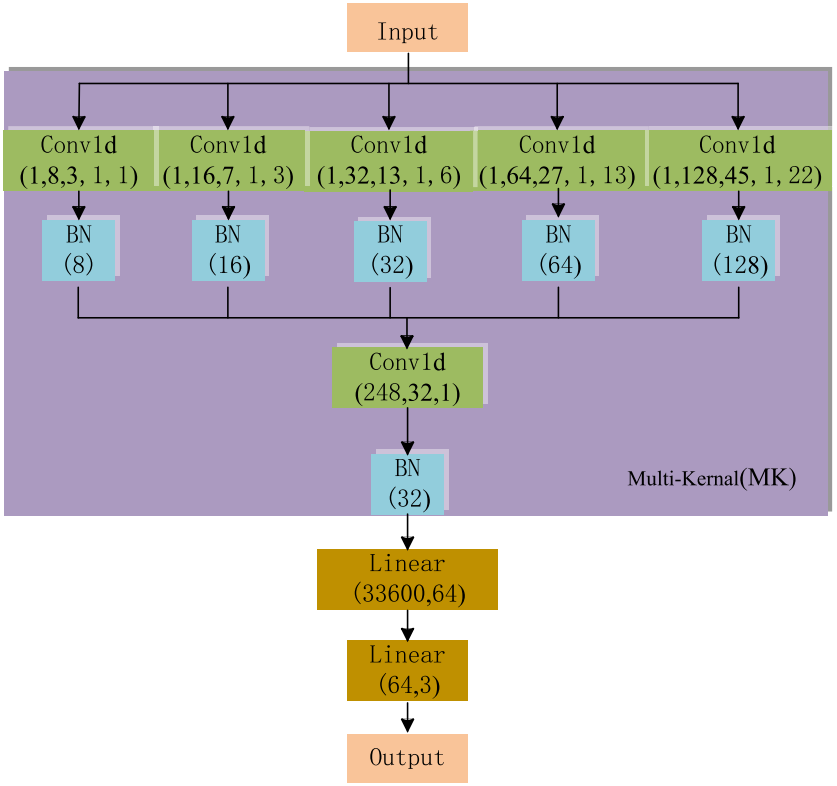


Fig. 3. MCNN framework diagram.

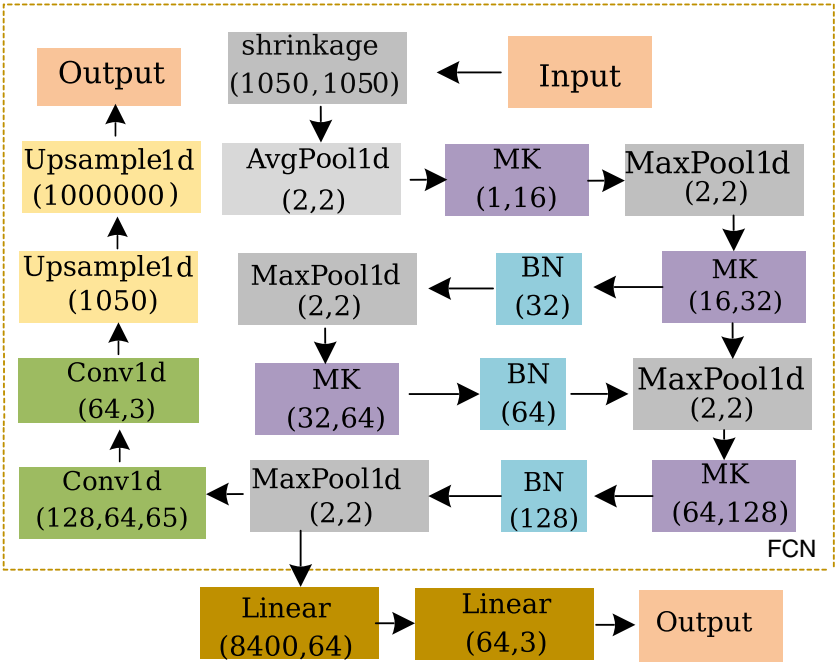


Fig. 4. FMCNN framework diagram.

3) ACC

ACC is the most commonly used performance measure in classification tasks, and it applies to both binary and multiclass tasks. The ACC rate is the proportion of correctly classified samples to the total number of samples, and its value represents the proportion of true results in the tested group, which is a combination of TP and TN. It measures the ACC of the classification model for classifying samples. A 99% ACC indicates that the test results are accurate in most cases, positive or negative. However, the formula for ACC means that even if SEN and SPE are high, it does not mean that the classification ACC is equally high. In addition to SEN and SPE, ACC is also determined by the proportion of such samples in the overall sample. Classification of sample classes that make up a small proportion of the overall sample may result in high SEN and SPE, but low ACC. ACC needs to be interpreted with caution. Its formula is as follows:

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

4) ROC and AUC

The samples are sorted according to the prediction results of the classification model, and the samples are predicted as positive examples one by one in this order, and the "False Positive Rate" (FPR) and "True Positive Rate" (TPR) are calculated each time. Take FPR as the abscissa and TPR as the ordinate to get the ROC curve. In classification problems, a threshold corresponds to an FPR and a TPR. When the threshold is changed, the corresponding FPR and FPT also change. When the threshold moves continuously, the ROC curve is obtained by plotting the corresponding continuously changing (FPR, TPR) in the graph. ROC describes the ability of the threshold to distinguish the subject, and also reflects the ability to distinguish positive examples and negative examples. A good classification model requires the ROC curve to be as close as possible to the upper left corner of the graph. According to this characteristic of ROC, the area under the ROC curve is calculated to obtain the area under the curve indicator. The formula of TPR is the same as that of SEN, and the formula of FPR is as follows:

$$FPR = \frac{FP}{TN + FP} \tag{5}$$

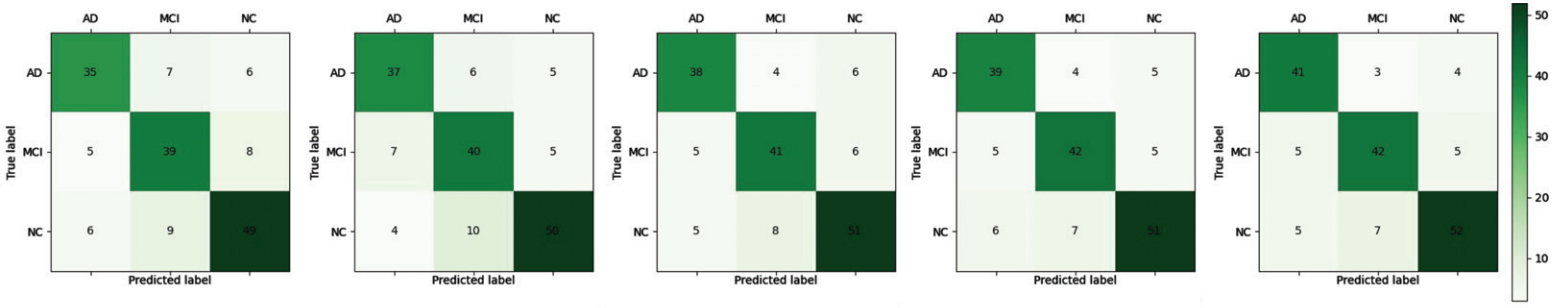
RESULTS

Our deep learning pipeline can predict AD status from DTI data by combining the segmentation concept in FCN with the multi-feature fusion concept in MCNN. Fiber tracking was performed on DTI of AD, MCI, and NC by deterministic fiber tracking after preprocessing of the DTI data. Then, the fibers in the region of interest are input into CNN, MCNN, and FMCNN for classification, and finally the corresponding high-risk fiber bundle probability map is output through FMCNN.

