Magnetic resonance radiomics-based deep learning model for diagnosis of Alzheimer's disease

Zengbei Yuan[1,#] 📷 , Na Qi[1,#], Xing Chen[1,#], Yingying Luo[1], Zirong Zhou[1], Jie Wang[2], Junhao Wu[2], Jun Zhao[1] and for the Alzheimer's Disease Neuroimaging Initiative[*]

Abstract

Introduction: The progression of Alzheimer's disease (AD) has been shown to significantly correlate with changes in brain tissue structure and leads to cognitive decline and dementia. Using radiomic features derived from brain magnetic resonance imaging (MRI) scan, we can get the help of deep learning (DL) model for diagnosing AD.

Methods: This study proposes the use of the DL model under the framework of MR radiomics for AD diagnosis. Two cross-racial independent cohorts from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (141 AD, 166 Mild Cognitive Impairment (MCI), and 231 normal control (NC) subjects) and Huashan hospital (45 AD, 35 MCI, and 31 NC subjects) were enrolled. We first performed preprocessing of MRI using methods such as spatial normalization and denoizing filtering. Next, we conducted Statistical Parametric Mapping analysis based on a two-sample t-test to identify regions of interest and extracted radiomic features using Radiomics tools. Subsequently, feature selection was carried out using the Least Absolute Shrinkage and Selection Operator model. Finally, the selected radiomic features were used to implement the AD diagnosis task with the TabNet model.

Results: The model was quantitatively evaluated using the average values obtained from five-fold cross-validation. In the three-way classification task, the model achieved classification average area under the curve (AUC) of 0.8728 and average accuracy (ACC) of 0.7111 for AD versus MCI versus NC. For the binary classification task, the average AUC values were 0.8778, 0.8864, and 0.9506 for AD versus MCI, MCI versus NC, and AD versus NC, respectively, with average ACC of 0.8667, 0.8556, and 0.9222 for these comparisons.

Conclusions: The proposed model exhibited excellent performance in the AD diagnosis task, accurately distinguishing different stages of AD. This confirms the value of MR DL radiomic model for AD diagnosis.

[1]Department of Nuclear Medicine, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China
[2]Department of Nuclear Medicine & PET Center, Huashan Hospital, Fudan University, Shanghai, China

[#]Zengbei Yuan, Na Qi1, and Xing Chen contributed equally to this work.

[*]Some of the data utilized in this study were sourced from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The investigators involved in the ADNI contributed to the design and implementation of the initiative and/or provided the data used in this research. However, they did not participate in the analysis or the preparation of this manuscript. A comprehensive list of ADNI investigators can be accessed at: http://adni.loni.usc.edu.

Corresponding author:
Jun Zhao, Department of Nuclear Medicine, Shanghai East Hospital, School of Medicine, Tongji University, 150th Jimo Road, Shanghai 200120, China.
Email: petcenter@126.com

Introduction

Alzheimer's disease (AD) is a common neurodegenerative disorder characterized by chronic, progressive memory decline and cognitive impairment. It is one of the most serious diseases affecting the elderly population.[1] The progression of AD is gradual, making early-stage diagnosis particularly challenging due to the absence of specific diagnostic markers, which often leads to its symptoms being mistaken for those of natural aging. Early identification and accurate diagnosis of prodromal AD are essential for clinical decision making and the development of future treatments.

Structural magnetic resonance imaging (sMRI), as a noninvasive imaging technique, provides excellent characterization of brain tissue structure and anatomical details.[2,3] Three-dimensional T1-weighted imaging (3D T1WI) provides detailed brain morphological features, such as cortical thickness and hippocampal volume, which are essential for detecting early atrophy and other morphological changes associated with AD. 3D T1WI offers advantages such as high image resolution, short scanning time, and high signal-to-noise ratio. Additionally, 3D T1WI is well-suited for extracting tissue structure and performing brain structural analysis based on gray and white matter features. The cortical changes in specific brain regions, as revealed by 3D T1WI data, serve as important references for the conversion from Mild Cognitive Impairment (MCI) to AD.[4] Studies have shown that structural changes in areas such as the entorhinal cortex and hippocampus exhibit high sensitivity and specificity in predicting the progression of AD.[5] Significant atrophy in brain regions such as the parahippocampal gyrus, amygdala, and prefrontal cortex occurs as early as the MCI stage, and 3D T1WI data can be used to assess the extent of cortical changes in these regions, thereby evaluating disease progression.

Radiomics, an image analysis method that extracts more microscopic information from imaging data, has been proposed. By converting intuitive imaging data into analyzable, high-throughput features, these quantitative characteristics can describe the heterogeneity of related diseases.[6] Radiomics technology provides quantitative information including first-order, shape, and texture from images.[7–9] First-order features: These features were derived by analyzing the histogram of voxel intensities within the region of interest (ROI). Key statistical metrics included mean, standard deviation, skewness, and kurtosis, which describe the overall distribution of grayscale values.[10] Texture features: Texture features were computed using the gray-level co-occurrence matrix (GLCM) and the gray-level difference matrix (GLDM). The GLCM was constructed by calculating the frequency of voxel pairs with specific intensity values at a given spatial offset. From the GLCM, features such as contrast, homogeneity, energy, and entropy were derived to characterize image properties like roughness and smoothness. Similarly, the GLDM was used to compute features based on the absolute differences in intensity levels between neighboring voxels. Shape features: These features described the geometric properties of brain structures, including volume, surface area, and surface smoothness.[6] These metrics were particularly useful for capturing morphological changes associated with AD. The radiomics technology primarily involves designing automated feature extraction algorithms that extract microscopic information from standardized images for subsequent feature analysis. The radiomics framework typically includes the following steps: (1) data acquisition and preprocessing; (2) segmentation of ROI; (3) feature extraction from specific regions; (4) feature selection; and (5) model validation. By deeply mining MR radiomic information, the traditional interpretation of clinical imaging can shift from subjective evaluation to quantitative analysis, supporting prognosis assessment, multimodal information fusion, and correlation analysis.

In recent years, the development of artificial intelligence, particularly deep learning (DL), has provided new approaches for medical image analysis and the extraction of highly sensitive features to aid in the diagnosis of AD. Convolutional Neural Networks (CNNs) and other DL models have been successfully applied to MRI to identify subtle structural changes associated with AD.[11] These models can automatically learn complex, hierarchical features from raw image data, significantly improving in both efficiency and accuracy compared to traditional manual feature extraction methods.[12] Integrating radiomic features into DL frameworks further enhances the predictive power of these models. Chaddad et al.[13] constructed a CNN model to extract MR radiomic features from brain regions for AD versus normal control (NC) classification, achieving a classification area under the curve (AUC) of 92.58% and an accuracy (ACC) of 84.45%. Feng et al.[14] used the Least Absolute Shrinkage and Selection Operator (LASSO) model to select radiomic features from 3D T1WI images and then built a logistic regression model for the AD versus NC classification task, achieving an AUC of 0.72 and an ACC of 0.68. Chen et al.[15] used logistic regression to extract hippocampal positron emission tomography (PET) and MR radiomic features, demonstrating good diagnostic performance with an AUC of 86.00% and ACC of 80.60% in the MCI versus NC classification task. Leandrou et al.[16] constructed a trained and integrated XGBoost ensemble model using MR radiomic features from the entorhinal cortex and hippocampus for AD classification, achieving an F1 score of 0.95 for AD versus NC.

Radiomic features including various quantitative measures such as texture, shape, and intensity changes extracted from medical images, provide a comprehensive representation of potential pathology. However, radiomic features are often high-dimensional and heterogeneous, posing challenges for traditional machine learning (ML) models in

terms of feature selection and interpretability. TabNet, a DL architecture specifically designed for tabular data, leverages an attention mechanism to selectively focus on the most relevant features while ignoring noise or redundant information. This makes it particularly well-suited for handling radiomic features, which typically consist of numerous interrelated variables with varying degrees of importance. By automatically identifying and prioritizing key radiomic features, TabNet not only improves classification performance but also enhances model interpretability, enabling clinicians to better understand the underlying pathological mechanisms of AD. Park et al.[17] utilized the TabNet and XGBoost models in combination with radiomic features extracted from 3D T1WI data, achieving a classification AUC of up to 0.951 for AD vs NC. By combining the strengths of ML and radiomics, researchers have developed powerful models capable of distinguishing between NC, MCI, and diagnosed AD patients. Cheung et al.[18] extracted 107 radiomic features from brain MRI data obtained from two publicly available datasets and employed classical ML models for AD classification, achieving a maximum classification ACC of 76.57%. Yang et al.[19] focused on extracting radiomic features from the hippocampus and parahippocampal gyrus regions of brain MRI. Using ML models for AD versus NC classification, they achieved an AUC of 0.954. Feng et al.[20] extracted radiomic features from the left and right hippocampal regions of brain MRI and applied logistic regression for the classification of amnestic MCI (aMCI), achieving a maximum AUC of 0.79. Du et al.[21] developed and validated radiomic classification models for late-onset AA (LOAD) versus young controls (YC), LOAD versus old controls (OC), and early-onset Alzheimer's disease (EOAD) versus LOAD. The models achieved AUCs of 0.91, 0.78, and 0.86, respectively, in the validation set. Liu et al.[22] selected 80 cases each of AD and NC, extracting radiomic features from whole-brain 3D T1WI, and achieved a maximum classification ACC of 0.962. Zhou et al.[23] employed the U2-Net model for hippocampal segmentation and radiomic feature extraction, achieving a classification AUC of 0.97 using ML models for AD versus NC. Feng et al.[24] extracted 3360 radiomic features from amygdala MRI and combined them with logistic regression, achieving a maximum classification AUC of 0.93 for AD versus NC. Peng et al.[25] selected 341 patients with MCI and divided them into two groups based on whether they progressed to AD within 8 years. Using radiomic features extracted from brain PET images and an Support Vector Machine (SVM) model, they achieved a maximum classification AUC of 0.865. These models not only enhance the accuracy of AD diagnosis but also show potential for detecting changes in imaging features as the disease progresses, which is crucial for timely intervention and management. Using DL models to automate the diagnostic process also enhances the standardization and reproducibility of results, reducing dependence on subjective interpretation and providing clinicians with a more consistent diagnostic tool.

This study aimed to develop a MR-based DL radiomics model for the classification of AD by integrating multiscale radiomic feature extraction with a TabNet architecture. Specifically, 3D-T1WI data were preprocessed to identify brain regions exhibiting significant differences across AD stages. Radiomic features were subsequently extracted, optimized through rigorous feature selection, and fed into the TabNet model, which is renowned for its interpretability and efficiency in handling high-dimensional data. The novelty of our approach lies in the seamless combination of radiomic feature extraction, advanced feature selection techniques, and the utilization of TabNet, which enhances both predictive performance and interpretability in AD diagnosis. Our findings demonstrate that this integrated framework achieves superior performance across various classification tasks, distinguishing between different AD stages with high accuracy. This work is expected to contribute significantly to the early detection of AD, providing clinicians with a robust tool to support timely and informed decision making in clinical practice.

Material and methods

Enrolled participants

A total of 649 subjects were included in this study: 538 subjects sourced from the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) database (ADNI-2 and ADNI-GO)[26] (https://adni.loni.usc.edu/) and 111 subjects from the Chinese Preclinical Alzheimer's Disease Study (C-PAS) database.[27] The demographic information of the included participants is provided in Table 1. The study was conducted in accordance with the Declaration of Helsinki. These sample numbers from ADNI were determined through random sampling, aiming to achieve a balanced representation of each group while maintaining a sufficiently large sample size for robust statistical analysis. The process was approved by the Institutional Review Boards (IRBs) at each participating site, and all participants provided written informed consent. The study of C-PAS was approved by the Institutional Ethics Review Board of Huashan Hospital (ethical code number: HS-KY-2017-406) and the clinical trial number was not applicable.

ADNI subjects were scanned using 3T scanners manufactured by Philips, Siemens, and GE, all acquiring T1WI images of the head. The protocol parameters were listed as follows: 1.2 mm slice thickness; 256×256 scanning matrix; repetition time = 2300 ms; echo time = 2.98 ms; field of view = 240×240 mm^2; flip angle = 90 degree; and 256×256 reconstruction matrix.

Subjects from ADNI were selected based on specific inclusion and exclusion criteria established by the ADNI

Table 1. The demographic information of enrolled subjects (mean ± SD).

	Groups	Age	Gender (M/F)	Total (N)	MMSE
ADNI	NC	75.85 ± 6.39	112/119	231	28.96 ± 1.78
	MCI	75.40 ± 8.27	94/72	166	26.74 ± 3.41
	AD	74.42 ± 8.35	77/64	141	23.06 ± 2.04
C-PAS	NC	67.42 ± 8.94	8/23	31	28.79 ± 1.15
	MCI	72.40 ± 8.60	18/17	35	25.44 ± 2.19
	AD	70.33 ± 7.95	11/34	45	17.89 ± 5.07

AD: Alzheimer's disease; ADNI: Alzheimer's Disease Neuroimaging Initiative; C-PAS: Chinese Preclinical Alzheimer's Disease Study; F: female; M: male; MCI: Mild Cognitive Impairment; MMSE: Mini-Mental State Examination; N: number of subjects in each class; NC: normal control.

consortium, including age range, cognitive status, and MRI availability, while excluding major neurological or psychiatric disorders and MRI contraindications.[28] The C-PAS database and images were acquired using a 3T uPMR790 TOF scanner (United Imaging, China). C-PAS is an observational longitudinal study conducted in Shanghai, China, commencing in April 2019. Participants were clinically diagnosed by experienced neurologists using the 2011 National Institute on Aging and Alzheimer's Association (NIAAA) diagnostic criteria for suspected AD. In addition, MCI was diagnosed if the participant met one of the following criteria: (1) at least one impaired cognitive domain, which implies impaired scores (>1 standard deviation (SD) below the age-corrected normative mean) on all neuropsychological tests in the same domain; (2) impaired scores (>1 SD) in each of the three cognitive domains. The participants for NC were all recruited from communities in Shanghai. Participants with NC were identified based on our previous study; specifically, those who did not meet the criteria for AD or MCI were identified as individuals with NC. During image acquisition, subjects positioned supine with cotton balls inserted in both ears to reduce noise, and sponge pads placed between the head and the coil to limit head movement. Subjects were instructed to remain quiet and keep their eyes closed during imaging. The imaging parameters were: 1.0 mm slice thickness; 256×256 scan matrix; repetition time = 2300 ms; echo time = 3.00 ms; field of view = 230×230 mm^2; flip angle = 10 degree; and 256×256 reconstruction matrix.

Methods

This study applied an MR radiomics-based DL model approach to perform AD diagnosis tasks, with the experimental workflow outlined in Figure 1. (1) Sample distribution: subjects were selected from the ADNI and C-PAS databases to form the experimental group; (2) Image preprocessing: raw 3D T1WI were preprocessed, including

steps such as skull stripping and image registration, to standardize all 3D T1WI into uniform specifications; (3) ROI determination and feature extraction: ROI were identified using two-sample t-tests, and radiomic features were extracted with Radiomics tools; (4) Feature selection: specific radiomic features were selected using the LASSO model; and (5) Model performance validation: based on the selected radiomic features, the TabNet model was trained and optimized. Model training parameters were adjusted to achieve optimal classification performance for diagnosing AD.