The first row in Fig. 5 shows the confusion matrices of UNC_L, UNC_R, UNC, CC, and CG+UNC classification models, respectively, and the second row of four graphs show confusion matrix for the CG, CC+UNC, CG+CC, CG+CC+UNC, and WM classification model, respectively. The horizontal axis of the confusion matrix is the predicted label, and the vertical axis is the actual label. Visualize the color of the values in the squares by different color. The closer to dark green the larger the sample size that falls in that area. The third row of Fig. 5 shows the ROC and P-R curves of UNC_L, UNC_R, UNC, CC, CG+UNC, CG, CC+UNC, CG+CC, CG+CC+UNC, and WM after model classification, respectively. The closer the curve is to the upper left corner, the more stable the performance of the model is. The abscissa of the P-R curve is the recall rate, and the ordinate is the precision. The curve reflects the comprehensive index of recall rate and precision. The results show that feeding the fibers of the WM into the CNN gives the best results.

The first row in Fig. 6 shows the confusion matrices of UNC_L, UNC_R, UNC, CC, and CG+UNC classification models, respectively, and the second row of four graphs show CG, CC+UNC, CG+CC, and CG+CC+UNC, respectively and confusion matrix for the WM classification model. Visualize the color of the values in the squares by color. The closer to dark green the larger the sample size that falls in that area. The third row of Fig. 6 shows the ROC and P-R curves of UNC_L, UNC_R, UNC, CC, CG+UNC, CG, CC+UNC, CG+CC, CG+CC+UNC, and WM after model classification, respectively. The results show that feeding the fibers of WM into MCNN gives the best results.

The graphs in the first row in Fig. 7 show the confusion matrices of the UNC_L, UNC_R, UNC, CC, and CG+UNC classification models, respectively, and the four graphs in the second row show CG, CC+UNC, CG+CC, and CG+CC+UNC, respec-



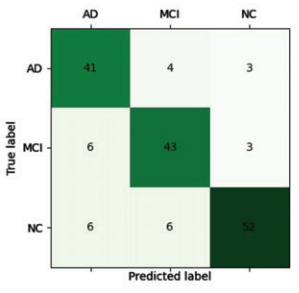
(a)confusion matrix of UNC_R

(b)confusion matrix of UNC_L

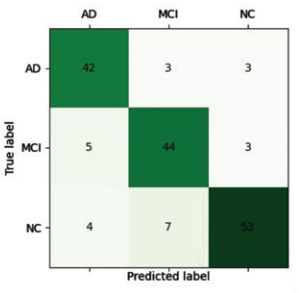
(c)confusion matrix of UNC

(d)confusion matrix of CC

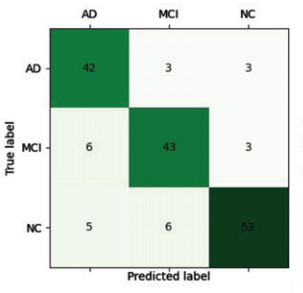
(e)confusion matrix of CG+UNC



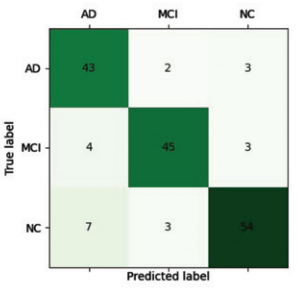
(f)confusion matrix of CG



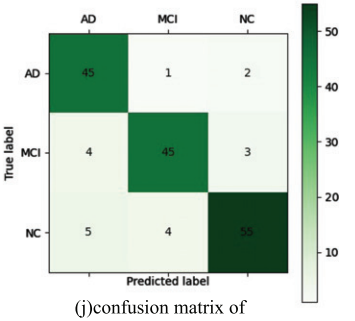
(g)confusion matrix of CC+UNC



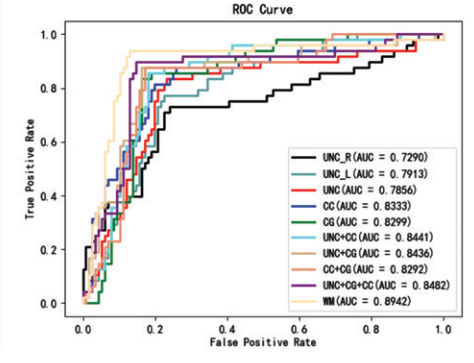
(h)confusion matrix of CC+CG



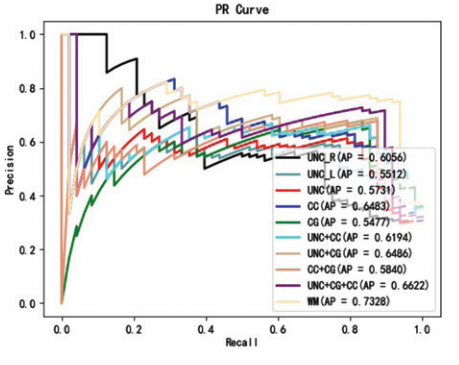
(i)confusion matrix of UNC+CC+CG



(j)confusion matrix of WM

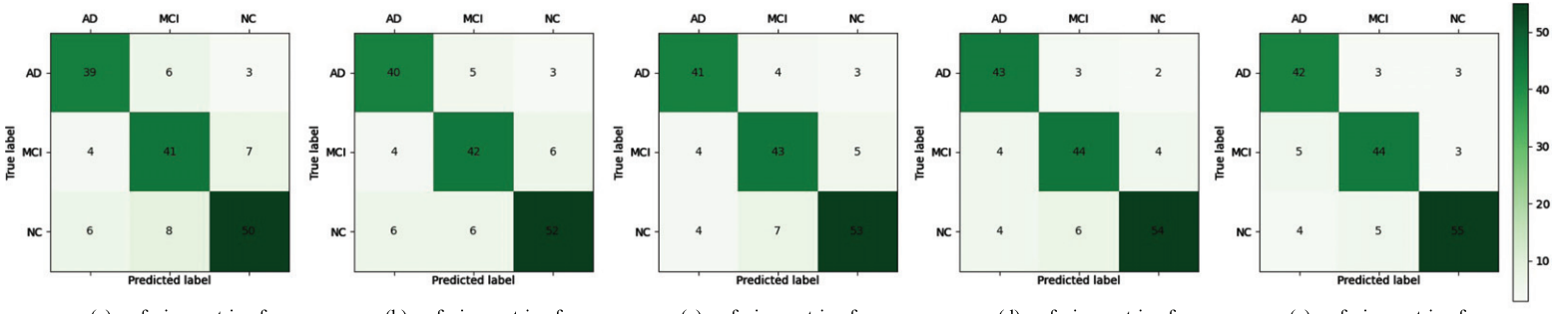


(k)ROC Curve



(l)P-R Curve

Fig. 5. Performance evaluation metrics for CNN models.