Image preprocessing. The purpose of preprocessing is to eliminate image distortions caused by random errors, such as low-frequency background noise. In this study, preprocessing of 3D T1WI were performed using the CAT12 tool based on SPM12 (https://www.nitrc.org/projects/cat/) on the MATLAB R2020b platform. First, the DCM2NII tool in MRIcon was used to convert the original DICOM images of each sample into a single NIfTI file. For each sample, the preprocessing process began with initial steps, including denoising filtering using resampling methods to adapt to low-resolution images and intensity nonuniformity correction to remove brightness and contrast variations. Subsequently, the Adaptive Probability Region-Growing (APRG) method was employed to remove nonbrain tissue and retain intracranial structures. Then, the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) registration method was used to align all images to the MNI152 space. This approach ensures precise alignment of anatomical structures while minimizing distortions and preserving morphological information.[29] Finally, the intensity normalization was applied using a bias field correction approach implemented in CAT12 to standardize voxel intensity distributions and minimize scanner-related variability. The preprocessed images had voxel dimensions of $181 \times 217 \times 181$, with each voxel measuring $1 \times 1 \times 1$ mm^3.
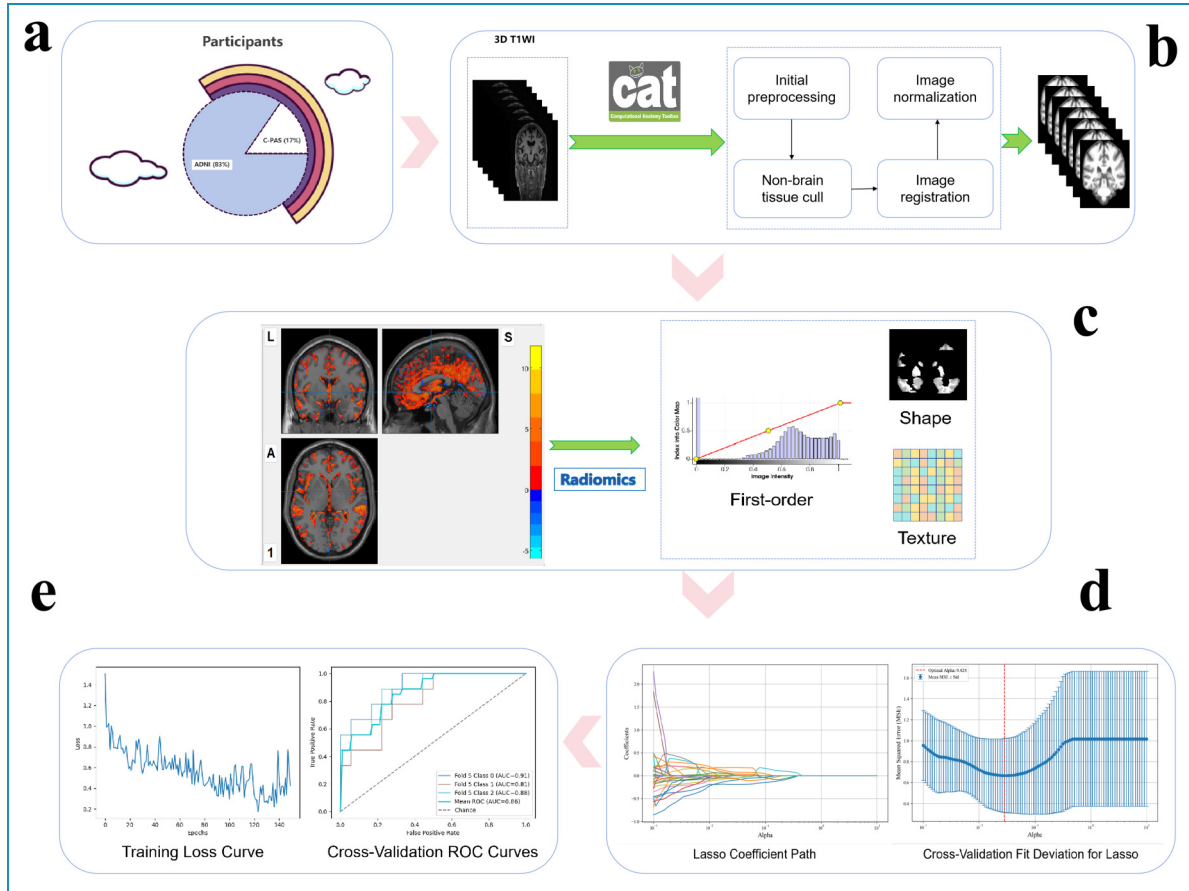
Figure 1. (a) Sample distribution; (b) image preprocessing; (c) ROI determination and feature extraction; (d) feature selection; and (e) model performance validation. ROI: regions of interest.

Definition of ROI. In this study, we focused on localizing brain regions that are strongly associated with structural morphological changes induced by AD. To characterize the morphological differences in brain tissue between AD, MCI, and NC, a group comparison was performed based on 141 AD, 166 MCI and 231 NC subjects randomly selected from the ADNI cohort. These samples were omitted from the subsequent feature selection and classification analyses and were not utilized in any of the model training, validation, or testing procedures. A two-sample t-test in SPM12 was used for intergroup comparison, with a peak threshold set at $p < 0.01$ and Family-Wise Error (FWE) correction applied across the entire brain region, with a voxel threshold of 20. The results were localized and visualized using the DPABI tool (http://rfmri.org/DPABI),[30] as shown in Figure 2. The progression of AD is accompanied by significant morphological changes in certain brain regions. Based on the results of the two-sample t-test, we selected brain regions that have been confirmed to undergo morphological changes during the course of AD, including the hippocampus, amygdala, medial temporal lobe, precuneus, posterior cingulate gyrus, parietal lobe, frontal lobe, basal ganglia, and thalamus. These regions will be treated as ROI in subsequent analyses.[31–35]

Radiomic features extraction and selection. Radiomic features were extracted from the C-PAS dataset based on ROIs identified using the ADNI cohort. In this study, radiomic features from MRI were extracted using the Python-based Radiomics tool (https://github.com/mvallieres/radiomics).[36] A total of 107 MR radiomic features were extracted, including 14 shape features, 18 first-order features, 24 texture features from the GLCM, 14 texture features from the GLDM, 16 texture features from the Gray-Level Run-Length Matrix (GLRLM), 16 texture features from the Gray-Level Size Zone Matrix (GLSZM), and 5 texture features from the Neighborhood Gray-Tone Difference Matrix (NGTDM). The detailed mathematical definitions of the radiomics matrices have been previously reported.[37] In order to select the most relevant features for AD diagnosis from high-dimensional radiomic features, we employed the LASSO regression method which was employed to select the most relevant radiomic features for distinguishing AD patients from non-AD individuals
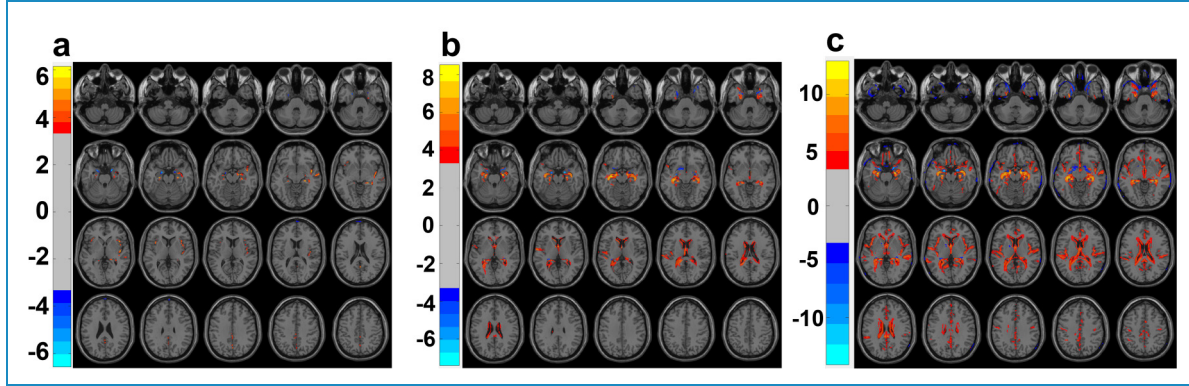
Figure 2. Results of the two-sample t-test between AD, MCI, NC, (a) AD versus MCI; (b) MCI versus NC; (c) AD versus NC. The color bars represent t-value, the higher the absolute t-value, indicating a more significant difference. AD: Alzheimer's disease; MCI: Mild Cognitive Impairment; NC: normal control.
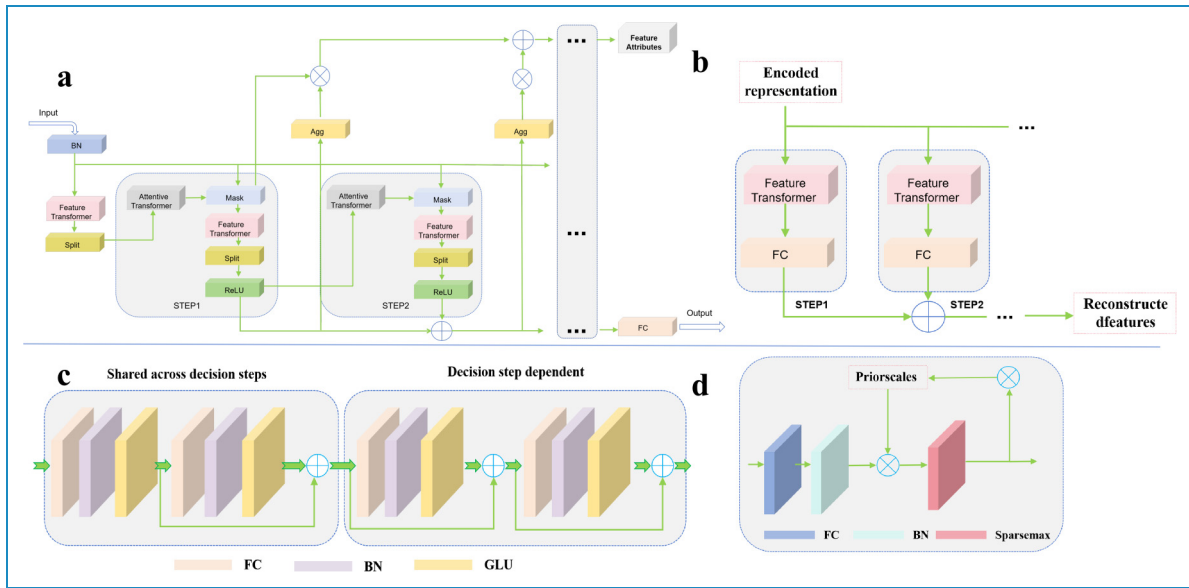