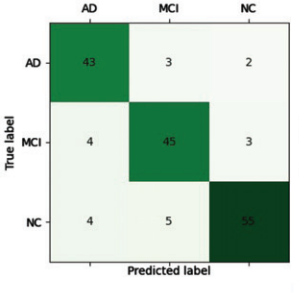
(a)confusion matrix of UNC_R

(b)confusion matrix of UNC_L

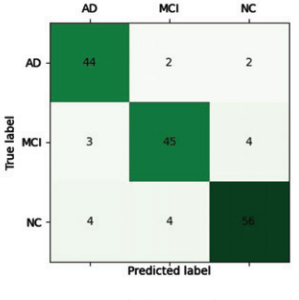
(c)confusion matrix of UNC

(d)confusion matrix of CC

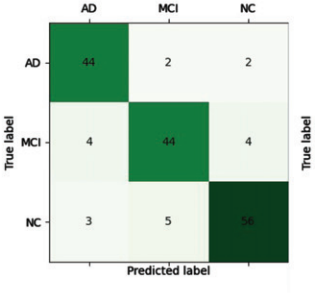
(e)confusion matrix of CG+UNC



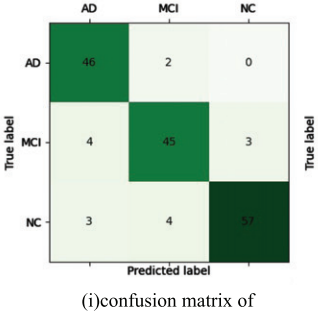
(f)confusion matrix of CG



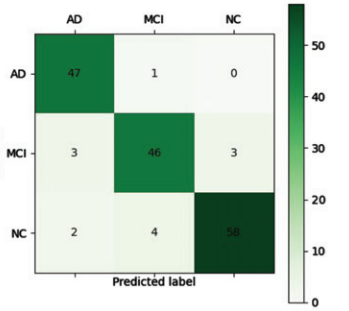
(g)confusion matrix of CC+UNC



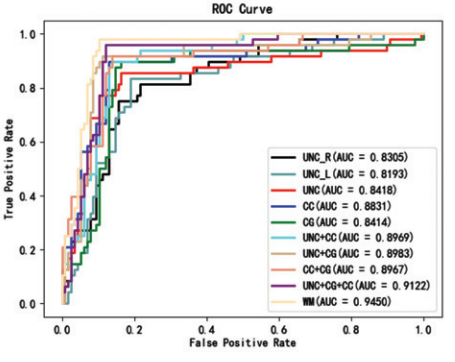
(h)confusion matrix of CC+CG



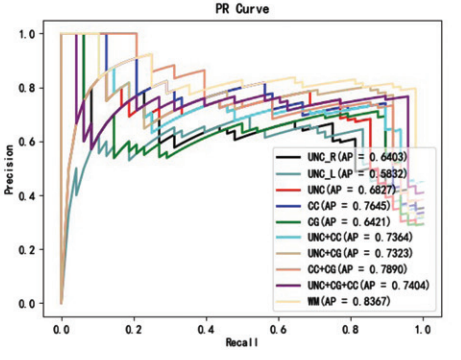
(i)confusion matrix of UNC+CC+CG



(j)confusion matrix of WM

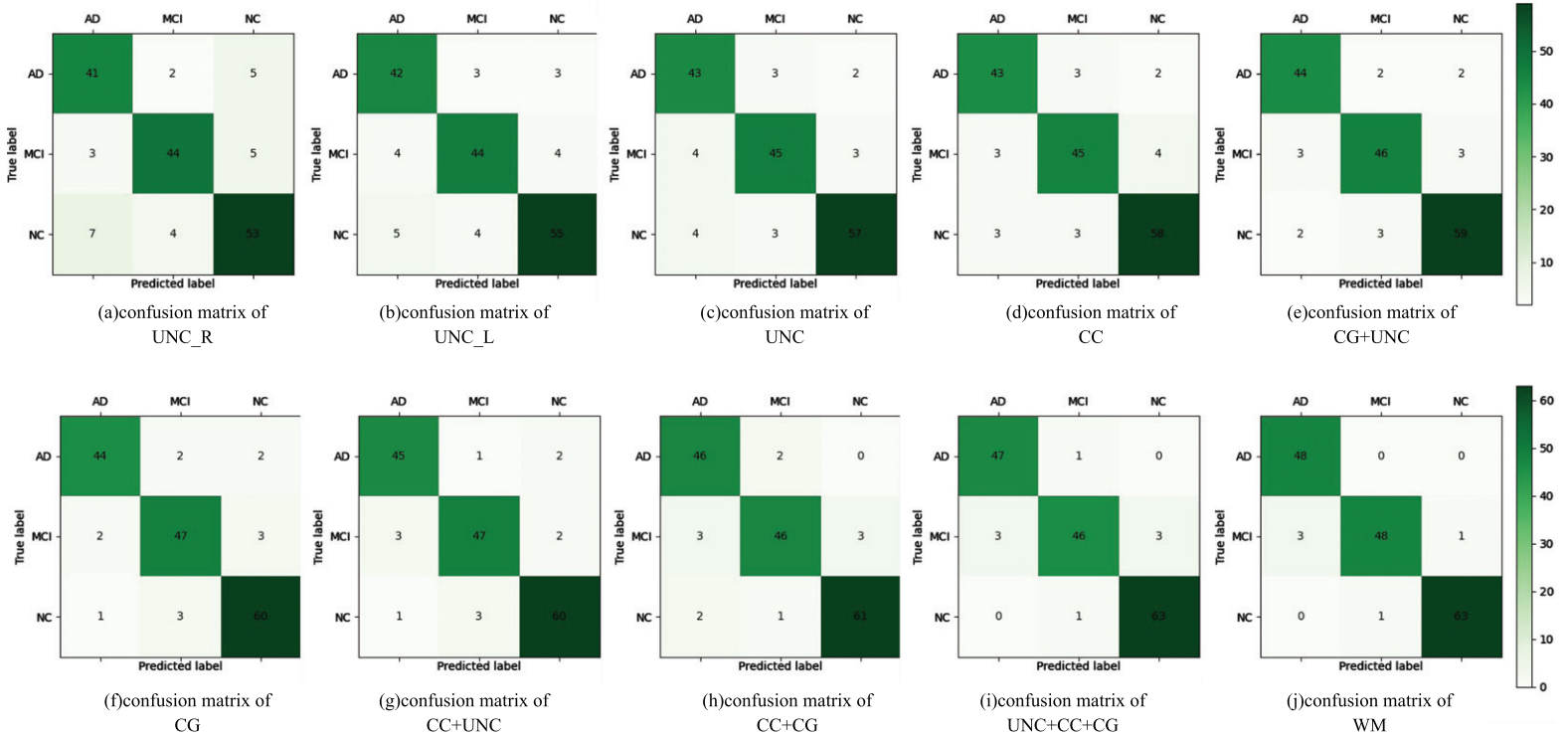


(k)ROC Curve



(l)P-R Curve

Fig. 6. Performance evaluation metrics for MCNN models.



(a)confusion matrix of UNC_R

(b)confusion matrix of UNC_L

(c)confusion matrix of UNC

(d)confusion matrix of CC

(e)confusion matrix of CG+UNC

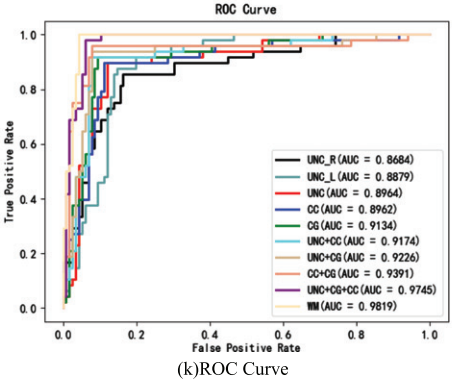
(f)confusion matrix of CG

(g)confusion matrix of CC+UNC

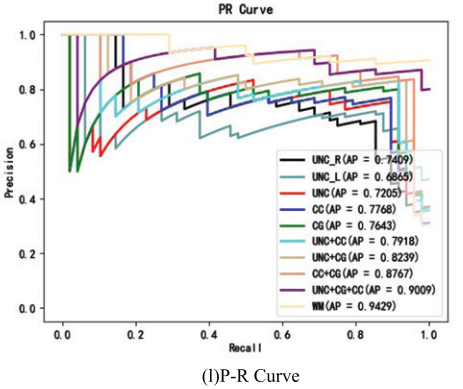
(h)confusion matrix of CC+CG

(i)confusion matrix of UNC+CC+CG

(j)confusion matrix of WM



(k)ROC Curve



(l)P-R Curve

Fig. 7. Performance evaluation metrics for FMCNN models.

Table 1
Model performance evaluation

ROI	CNN			MCNN			FMCNN		
	ACC%	SEN%	SPE%	ACC%	SEN%	SPE%	ACC%	SEN%	SPE%
UNC_R	75.00	90.52	72.92	79.27	91.38	81.25	84.15	91.38	85.42
UNC_L	77.44	90.52	77.08	81.71	91.38	83.33	85.98	92.24	87.50
UNC	79.27	91.38	79.17	83.54	93.10	85.42	88.41	93.10	89.58
CC	80.49	90.52	81.25	85.98	93.10	89.58	89.02	94.83	89.58
CG	82.32	91.38	85.42	85.98	92.24	87.50	90.85	95.69	91.67
UNC+CC	82.93	89.66	85.42	87.20	93.10	89.58	92.07	97.41	91.67
UNC+CG	84.76	92.24	87.50	88.41	93.97	91.67	92.68	96.55	93.75
CC+CG	84.15	90.52	87.50	87.80	93.97	91.67	93.29	95.69	95.83
UNC+CC+CG	86.59	90.52	89.58	90.24	93.97	95.83	95.12	97.41	97.92
WM	88.41	92.24	93.75	92.07	95.69	97.92	96.95	97.41	100.00

tively, and confusion matrix for the WM classification model. The horizontal axis of the confusion matrix is the predicted label, and the vertical axis is the actual label. Visualize the color of the values in the squares by color. The closer to dark green the larger the sample size that falls in that area. The third row of Fig. 7 shows the ROC and P-R curves of UNC_L, UNC_R, UNC, CC, CG+UNC, CG, CC+UNC, CG+CC, CG+CC+UNC, and WM after model classification, respectively. The results show that feeding the fibers of WM into FMCNN gives the best results.

The performance evaluation metrics in Table 1 confirm the consistent, high classification performance of deep learning models on external datasets. Among them, WM showed the best classification performance. It shows that in the classification of AD, NC, and MCI, for the three fiber bundles of interest, CC, CG and UNC, the fiber bundles in CG show obvious differences, while the fiber bundles in CC have little change in AD and NC. For the two ROIs of UNC_L and UNC_R, the ACC, SPE, and SEN of UNC_L were higher than those of UNC_R, indicating that the UNC on the left side of the brain contributed more to AD classification than the UNC on the right side of the brain, a conclusion that has been widely accepted. This is consistent with the conclusion that the left brain controls memory function. The experimental results show that, compared with CC and UNC, the fiber bundles in CG have a higher predictive ability for AD, which is consistent with the conclusion in previous studies [33] that CG is the most relevant WM brain region for cognitive memory. For the four combinations of CG+CC, CG+UNC, CG+CC+UNC, and CC+UNC, CG+CC+UNC showed higher accuracy, indicating that the integrated WM features of these three regions are useful in distinguishing AD from MCI and NC have a greater contribution. For all input features, WM showed the best classification results, indicating the importance of WM fibers in AD classification.

After demonstrating the effectiveness of FMCNN for AD and MCI diagnosis, we output the fiber probability maps. The four lines in Fig. 8 respectively represent the UNC quantitative anisotropy (QA) map, the QA map + the map of fiber bundles, the fiber bundle map and the grayscale display map of high-risk fiber bundles processed by FCMNN. The three columns of Fig. 8a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively; Fig. 8d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 8g-i show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. For the color representation of the last row in Fig. 8, its gray value ranges from 0 to 1. The closer the value is to 1, the higher the weight of the fiber in this category, the closer it is to 0, indicating that the fiber is more likely not to be classified into this category. From Fig. 8g-i, it can be clearly seen that the fibers on the left side of the UNC are closer to white than the right side, indicating that the right side contributes more to the classification of NC.

The four rows of Fig. 9 respectively represent the CC QA map, the QA map + the map of fiber bundles, the fiber bundle map, and the grayscale display map of high-risk fiber bundles processed by FCMM. The three columns in Fig. 9a-c represent the cross-section, coronal plane, and sagittal plane of the AD, respectively. Figure 9d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 9g-i show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. According to the last row of Fig. 9a-c, it can be found that the body of corpus callosum and the major forceps of corpus callosum are highly corre-

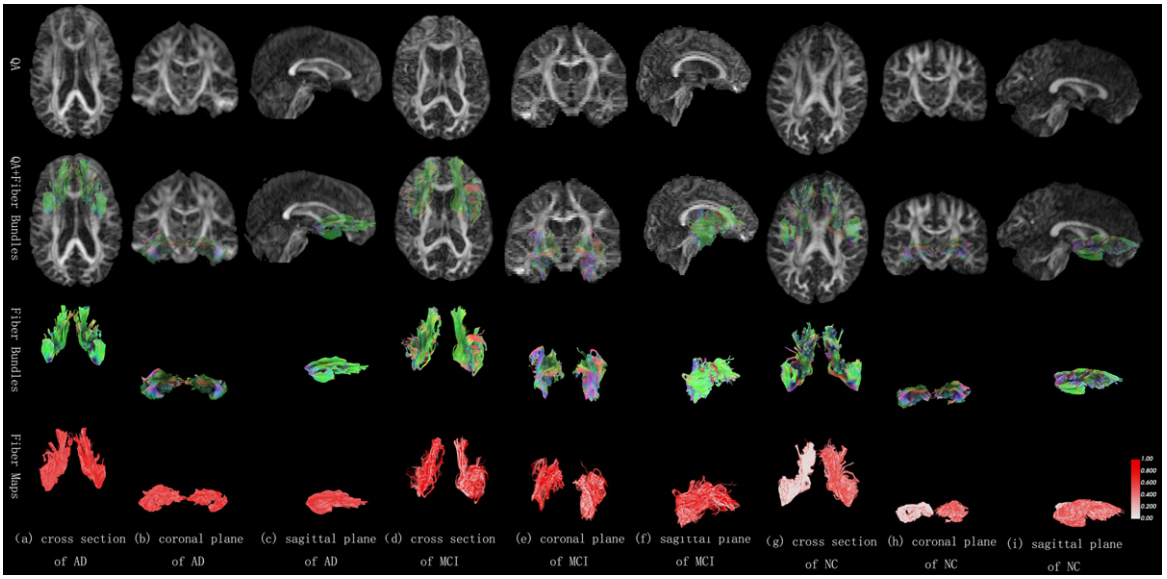


Fig. 8. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of UNC.

lated with AD positivity. According to the last row of Fig. 9d-f, it can be found that the body of corpus callosum and the tapetum of the corpus callosum are highly correlated with AD and MCI positivity.

The four lines of Fig. 10 respectively represent the CG+UNC QA map, the QA map and the fiber bundle overlay map, the fiber bundle map and the grayscale display map of the high-risk fiber bundles processed by FCMM. The three columns in Fig. 10a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 10d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 10g-i show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. From the probability maps in the last row of Fig. 10a-c, it can be seen that except for the right side of the UNC and the cingulate bundle fibers connecting the hippocampus and parietal lobe, other fibers are highly correlated with AD positivity. From the probability map in the last row of Fig. 10d-f, we can see that the fibers of the left cingulate tract connecting the frontal and parietal lobes, and the right cingulate tract connecting the parahippocampal gyrus and the parietal lobe fibers, fibers of the left cingulate tract connecting the parahippocampal gyrus and the parietal lobe were highly correlated with AD and MCI positivity.

The four rows of Fig. 11 respectively represent the CG QA map, the QA map + the map of fiber bundles, the fiber bundle map, and the grayscale display image of the high-risk fiber bundles processed by FCMM. The three columns in Fig. 11a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 11d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 11g-i, show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. From the comparison of the last two rows of Fig. 11, it is found that in the AD fiber bundle probability map, the left fiber bundle connecting the frontal and the parahippocampal gyrus, the left fiber bundle connecting the frontal and the parietal lobe, and the left fiber bundle connecting parahippocampal gyrus and the parietal lobe and the fiber tracts surrounding the paraolfactory area were significantly more whitish than the other fiber tracts, indicating that the fiber tracts in other parts contributed more to AD classification than the fiber tracts in these parts. In the NC fiber bundle probability map, the fiber bundles connecting the frontal and parietal lobes and the fiber bundles surrounding the parolfactory area are significantly redder than the fiber bundles in other areas, indicating that in addition to these two areas, other areas are contributing more to NC classification.