Figure 3. The architecture of TabNet. (a) encoder, (b) decoder, (c) feature transformer, and (d) attentive transformer. BN: batch normalization; FC: fully connected; GLU: gated linear unit.

(NC and MCI combined).By introducing the L1 regularization term, LASSO is able to shrink the regression coefficients of less important features to zero, thereby achieving feature selection and dimensionality reduction. Nine nonzero coefficient features were selected in total, including 2 shape features, 3 first-order features, 1 GLDM texture feature, and 3 GLSZM texture features.

Model construction of TabNet. The TabNet model selected in this study is a neural network method proposed by the Google team for processing tabular data. It was originally designed to create interpretable models for tabular data using a DL attention network.[38] The main advantage of TabNet lies in its combination of end-to-end neural network learning and representation learning, along with the interpretability and sparse feature selection characteristics similar to those of tree models. TabNet achieves this by constructing a neural network with a decision manifold similar to that of tree models, thereby incorporating the benefits of decision trees. Each layer of TabNet corresponds to a step in the decision tree: Mask Layer: A sparse attention mechanism that selects a subset of input features at each decision step; the fully connected (FC) layer with Rectified Linear Unit (ReLU) corresponds to threshold judgment; Aggregation Layer: Summed outputs from all decision steps, followed by a Softmax layer for classification. The input to TabNet consists of feature dimensions $B \times D$, with a batch size B and feature dimensions D as

inputs, and the model output is a single vector or number, facilitating classification or regression tasks.

The main architecture of TabNet consists of two primary components: the encoder and the decoder.[39] The TabNet encoder (as shown in Figure 3(a)) consists of multiple feature processing modules (Feature Transformer), attention modules (Attentive Transformer), and feature masks.[40] The input to the encoder consists of preprocessed numerical features, and the output includes the encoded features and the data used for the final decision. The STEP in the figure represents a decision step, where each step receives all input features and selects a subset of features for computation. Each STEP receives data features as input and applies weights to these features based on the output from the previous step. The output from each step is accumulated and used for the final decision. The TabNet decoder is responsible for reconstructing the encoded features to the original data table features. Each step in the decoder consists of a Feature Transformer module and a FC layer (as shown in Figure 3(b)). By passing the masked portion of the data through the encoder and then decoding it, the final decoded data is multiplied to obtain the predicted values. The model parameters are adjusted based on the error between the actual values of the masked portion and the predicted values.

The structure of the Feature Transformer module is shown in Figure 3(c). It consists of FC layer, batch normalization (BN) layer, and a gated linear unit (GLU) layer. The module employs skip connections to enhance the effectiveness of training and scales the gradients during backpropagation to prevent excessively large gradients, which could increase the variance of the weights and compromise training stability.[41,42] The Feature Transformer module can be divided into shared and independent GLUs. The shared GLUs maintain parameter sharing during training, while the independent GLUs are trained separately. This structure allows the model to use the same layers to process common features in the early stages of training, and then choose different layers to capture independent features. The Attentive Transformer module (as shown in Figure 3(d)) consists of an FC layer, BN layer and a function.[43] Its purpose is to determine which features should be prioritized at each decision step by encoding features, generating attention weights, performing feature selection, and making recursive decisions. This dynamic adjustment of feature selection enhances the model's interpretability.

Experimental

Experimental setup

The parameter settings for the model training process were detailed as follows: Model training was performed using the selected MR radiomic features. In 3-way classification: Epochs: 150, Batch Size: 32, Optimizer: Adam,[44] Loss Function: Cross-Entropy Loss, Learning Rate Schedule: 2e-2; In binary classification : Epochs: 250, Batch Size: 32, Optimizer: Adam, Loss Function: Cross-Entropy Loss, Learning Rate Schedule: 2e-2.

The experiments were conducted on a Windows 11 operating system, using Python 3.9, PyTorch 2.0, and CUDA 11.8 environments. The hardware configuration was as follows: CPU: Intel(R) Core(TM) i7-14700KF; GPU: NVIDIA GeForce RTX 4070 Ti.

Results

This study extracts radiomic features from 3D T1WI and performs AD diagnosis using the TabNet model. The evaluation metrics selected for this study include AUC and ACC, the AUC value using the One-vs-Rest (OvR) strategy. In addition to the classification AUC, the ACC is also known to be beneficial to diagnosis. All results were averaged over five-fold cross-validation. Additionally, we conducted two distinct sets of experiments to evaluate both the performance of the proposed model and the sensitivity of different radiological features in diagnosing AD. Specifically, we compared the performance of various ML models and examined the impact of different combinations of radiomic features. Both comparative and main experimental (TabNet) groups maintained consistent training environments, samples, and model configurations and using the average results from the five-fold cross-validation.

To further verify the sensitivity of different radiomic features for AD diagnosis, we combined various radiomic features and designed control groups under different feature combinations. The results of 3-way classification (AD vs. MCI vs. NC), as shown in Tables 2 and 3 and Figure 4, indicate that the combined radiomic feature set (first-order, shape, and texture) achieved the highest average AUC (0.8728) and average ACC (0.7111). Moreover, to effectively address the potential impact of dataset imbalance on result evaluation, we have introduced the F1-score as an additional evaluation metric. In the 3-way classification (AD vs. MCI vs. NC), the TabNet model obtained a mean F1 score of 0.7111 under the same five-fold cross-validation setting. Among the individual radiomic feature control groups, the first-order features demonstrated the highest average AUC (0.7918) and average ACC (0.5778). The classification performance of the two combined radiomic feature pairs was superior to that of individual features. Given the interpretability of the TabNet model, the feature weights of different radiomic features were derived from the model, as illustrated in Figure 5. Among all radiomic features, the first-order features were assigned the highest weights.

To select the baseline models, we referred to previous studies that have examined the efficacy of these models in the context of AD diagnosis, and selected classical ML models (SVM, Logistic Regression (LR), K-Nearest

Table 2. Diagnosis results of AUC under five-fold cross-validation of three-way classification.

Group	First-order + shape + texture	First-order + texture	First-order + shape	Shape + texture	First-order	Texture	Shape
Fold0	0.8272	0.8786	0.9177	0.6728	0.8539	0.6852	0.6728
Fold1	0.8395	0.8045	0.784	0.6399	0.7531	0.6872	0.7058
Fold2	0.8704	0.7202	0.8292	0.8519	0.7593	0.6708	0.6543
Fold3	0.9156	0.8292	0.858	0.8272	0.8045	0.8601	0.8189
Fold4	0.9115	0.8909	0.7737	0.8272	0.7881	0.6523	0.7551
Mean ± standard	0.8728 ± 0.0361	0.8247 ± 0.0611	0.8325 ± 0.0524	0.7638 ± 0.0888	0.7918 ± 0.0363	0.7111 ± 0.0755	0.7214 ± 0.0596

AUC: area under the curve.

Table 3. Diagnosis results of ACC under five-fold cross-validation in different groups.

Group	First-order + shape + texture	First-order + texture	First-order + shape	Shape + texture	First-order	Texture	Shape
AD vs. MCI	0.8667	0.7889	0.7667	0.8222	0.6889	0.7111	0.6889
MCI vs. NC	0.8556	0.8333	0.7778	0.7667	0.7222	0.7444	0.6889
AD vs. NC	0.9222	0.9000	0.8222	0.8444	0.8222	0.7667	0.8000
AD vs. MCI vs. NC	0.7111	0.6741	0.6519	0.6296	0.5778	0.5630	0.5185

ACC: average accuracy curve; AD: Alzheimer's disease; MCI: Mild Cognitive Impairment; NC: normal control.
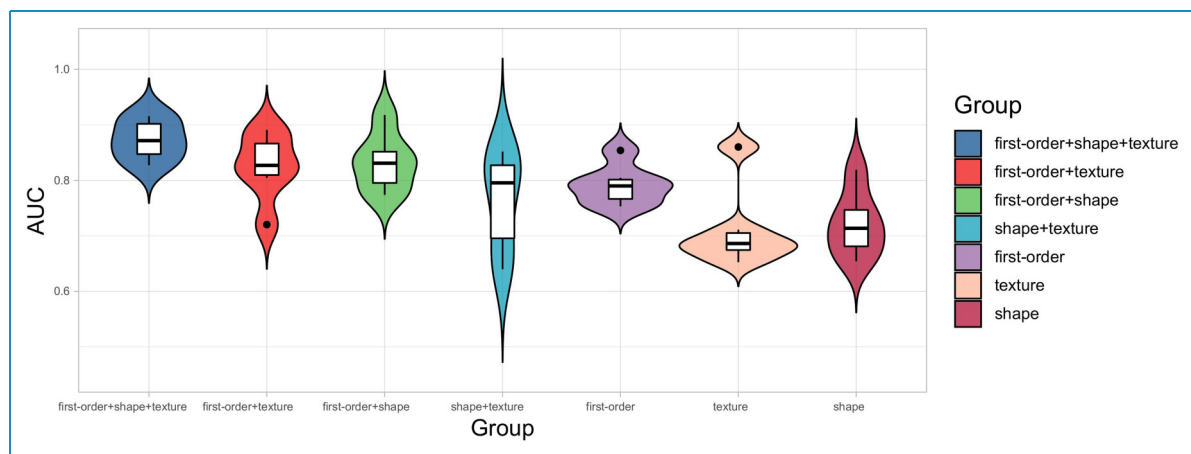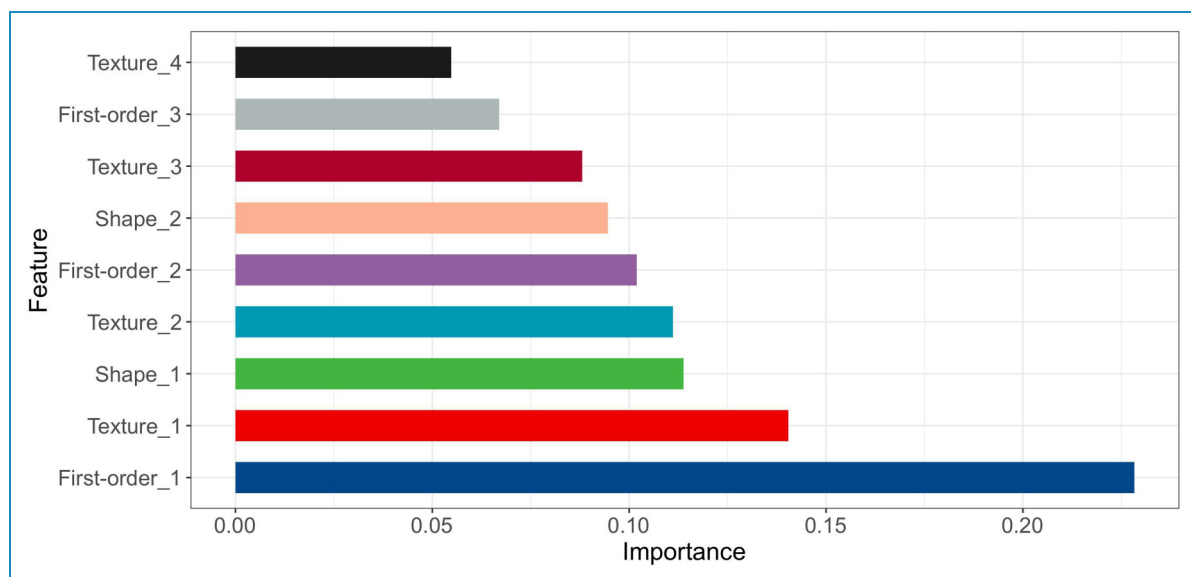

Figure 4. Violin plot of the metrics from the three-way classification experiments showing model performance.

Figure 5. Feature importance ranking from TabNet.

Table 4. Performance comparison in different models of three-way classification.

Model	SVM	LR	KNN	NB	RF	TabNet
AUC	0.8111	0.7938	0.7152	0.7267	0.8412	0.8728
ACC	0.6741	0.5704	0.5185	0.5259	0.7037	0.7111

ACC: average accuracy curve; AUC: area under the curve; KNN: K-Nearest Neighbors; LR: Logistic Regression; NB: Naive Bayes; RF: Random Forest; SVM: Support Vector Machine.

Neighbors (KNN), Naive Bayes (NB), and Random Forest (RF)) as the control groups.[45–47] To ensure a fair comparison, the baseline methods were trained and evaluated using the exact same preprocessed data as the proposed model. Both the proposed model and the baseline methods were trained under identical conditions, including the same training data and computational resources. The performance of all models was assessed using five-fold cross-validation, and the reported results represent the average performance across all folds. This approach ensures a robust and unbiased comparison between the proposed model and the baseline methods. As shown in Table 4, our findings demonstrated that the proposed model base on combined radiomic feature set (first-order, shape, and texture) significantly outperforms the control groups, achieving a maximum improvement in average AUC of 0.1576 and average ACC of 0.1926. In addition, under five-fold cross-validation, the TabNet model achieved mean F1 scores of 0.8647 for AD vs. MCI, 0.8480 for MCI vs. NC, and 0.9178 for AD vs. NC.

Binary classification experiments (AD vs. MCI, MCI vs. NC, and AD vs. NC) were also conducted, with results detailed in Tables 3 and 5 to 7 and Figure 6. For the AD versus MCI group, the proposed model achieved an average AUC of 0.8778 and an average ACC of 0.8667. For the MCI versus NC group, the model achieved an average AUC of 0.8864 and an average ACC of 0.8556. For the AD versus NC group, the model achieved an average AUC of 0.9506 and an average ACC of 0.9222. Among these tasks, the combined radiomic features (first-order, shape, and texture) yielded the better average AUC and ACC values. When individual radiomic features were utilized, the first-order features consistently demonstrated the highest average AUC, consistent with the 3-way classification results.

Additionally, binary classification for different AD subgroups using ML models, and the results shown in Table 8 showed that TabNet outperformed other control methods. Specifically, in the AD versus MCI comparison, the maximum increases in average AUC and average ACC over the control groups were 0.1247 and 0.1889, respectively. For the MCI versus NC comparison, these increases were 0.2271 and 0.2889, respectively. Finally, in the AD versus NC comparison, the maximum increases in average AUC and average ACC were 0.0802 and 0.1444, respectively.

Moreover, we conducted ablation experiments without the LASSO-based feature selection step. The results show that the inclusion of LASSO significantly improves model performance, confirming its role in reducing noise and enhancing interpretability. These findings have been added to Table 9.

Discussion

In this study, we selected 3D T1WI data to extract MR radiomic features and constructed an MR radiomics-based DL model for AD diagnosis. Our proposed model had the highest overall classification AUC across diagnostic tasks compared with other models.