The four rows of Fig. 12 represent the QA map of CC+UNC, the QA map + the map of fiber bundles, the fiber bundle map, and the grayscale display map of high-risk fiber bundles processed by FCMM. The three columns in Fig. 12a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 12d-f show the cross-

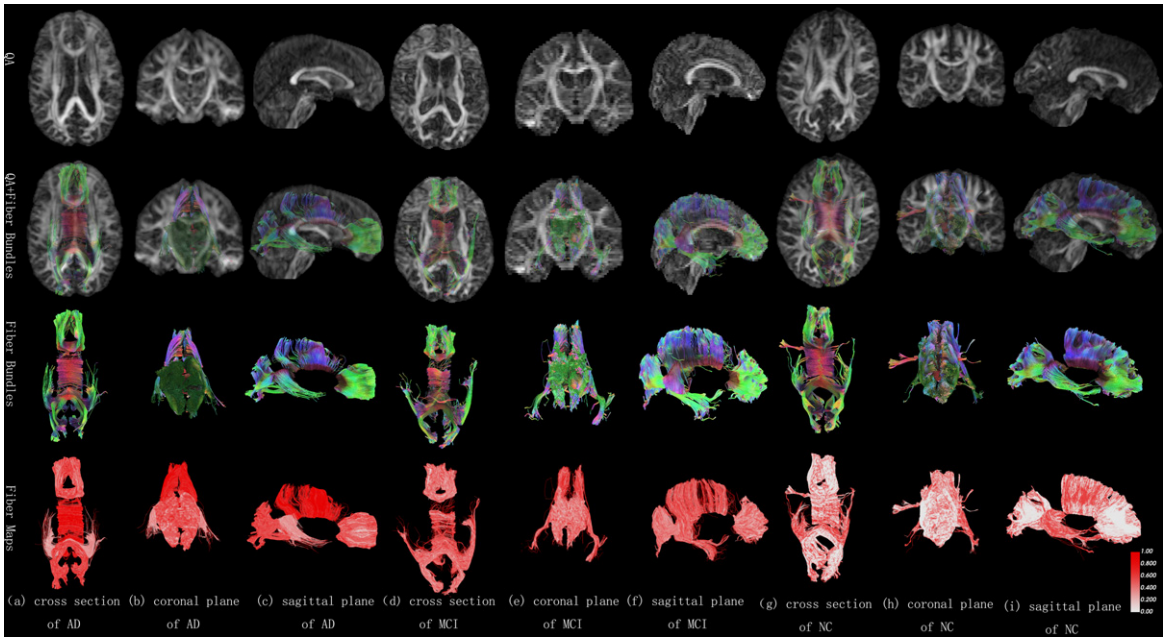


Fig. 9. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CC.

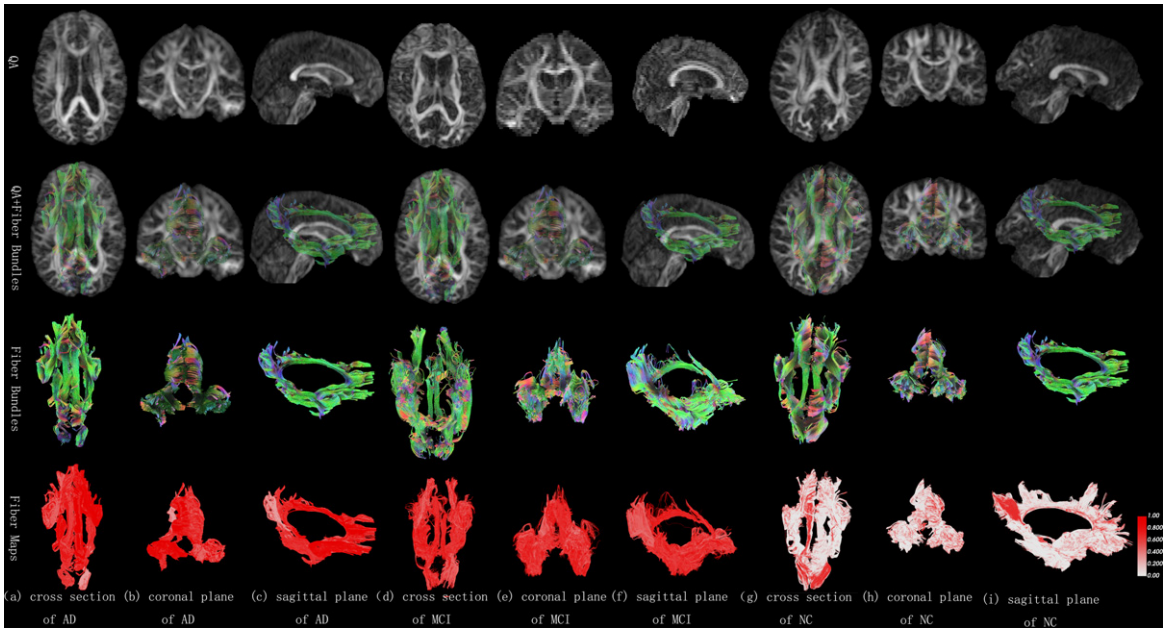


Fig. 10. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CG+UNC.

sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 12g-i, show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. From the probability maps in Fig. 12a-c, we found expect the tapetum of the corpus callosum and the fibers on the right side of the UNC, the fibers in other parts were highly correlated with AD positivity.

The four rows of Fig. 13 respectively represent the CG+CC QA map, the QA map + the fiber bundle map, the fiber bundle map, and the grayscale display map of the high-risk fiber bundles processed by

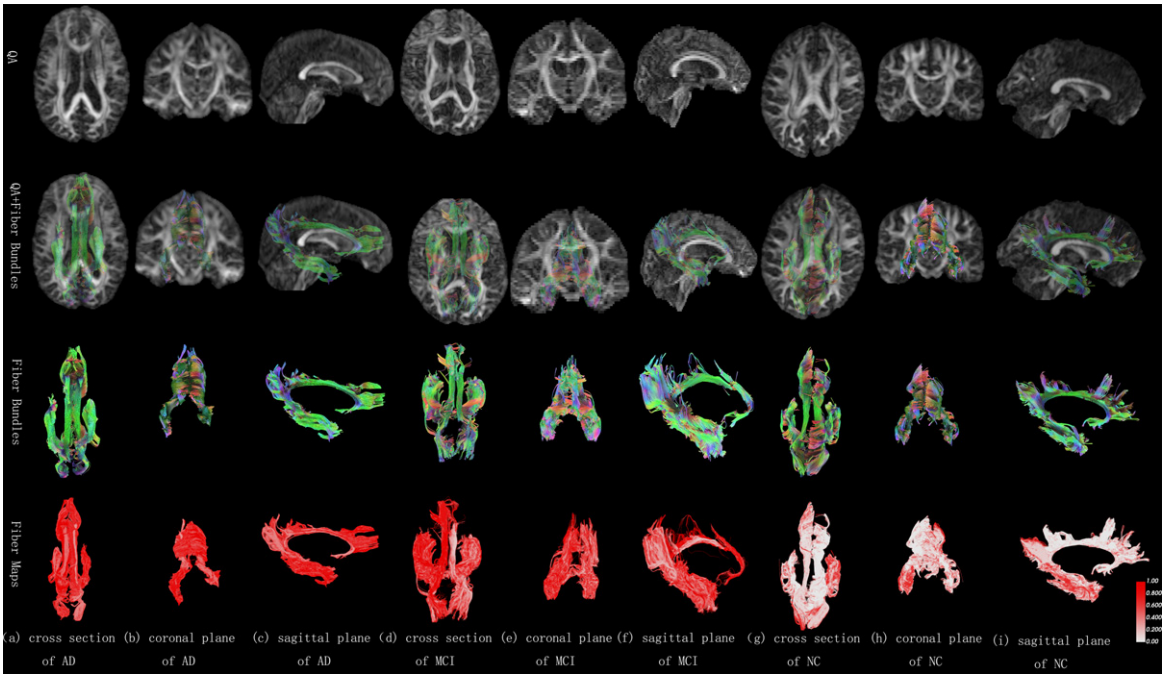


Fig. 11. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CG.