Currently, MRI is widely used for tasks such as the diagnosis of different stages of AD and the prediction of the conversion from MCI to AD.[48] Compared to commonly used imaging techniques such as PET, MR provides reliable diagnostic information without the risk of radiation exposure,[49] making it more suitable for long-term follow up and monitoring. DL models can automatically extract core features from radiomics, and this end-to-end learning framework offers higher efficiency and superior performance when processing large-scale datasets.[1] Unlike traditional ML methods, DL models are particularly well-suited for handling high-dimensional data, enabling them to capture subtle structures and patterns in images. This capability is especially important for diagnosing complex neurological diseases such as AD.[50]

This study highlights the differential performance of first-order, shape, and texture radiomic features under the TabNet model for the diagnosis of AD, MCI, and NC. Our findings indicate that while first-order features achieved the best performance when applied independently, combining first-order, shape, and texture features resulted in the highest classification accuracy for the AD classification. These results are consistent with prior studies[51,52] showing that first-order features, which describe the intensity distribution of voxel values, are particularly effective for detecting global brain tissue alterations associated with AD, such as atrophy and cortical thinning.[53] The synergistic performance observed when combining the three feature types can be attributed to the complementary information they provide. Shape features capture macroscopic structural changes such as hippocampal atrophy, which are hallmark indicators

Table 5. The results of AUC of the binary classification (AD vs. MCI).

Group	First-order + shape + texture	First-order + texture	First-order + shape	Shape + texture	First-order	Texture	Shape
Fold0	0.9753	1.0000	0.9630	0.9259	0.8519	0.8642	0.7037
Fold1	0.8951	0.9012	0.9136	0.8889	0.8395	0.7037	0.7901
Fold2	0.7654	0.9012	0.9136	0.8395	0.8889	0.6543	0.6543
Fold3	0.8889	0.8519	0.8395	0.8519	0.7284	0.7531	0.642
Fold4	0.8642	0.6543	0.5309	0.8148	0.4938	0.7284	0.7407
Mean ± standard	0.8778 ± 0.0674	0.8617 ± 0.1143	0.8321 ± 0.1557	0.8642 ± 0.0390	0.7605 ± 0.1437	0.7407 ± 0.0698	0.7062 ± 0.0549

AUC: area under the curve; AD: Alzheimer's disease; MCI: Mild Cognitive Impairment.

Table 6. The results of AUC of the binary classification (MCI vs. NC).

Group	First-order + shape + texture	First-order + texture	First-order + shape	Shape + texture	First-order	Texture	Shape
Fold0	0.8889	0.9012	0.8272	0.7901	0.8642	0.7160	0.7654
Fold1	0.8025	0.8395	0.8395	0.9012	0.7778	0.6790	0.8025
Fold2	0.9259	1.0000	0.9753	0.9259	0.7037	0.8148	0.6914
Fold3	1.0000	0.9259	0.9630	1.0000	0.9877	0.8519	0.9630
Fold4	0.8148	0.7654	0.7531	0.6296	0.6543	0.7531	0.7284
Mean ± standard	0.8864 ± 0.0730	0.8864 ± 0.0794	0.8716 ± 0.0850	0.8494 ± 0.1289	0.7975 ± 0.1186	0.7630 ± 0.0632	0.7901 ± 0.0940

AUC: area under the curve; MCI: Mild Cognitive Impairment; NC: normal control.

Table 7. The results of AUC of the binary classification (AD vs. NC).

Group	First-order + shape + texture	First-order + texture	First-order + shape	Shape + texture	First-order	Texture	Shape
Fold0	0.9506	1.0000	0.9630	0.9012	0.9753	0.8272	0.9012
Fold1	1.0000	1.0000	0.9506	1.0000	0.8395	0.9136	0.8765
Fold2	1.0000	0.9753	0.9877	0.8519	0.9877	0.8642	0.9630
Fold3	0.9012	0.8519	0.9506	0.9259	0.8395	0.7901	0.8889
Fold4	0.9012	0.9012	0.9136	0.8395	0.8272	0.8025	0.7407
Mean ± standard	0.9506 ± 0.0442	0.9457 ± 0.0592	0.9531 ± 0.0239	0.9037 ± 0.0576	0.8938 ± 0.0718	0.8395 ± 0.0449	0.8741 ± 0.0730

AUC: area under the curve; AD: Alzheimer's disease; NC: normal control.
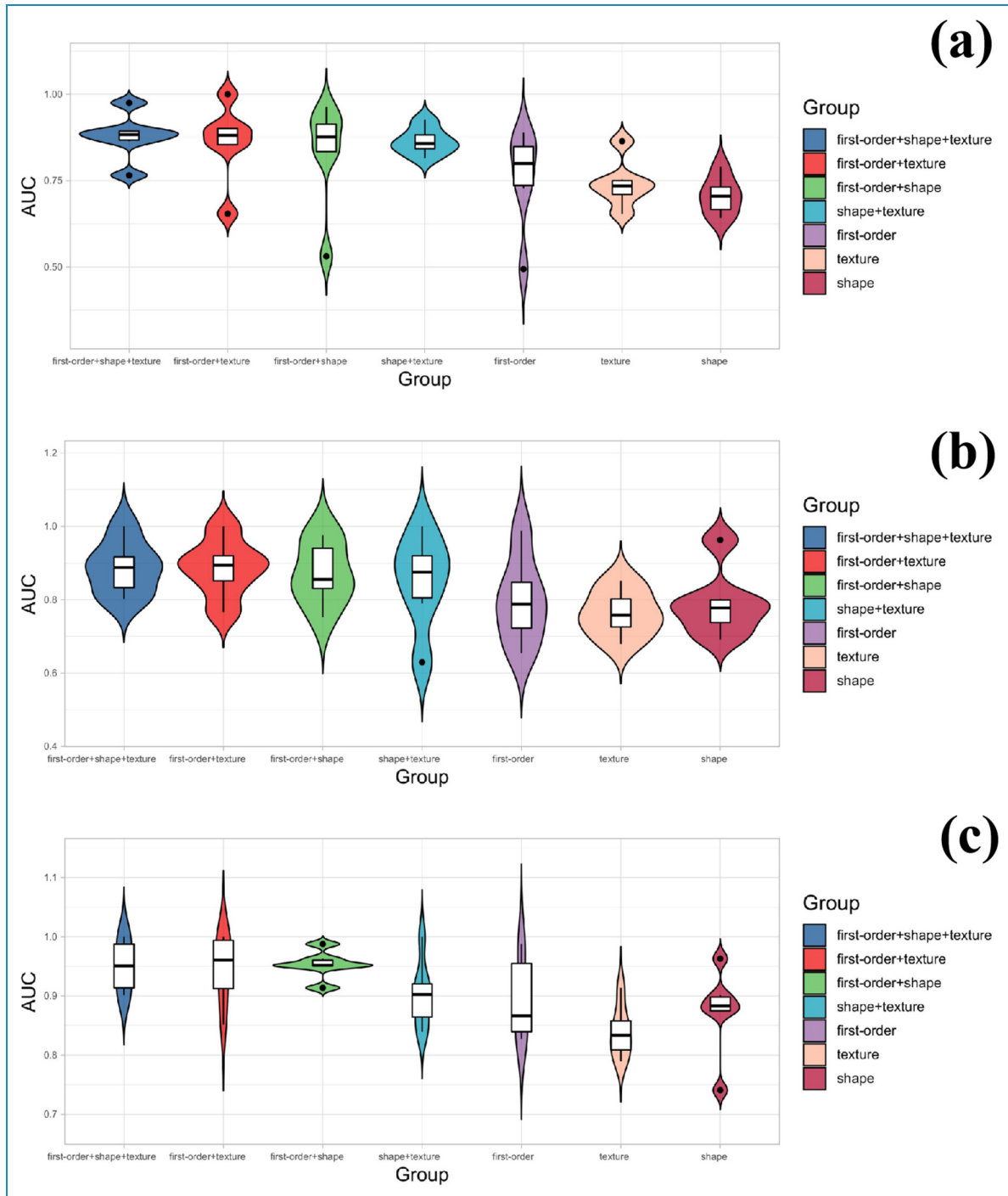
Figure 6. Violin plot of the metrics from the binary classification experiments showing model performance: (a) AD versus MCI; (b) MCI versus NC; (c) AD versus NC. AD: Alzheimer's disease; MCI: Mild Cognitive Impairment; NC: normal control.

of AD.[54] Meanwhile, texture features quantify microstructural heterogeneity, reflecting subtle changes in gray matter integrity and connectivity that are often early indicators of neurodegeneration. The combination of these features enables a more holistic representation of the underlying pathophysiology, allowing the TabNet model to achieve superior classification performance.

During the training process, we compared the selected TabNet model with four classic ML models. The results showed that the TabNet model significantly outperformed the control group in both binary classification (AD vs. MCI, AD vs. NC, MCI vs. NC) and 3-way classification (AD vs. MCI vs. NC). TabNet employs a "sequential attention" mechanism, enabling the model to dynamically focus

Table 8. Performance comparison of different models of binary classification.

		SVM	LR	KNN	NB	RF	TabNet
AD vs. MCI	AUC	0.8395	0.8074	0.7531	0.7654	0.8642	0.8778
	ACC	0.7778	0.7222	0.6778	0.6889	0.7778	0.8667
MCI vs. NC	AUC	0.8222	0.6914	0.6593	0.7012	0.8815	0.8864
	ACC	0.7556	0.6667	0.6444	0.5667	0.7889	0.8556
AD vs. NC	AUC	0.9086	0.9235	0.8704	0.8765	0.9284	0.9506
	ACC	0.8333	0.8333	0.8000	0.7778	0.8111	0.9222

ACC: average accuracy curve; AUC: area under the curve; KNN: K-Nearest Neighbors; LR: Logistic Regression; NB: Naive Bayes; RF: Random Forest; SVM: Support Vector Machine.

Table 9. Ablation experimental results of LASSO.

		AD vs. MCI vs. NC	AD vs. MCI	MCI vs. NC	AD vs. NC
AUC	Remove LASSO	0.6303	0.4667	0.5810	0.7042
	LASSO	0.8728	0.8778	0.8864	0.9506
ACC	Remove LASSO	0.4874	0.4875	0.5626	0.6592
	LASSO	0.7111	0.8667	0.8556	0.9222
F1	Remove LASSO	0.4874	0.4912	0.5580	0.7054
	LASSO	0.7111	0.8647	0.8480	0.9178

ACC: average accuracy curve; AD: Alzheimer's disease; AUC: area under the curve; LASSO: Least Absolute Shrinkage and Selection Operator; MCI: Mild Cognitive Impairment; NC: normal control.

on different features (e.g. first-order, shape, and texture) at each decision step. This dynamic feature selection process enhances learning efficiency by prioritizing the most relevant features for AD diagnosis. For instance, our findings reveal that first-order features significantly contribute to the model's performance, likely due to their ability to capture intensity-based variations in MRI data. Furthermore, TabNet quantifies the importance of specific features across the entire training set, thereby facilitating global interpretability. This capability not only improves the transparency of the model's decision-making process but also offers valuable insights into the biological relevance of radiomic features. As illustrated in Figure 5, the distribution of feature weights underscores the dominance of first-order features in distinguishing AD from MCI and NC. The selective attention mechanism ensures the efficient utilization of high-dimensional radiomic data, capturing subtle patterns that traditional ML methods may overlook. This attribute is particularly critical for addressing the complex and heterogeneous nature of AD-related imaging data, further enhancing the model's diagnostic accuracy and robustness.

This study has several limitations and directions for future improvement. First, the success of radiomics depends on multicenter and multimodal datasets. However, existing MR-based radiomics studies for AD primarily rely on the ADNI dataset. Future research could consider including additional databases such as AlzData and OASIS for cross-validation and increasing the sample size of independent validation datasets to improve the generalization ability of DL models.[55,56] Second, the ability of MR radiomics to explain underlying neurobiological mechanisms is limited, particularly in terms of the lack of interpretation of potential biological and molecular mechanisms. Further research is needed to enhance the interpretability of results and correlate them with the pathophysiological mechanisms and clinical manifestations of AD. Finally, the progression of AD is almost irreversible, and the ultimate goal of most AD research is to identify the disease progression as early as possible and intervene in a timely manner to prevent deterioration in high-risk populations. Therefore, obtaining longitudinal follow-up data will be crucial in future studies.

Overall, our study demonstrates that DL can accurately diagnose AD using MR radiomic features. The used DL model was validated on an independent dataset, and the results indicate that the model can serve as a valuable decision support tool for radiology researchers and clinicians in the early diagnosis of AD using MR radiomic features.

Conclusion

This study proposes a DL-based diagnostic model for AD using MR radiomics features. This model effectively

extracts high-sensitivity AD-related features from MR radiomic data. Compared to existing methods, the model demonstrates significant advantages in AD diagnosis, achieving higher AUC in comprehensive classification tasks across different stages of AD. The method validates the value of MR radiomics features. In the future, the radiomics-based DL model has the potential to become a practical tool for computer-assisted diagnosis of AD.

Guarantor

JZ.

ORCID iD

Zengbei Yuan (iD) https://orcid.org/0000-0002-9254-8430

Abbreviations

3D T1WI	three-dimensional T1-weighted imaging
ACC	accuracy
AD	Alzheimer's disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
aMCI	amnestic mild cognitive impairment
APRG	Adaptive Probability Region-Growing
AUC	area under the curve
BN	batch normalization
CNNs	Convolutional Neural Networks
C-PAS	Chinese Preclinical Alzheimer's Disease Study
DARTEL	Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra
DL	deep learning
DPC	Publications Committee DPC
EOAD	early-onset Alzheimer's disease
FC	fully connected
FWE	Family-Wise Error
GLCM	Gray-Level Co-occurrence Matrix
GLDM	Gray-Level Dependence Matrix
GLRLM	Gray-Level Run-Length Matrix
GLSZM	Gray-Level Size Zone Matrix
GLU	gated linear unit
IRBs	Institutional Review Boards
KNN	K-nearest neighbors
LASSO	Least Absolute Shrinkage and Selection Operator
LOAD	late-onset Alzheimer's disease
LR	logistic regression
MCI	Mild Cognitive Impairment
ML	machine learning
NB	Naive Bayes
NC	normal control
NGTDM	Neighborhood Gray-Tone Difference Matrix
NIAAA	National Institute on Aging and Alzheimer's Association
OC	old controls
OvR	One-vs-Rest
PET	Positron emission tomography
ReLU	Rectified linear unit
RF	Random Forest
ROI	regions of interest
SD	standard deviation
sMRI	Structural Magnetic Resonance Imaging
SVM	support vector machine
YC	young controls

Statements and declarations

Conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

References

1. Zhou P, Zeng R, Yu L, et al. Deep-learning radiomics for discrimination conversion of Alzheimer's disease in patients with mild cognitive impairment: a study based on 18F-FDG PET imaging. *Front Aging Neurosci* 2021; 13: 764872.

2. Salvatore C, Battista P and Castiglioni I. Frontiers for the early diagnosis of AD by means of MRI brain imaging and support vector machines. *Curr Alzheimer Res* 2016; 13: 509–533.

3. Basher A, Kim BC, Lee KH, et al. Volumetric feature-based Alzheimer's disease diagnosis from sMRI data using a convolutional neural network and a deep neural network. *IEEE Access* 2022; 9: 29870–29882.

4. Yin TT, Cao MH, Yu JC, et al. T1-weighted imaging-based hippocampal radiomics in the diagnosis of Alzheimer's disease. *Acad Radiol* 2024; 31: 5183–5192.

5. Végh MJ, Heldring CM, Kamphuis W, et al. Reducing hippocampal extracellular matrix reverses early memory deficits in

a mouse model of Alzheimer's disease. *Acta Neuropathol Commun* 2014; 2: 76.

6. Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2015; 278: 563–577.

7. Amadasun M and King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern Syst* 1989; 19: 1264–1274.

8. Galloway MM. Texture analysis using grey level run lengths. *Recon Technical Report N* 1974; 4: 75.

9. Haralick RM and Shanmugam K. Textural features for image classification. *IEEE Trans Syst Man Cybern Syst* 1973; SMC-3: 610–621.

10. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012; 48: 441–446.

11. Khojaste-Sarakhsi M, Haghighi SS, Ghomi SMTF, et al. Deep learning for Alzheimer's disease diagnosis: a survey. *Artif Intell Med* 2022; 130: 102332.

12. Sun W, Zheng B and Qian W. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. *Comput Biol Med* 2017; 89: 530–539.

13. Chaddad A, Desrosiers C and Niazi T. Deep radiomic analysis of MRI related to Alzheimer's disease. *IEEE Access* 2018; 6: 58213–58221.