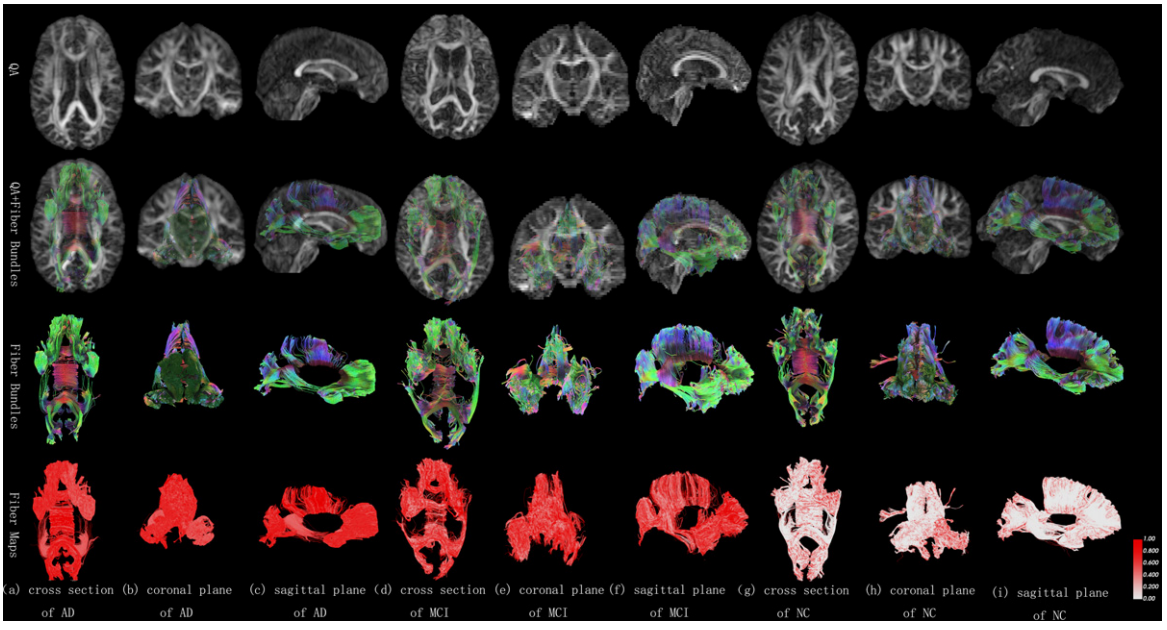


Fig. 12. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CC+UNC.

FCMM. The three columns in Fig. 13a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 13d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 13g-i show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. From the probability maps in Fig. 13a-c, it was found that fibers in other parts except the major forceps of corpus callosum were highly correlated with AD positivity. From the prob-

ability maps of Fig. 13g-i, it is found that the fibers of the body of CC and the CG surrounding the hippocampal are highly correlated with AD and MCI positivity.

The four rows of Fig. 14 respectively represent the QA map, QA map, and fiber bundle overlay map of CG+CC+UNC, fiber bundle map, and grayscale display map of high-risk fiber bundles processed by FCMM. The three columns in Fig. 14a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 14d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 14g-i, show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. By comparing the probability maps in Fig. 14a-f, it is found that the left side of the UNC and the CG connecting the hippocampus and the parietal were highly correlated with AD positivity. From the probability maps in Fig. 14a-c, it is found that other fiber bundles except the tapetum of the corpus callosum are highly correlated with AD positivity.

The four rows of Fig. 15 respectively represent the QA map, QA map + fiber bundle map of WM, fiber bundle map, and grayscale display map of high-risk fiber bundles processed by FCMM. The three columns in Fig. 15a-c represent the cross-section, coronal plane, and sagittal plane of the AD image, respectively. Figure 15d-f show the cross-sectional, coronal, and sagittal planes of the MCI images, respectively; and Fig. 15g-i, show the cross-sectional, coronal, and sagittal planes of the NC images, respectively. Combined with the probability map in the last row of Fig.15, it was found that the area of fiber tracts connecting the gray matter portion of WM was highly correlated with AD and MCI positivity.

DISCUSSION

Our method is compared with other methods in the literature, as shown in Table 2. Among them, Yida et al. collected and cleaned the data of the four hospital sites and applied it to the competition, and finally selected the top 5 from the competition for comparative analysis. The model used in the first place is XGBoot, which obtained 82.36% ACC, 86.36% SEN, 78.05% SPE, and 0.88 AUC. Ghaidaa et al. downloaded the tested DTI data from the ADNI database, first performed feature extraction on the data by SHIFT and SURFT, then performed feature selection on the data by BoW, and classified it by SVM, finally get 89% ACC for the classification of AD, MCI, and NC. Bahare et al. classified AD and NC by KSVM with 95.8% ACC, 95.8% SEN, 95.8% SPE, and 0.99 AUC. Ying et al. performed the classification to obtain an ACC of 91.92% and an AUC of 0.95. Li et al. extracted features from the data through VGG16, and then classified the features through SVM, obtaining an ACC of 94.2%, a SEN of 97.3%, a SPE of 92.9%, and the AUC of 0.95. Through our method, we obtained an ACC of 96.95%, a SEN of 97.41%, a SPE of 100% and an AUC of 0.98. The ACC, SEN, and SPE obtained by our method were better than those of methods in Table 2, higher by 1.15%, 1.61%, and 4.2%.

Advantages of this paper: 1) Considering that AD pathological features (senile plaques and filament winding) are abnormal protein deposits around neurons and synapses, we extracted neural information from nerve fibers for classification; 2) We used the coordinates of the fibers points as input features, which not only makes the input feature dimension low, but also retains the direction information of the fiber direction; 3) MCNN fuses abstract features of multiple scales, and FCN can locate the classification probability of each voxel, so this paper fused these two frameworks to achieve an accurate diagnosis of AD and MCI while capturing fiber-related disease status in brain ROIs.

This experiment also has certain limitations. We considered a case-control population in which three cognitively normal or diagnosed (AD and MCI) subpopulations were preselected. This situation does not fully represent the standard clinical decision-making process neurologists face. Patients typically present with a set of symptoms and the results of standard neurological tests that suggest a spectrum of neurodegenerative disorders rather than a binary situation. Therefore, our method cannot be directly applied to its current state, but rather serves as a first step towards establishing a more comprehensive framework to characterize the multiple etiologies of neurodegeneration. Secondly, the single dataset ADNI used in the experimental data is not decentralized, and it is impossible to know the impact of different datasets on the results.

Conclusion

The deep learning framework of this experiment integrates fully convolutional network and multi-kernel convolutional network to accurately diagnose AD while generating disease probability maps for

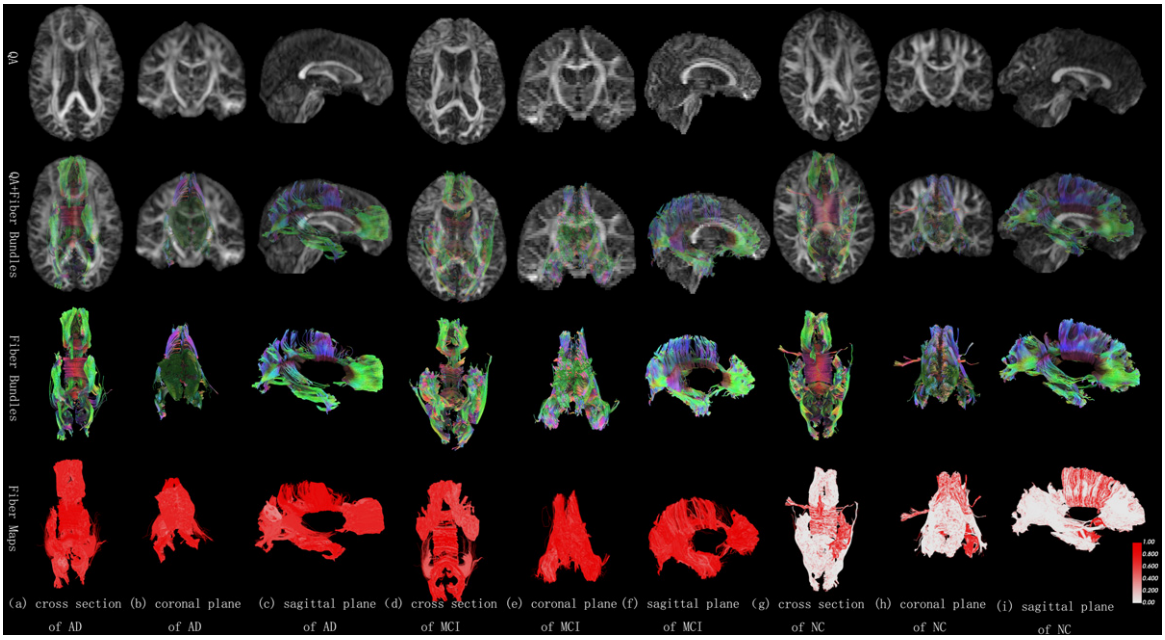


Fig. 13. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CG+CC.

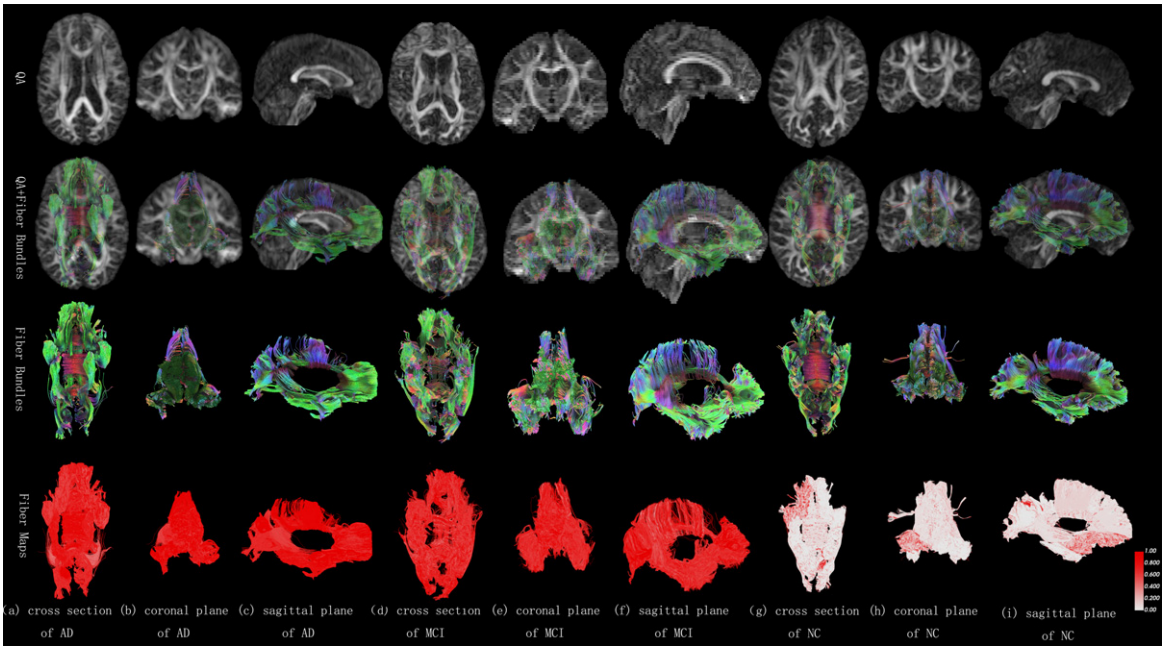


Fig. 14. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of CG+CC+UNC.

subjects with AD status and NC. Our model outputs probabilistic maps of fiber tracts of intuitive regions of interest, which can help in the direction of explainable artificial intelligence in medicine and deriving individualized phenotypes of occult diseases from traditional diagnostic tools. Disease probability maps provide a way to track areas of distinct brain fiber tracts associated with AD during diagnosis. Our model shows good predictive performance, yielding high and consistent values across all test datasets, according to the criteria of several different metrics. Thus, these findings demonstrate innovation in com-

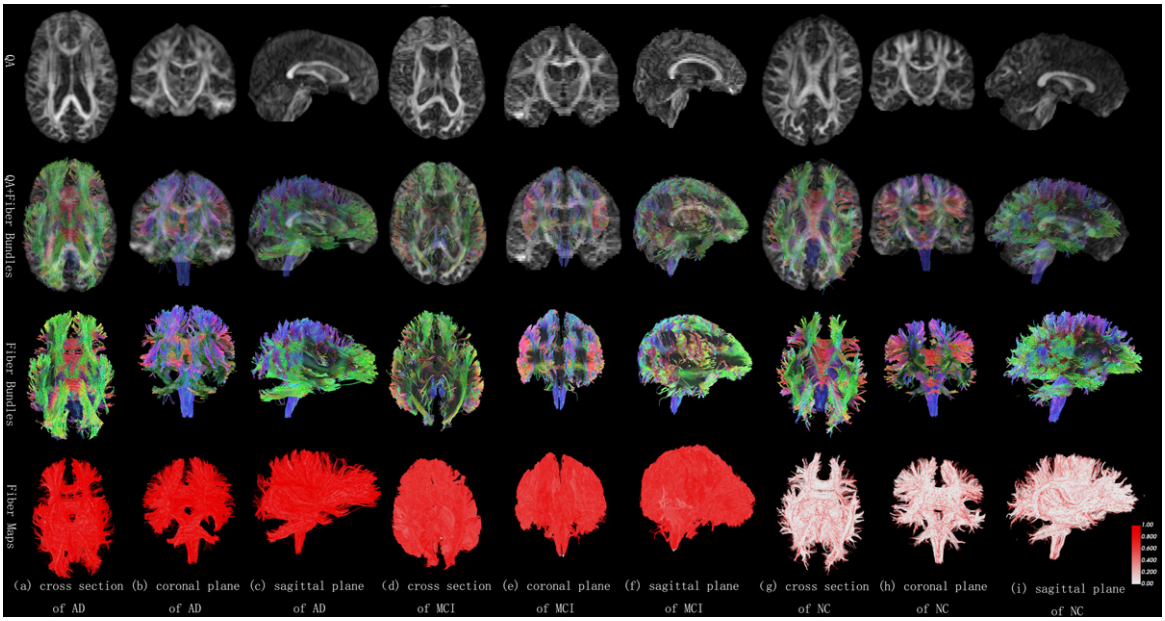


Fig. 15. QA, QA and fiber bundle overlay map, fiber bundle map, and fiber bundle probability map for AD, MCI, and NC of WM.

Table 2
Classification framework model comparison

Method	Data	AD/MCI/NC	Model	Modality	ACC%	SEN%	SPE%	AUC
Yida et al. [34]	4 hospitals	318/0/279	XGBoost	DTI	82.35	86.36	78.05	0.88
Ghaidaa et al. [35]	ADNI	35/30/31	BoW+SVM	DTI	89	–	–	–
Bahare et al. [36]	ADNI	24/0/24	KSVM	DTI	95.8	95.8	95.8	0.99
Ying et al. [37]	ADNI	48/0/51	SVM	DTI	91.92	–	–	0.95
Li et al. [38]	ADNI	0/70/50	CNN+SVM	DTI+sMRI	94.2	97.3	92.9	0.95
Our method	ADNI	121/130/162	FMCNN	DTI	96.95	97.41	100	0.98

bining medicine and computing, while providing new insights into the field of computer vision, while also expanding the range of biomedical applications of neural networks.

The disease probability map is created by applying the softmax function element-wise to the final activation array generated by the FCN. This step enables the transformation of an abstract tensor encoding of neuroanatomical information into a probabilistic array, demonstrating the likelihood that different fiber tracts in the region of interest contribute to AD, depending on the local geometry. In other words, the model refined the conceptualization of AD-suggested morphologies across the region of interest, and then used this learned information in test cases to assess the contribution of each fiber to the development of AD-related pathophysiological processes. probability. Representing these probabilities simply as a coherent grayscale map, it is thus possible to predict bar-by-bar where disease-related changes are likely to exist. Furthermore, we believe that the broader concept of disease process mapping by deep learning has the potential to be applied to many fields of medicine. Our work builds on these advances by requiring only one imaging modality to map a set of raw pixel values to a homogenous neural information-preserving disease probability map.

In conclusion, our deep learning framework is able to obtain high-accuracy AD classification features from DTI data. The framework also applies to other neurodegenerative diseases.

ACKNOWLEDGMENTS

The authors have no acknowledgments to report.