14. Feng Q, Chen Y, Liao Z, et al. Corpus callosum radiomics-based classification model in Alzheimer's disease: a case-control study. *Front Neurol* 2018; 9: 618.

15. Chen Z, Bi S, Shan Y, et al. Multiparametric hippocampal signatures for early diagnosis of Alzheimer's disease using ^{18}F-FDG PET/MRI radiomics. *CNS Neurosci Ther* 2024; 30: e14539.

16. Leandrou S, Lamnisos D, Bougias H, et al. A cross-sectional study of explainable machine learning in Alzheimer's disease: diagnostic classification using MR radiomic features. *Front Aging Neurosci* 2023; 15: 1149871.

17. Park HY, Shim WH, Suh CH, et al. Development and validation of an automatic classification algorithm for the diagnosis of Alzheimer's disease using a high-performance interpretable deep learning network. *Eur Radiol* 2023; 33: 7992–8001.

18. Cheung EYW, Wu RWK, Chu ESM, et al. Integrating demographics and imaging features for various stages of dementia classification: feed forward neural network multi-class approach. *Biomedicines* 2024; 12: 896.

19. Yang M, Meng S, Wu F, et al. Automatic detection of mild cognitive impairment based on deep learning and radiomics of MR imaging. *Front Med (Lausanne)* 2024; 11: 1305565.

20. Feng Q, Song Q, Wang M, et al. Hippocampus radiomic biomarkers for the diagnosis of amnestic mild cognitive impairment: a machine learning method. *Front Aging Neurosci* 2019; 11: 323.

21. Du Y, Zhang S, Fang Y, et al. Radiomic features of the hippocampus for diagnosing early-onset and late-onset Alzheimer's disease. *Front Aging Neurosci* 2022; 13: 789099.

22. Liu S, Jie C, Zheng W, et al. Investigation of underlying association between whole brain regions and Alzheimer's disease: a research based on an artificial intelligence model. *Front Aging Neurosci* 2022; 14: 872530.

23. Zhou K, Piao S, Liu X, et al. A novel cascade machine learning pipeline for Alzheimer's disease identification and prediction. *Front Aging Neurosci* 2023; 14: 1073909.

24. Feng Q, Niu J, Wang L, et al. Comprehensive classification models based on amygdala radiomic features for Alzheimers' disease and mild cognitive impairment. *Brain Imaging Behav* 2021; 15: 2377–2386.

25. Peng J, Wang W, Song Q, et al. ^{18}F-FDG-PET Radiomics based on white matter predicts the progression of mild cognitive impairment to Alzheimer disease: a machine learning study. *Acad Radiol* 2023; 30: 1874–1884.

26. Zhao Y, Zhang J, Chen Y, et al. A novel deep learning radiomics model to discriminate AD, MCI and NC: an exploratory study based on tau PET scans from ADNI. *Brain Sci* 2022; 12: 1067.

27. Cui L, Huang L, Pan FF, et al. Chinese Preclinical Alzheimer's disease study (C-PAS): design and challenge from PET acceptance. *J Prev Alzheimers Dis* 2023; 10: 571–580.

28. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology* 2010; 74: 201–209.

29. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage* 2007; 38: 95–113.

30. Yan CG, Wang XD, Zuo XN, et al. DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 2016; 14: 339–351.

31. Henneman WJP, Sluimer JD, Barnes J, et al. Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. *Neurology* 2009; 72: 999–1007.

32. Cavedo E, Boccardi M, Ganzola R, et al. Local amygdala structural differences with 3T MRI in patients with Alzheimer disease. *Neurology* 2011; 76: 727–733.

33. Duara R, Loewenstein DA, Potter E, et al. Medial temporal lobe atrophy on MRI scans and the diagnosis of Alzheimer disease. *Neurology* 2008; 71: 1986–1992.

34. Ernst T, Chang L, Melchor R, et al. Frontotemporal dementia and early Alzheimer disease: differentiation with frontal lobe H-1 MR spectroscopy. *Radiology* 1997; 203: 829–836.

35. Aggleton JP, Pralus A, Nelson AJD, et al. Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to papez circuit. *Brain* 2016; 139: 1877–1890.

36. Ibrahim A, Vallieres M, Woodruff H, et al. Radiomics analysis for clinical decision support in nuclear medicine. *Semin Nucl Med* 2019; 49: 438–449.

37. Li Y, Jiang J, Lu J, et al. Radiomics: a novel feature extraction method for brain neuron degeneration disease using 18F-FDG PET imaging and its implementation for

Alzheimer's disease and mild cognitive impairment. *Ther Adv Neurol Disord* 2019; 12: 1756286419838682.

38. Arık SO and Pfister T. Tabnet: attentive interpretable tabular learning. *AAAI* 2021; 35: 6679–6687.

39. Wang Y, Jin C, Ma L, et al. A robust TabNet-based multi-classification algorithm for infrared spectral data of Chinese herbal medicine with high-dimensional small samples. *J Pharm Biomed Anal* 2024; 242: 116031.

40. Yan J, Xu T, Yu Y, et al. Rainfall forecast model based on the TabNet model. *Water (Basel)* 2021; 13: 1272.

41. Qi H, Hu Y, Fan R, et al. Tab-Cox: an interpretable deep survival analysis model for patients with nasopharyngeal carcinoma based on TabNet. *IEEE J Biomed Health Inform* 2024; 28: 4937–4950.

42. Wang S and Zhang X. Research on credit default prediction model based on TabNet-stacking. *Entropy (Basel)* 2024; 26: 861.

43. Roh HR and Lee JM. TabNet-based self-supervised fault diagnosis in multivariate time-series process data without labels. *IFAC-Papers OnLine* 2024; 58: 829–834.

44. Adam R, Dell'Aquila K, Hodges L, et al. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review. *Breast Cancer Res* 2023; 25: 87.

45. Zheng Y, Zhou D, Liu H, et al. CT-based radiomics analysis of different machine learning models for differentiating benign and malignant parotid tumors. *Eur Radiol* 2022; 32: 6953–6964.

46. Mirzaei G and Adeli H. Machine learning techniques for diagnosis of Alzheimer disease, mild cognitive disorder, and other types of dementia. *Biomed Signal Process Control* 2022; 72: 103293.

47. Yang C, Jiang ZK, Liu LH, et al. Pre-treatment ADC image-based random forest classifier for identifying resistant rectal adenocarcinoma to neoadjuvant chemoradiotherapy. *Int J Colorectal Dis* 2020; 35: 101–107.

48. Li H, Habes M, Wolk DA, et al. A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimers Dement* 2019; 15: 1059–1070.

49. Schmidt GP, Haug AR, Schoenberg SO, et al. Whole-body MRI and PET-CT in the management of cancer patients. *Eur Radiol* 2006; 16: 1216–1225.

50. Jo T, Nho K and Saykin AJ. Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. *Front Aging Neurosci* 2019; 11: 220.

51. Ding Y, Zhang C, Lan T, et al. Classification of Alzheimer's disease based on the combination of morphometric feature and texture feature. *IEEE Int Conf Bioinf Biomed* 2015: 409–412.

52. Chu T, Yang C, Lu M, et al. A new method for classification of Alzheimer's disease combined with structural magnetic resonance imaging texture features. *Sheng wu yi xue Gong Cheng xue za zhi=J Biomed Eng=Shengwu Yixue Gongchengxue Zazhi* 2019; 36: 94–100.

53. Feng F, Wang P, Zhao K, et al. Radiomic features of hippocampal subregions in Alzheimer's disease and amnestic mild cognitive impairment. *Front Aging Neurosci* 2018; 10: 290.

54. Wei Y, Huang N, Liu Y, et al. Hippocampal and amygdalar morphological abnormalities in Alzheimer's disease based on three Chinese MRI datasets. *Curr Alzheimer Res* 2020; 17: 1221–1231.

55. Zou T, Zhou X, Wang Q, et al. Associations of serum DNA methylation levels of chemokine signaling pathway genes with mild cognitive impairment (MCI) and Alzheimer's disease (AD). *PLoS ONE* 2023; 18: e0295320.

56. Marcus DS, Wang TH, Parker J, et al. Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J Cogn Neurosci* 2007; 19: 1498–1507.