FUNDING

This study received support from the Natural Science Foundation of Shanghai (18ZR1426900).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (https://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

DATA AVAILABILITY

The data supporting the findings of this study are openly available in ADNI at https://adni.loni.usc.edu/data-samples/access-data/. These data were derived from the following resources available in the public domain: https://adni.loni.usc.edu/.

REFERENCES

[1] Scheltens P, Blennow K, Breteler MM, De Strooper B, Frisoni GB, Salloway S, Van der Flier WM (2016) Alzheimer's disease. *Postgrad Med J* **388**, 505-517.

[2] Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM (2010) The clinical use of structural MRI in Alzheimer disease. *Nat Rev Neurol* **6**, 67-77.

[3] Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Trojanowski J Q (2013) Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol* **12**, 207-216.

[4] Harper L, F Barkhof, Scheltens P, Schott JM, Fox NC (2014) An algorithmic approach to structural imaging in dementia. *J Neurol Neurosurg Psychiatry* **85**, 692-698.

[5] Nordberg A (2004) PET imaging of amyloid in Alzheimer's disease. *Lancet Neurol* **3**, 519–527.

[6] Bohnen NI, Djang D, Herholz K, Anzai Y, Minoshima S (2012) Effectiveness and safety of 18F-FDG PET in the evaluation of dementia: A review of the recent literature. *J Nucl Med* **53**, 59-71.

[7] Mattsson N, Insel P S, Donohue M, Jögi J, Ossenkoppele R, Olsson T, Schöll M, Smith R, Hansson O (2019) Predicting diagnosis and cognition with 81F-AV-1451 tau PET and structural MRI in Alzheimer's disease. *Alzheimers Dement* **15**, 570-580.

[8] Ossenkoppele R, Smith R, Ohlsson T, Strandberg O, Mattsson N, Insel PS, Palmqvist S, Hansson O (2019) Associations between tau, Aβ, and cortical thickness with cognition in Alzheimer disease. *Neurology* **92**, e601-e612.

[9] Mckhann G (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 263-269.

[10] Beach TG, Monsell SE, Phillips LE, Kukull W (2012) Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J Neuropathol Exp Neurol* **71**, 266-273.

[11] Whitwell JL, Dickson DW, Murray ME, Weigand SD, Nirubol Tosakulwong BS, Senjem ML, Knopman DS, Boeve BF, Parisi JE, Petersen RC, Jack CR, Josephs KA (2012) Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: A case-control study. *Lancet Neurol* **11**, 868-877.

[12] Van de Pol L A, Hensel A, Barkhof F, Gertz HJ, Scheltens P, Van Der Flier WM (2006) Hippocampal atrophy in Alzheimer disease: Age matters. *Neurology* **66**, 236-238.

[13] Barkhof F, Polvikoski TM, Van Straaten ECW, Kalaria RN, Sulkava R,Aronen HJ, Niinistö L, Rastas S, Oinas M, Scheltens P, Erkinjuntti T (2007) The significance of medial temporal lobe atrophy: A postmortem MRI study in the very old. *Neurology* **69**, 1521-1527.

[14] Raji CA, Lopez OL, Kuller LH, Carmichael OT, Becker JT (2009) Age, Alzheimer disease, and brain structure. *Neurology* **73**, 8199-1905.

[15] Hinton G (2018) Deep learning—a technology with the potential to transform health care. *JAMA* **320**, 1101-1102.

[16] Topol EJ (2019) High-performance medicine: The convergence of human and artificial intelligence. *Nat Med* **25**, 44-56.

[17] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* **521**, 436-444.

[18] Qiu S, Chang G H, Panagia M, Gopal, DM, Au R, Kolachalama, VB (2018) Fusion of deep learning models of MRI scans, Mini-Mental State Examination, and logical memory test enhances diagnosis of mild cognitive impairment. *Alzheimers Dement (Amst)* **10**, 737-749.

[19] Castelvecchi D (2016) Can we open the black box of AI?. *Nature News* **538**, 20.

[20] Lella E, Lombardi A, Amoroso N, Diacono D, Maggipinto T, Monaco M, Roberto Bellotti R, Tangaro T (2020) Machine learning and DWI brain communicability networks for Alzheimer's disease detection. *Appl Sci* **10**, 934.

[21] Giannakakis G, Trivizakis E, Tsiknakis M, Marias K (2019) A novel multi-kernel 1D convolutional neural network for stress recognition from ECG. *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pp. 1-4.

[22] Long J, Shelhamer E, and Darrell T (2015) Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431-3440.

[23] Petersen R C, Aisen P S, Beckett L A, Donohue MC, Gamst AC, Harvey DC, Jack CR, Jagust WJ, Shaw LM, Toga AW, Trojanowski JQ, Weiner MW (2010) Alzheimer's disease neuroimaging initiative (ADNI): Clinical characterization. *Neurology* **74**, 201-209.

[24] Conturo T E, Lori N F, Cull T S, Akbudak E, Snyder AZ, Shimony JS, McKinstry RC, Harold Burton H, Raichle ME (1999) Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A* **96**, 10422-10427.

[25] Mori S, Crain BJ, Chacko VP, Van Zijl PC (1999) Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* **45,** 265-269.

[26] Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MF, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?. *Neuroimage* **34**, 144-155.

[27] Køster R N, Jesper R, Bente P (2018) The total number of myelinated nerve fibers is reduced in corpus callosum in brains from patients with Alzheimer's disease. *Neurobiol Aging* **69**, 58-64.

[28] Park CH, Kim SH, Jung HY (2019) Characteristics of the uncinate fasciculus and cingulum in patients with mild cognitive impairment: Diffusion tensor tractography study. *Brain Sci* **9**, 377.

[29] Lo Buono V, Palmeri R, Corallo F, Allone C, Pria D, Bramanti P, Marino S (2020) Diffusion tensor imaging of white matter degeneration in early stage of Alzheimer's disease: A review. *Int J Neurosci* **130**, 243-250.

[30] Kamal S, Park I, Kim Y J, Kim, YJ, Lee U (2021) Alteration of the corpus callosum in patients with Alzheimer's disease: Deep learning-based assessment. *PLoS One* **16**, e0259051.

[31] Feng Q, Chen Y, Liao Z, Jiang H, Mao D, Wang M, Yu E, Ding Z (2018) Corpus callosum radiomics-based classification model in Alzheimer's disease: A case-control study. *Front Neurol* **9**, 618.

[32] Han C, Duan Y, Tao X, Tao X, Lu J (2019) Dense convolutional networks for semantic segmentation. *IEEE Access* **7**, 43369-43382.

[33] Fellgiebel A, Müller M J, Wille P, Dellani, PR, Scheurich A, Schmidt LG, Stoeter P (2005) Color-coded diffusion-tensor-imaging of posterior cingulate fiber tracts in mild cognitive impairment. *Neurobiol Aging* **26**, 1193-1198.

[34] Qu Y, Wang P, Liu B, Song C, Wang D, Yang H, Zhang Z, Chen P, Kang X, Du K, Yao H, Zhou B, Han T, Zuo N, Han Y, Lu J, Yu C, Zhang X, Jiang T, Zhou Y, Liu Y (2021) AI4AD: Artificial intelligence analysis for Alzheimer's disease classification based on a multisite DTI database. *Brain Disord* **1**, 100005.

[35] Eldeeb GW, Zayed N, Yassine IA (2018) Alzheimer's disease classification using bag-of-words based on visual pattern of diffusion anisotropy for DTI imaging. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE*, pp. 57-60.

[36] Bigham B, Zamanpour SA, Zare H, Initiative F (2022) Features of the superficial white matter as biomarkers for the detection of Alzheimer's disease and mild cognitive impairment: A diffusion tensor imaging study. *Heliyon* **8**, e08725.

[37] Zhang YT, Liu SQ (2018) Individual identification using multi-metric of DTI in Alzheimer's disease and mild cognitive impairment. *Chinese Physics B* **27**, 088702.

[38] Kang L, Jiang J, Huang J, Zhang T (2020) Identifying early mild cognitive impairment by multi-modality MRI-based deep learning. *Front Aging Neurosci* **12**, 206.