Research Article

Prediction of Early Alzheimer Disease by Hippocampal Volume Changes under Machine Learning Algorithm

Qun Shang [ID], Qi Zhang [ID], Xiao Liu [ID], and Lingchen Zhu [ID]

Department of Radiology, Zibo Central Hospital, Zibo, 255000 Shandong, China

Correspondence should be addressed to Lingchen Zhu; 6180205009@stu.jiangnan.edu.cn

This research was aimed at discussing the application value of different machine learning algorithms in the prediction of early Alzheimer's disease (AD), which was based on hippocampal volume changes in magnetic resonance imaging (MRI). In the research, the 84 cases in American Alzheimer's disease neuroimaging initiative (ADNI) database were selected as the research data. Based on the scoring results of cognitive function, all cases were divided into three groups, including cognitive function normal (normal group), early mild cognitive impairment (e-MCI group), and later mild cognitive impairment (l-MCI group) groups. Each group included 28 cases. The features of hippocampal volume changes in MRI images of the patients in different groups were extracted. The samples of training set and test set were established. Besides, the established support vector machine (SVM), decision tree (DT), and random forest (RF) prediction models were used to predict e-MCI. Metalinear regression was utilized to analyze MRI feature data, and the predictive accuracy, sensitivity, and specificity of different models were calculated. The result showed that the volumes of hippocampal left CA1, left CA2-3, left CA4-DG, left presubiculum, left tail, right CA2-3, right CA4-DG, right presubiculum, and right tail in e-MCI group were all smaller than those in normal group ($P < 0.01$). The corresponding volume of hippocampal subregions in l-MCI group was remarkably reduced compared with that in normal group ($P < 0.001$). The volumes of regions left CA1, left CA2-3, left CA4-DG, right CA2-3, right CA4-DG, and right presubiculum were all positively correlated with logical memory test-delay recall (LMT-DR) score ($R^2 = 0.1702$, 0.3779, 0.1607, 0.1620, 0.0426, and 0.1309; $P < 0.001$). The predictive accuracy of training set sample by DT, SVM, and RF was 86.67%, 93.33%, and 98.33%, respectively. Based on the changes in the volumes of left CA4-DG, right CA2-3, and right CA4-DG, the predictive accuracy of e-MCI and l-MCI by RF model was both higher than those by DT model ($P < 0.01$). Besides, the predictive accuracy, sensitivity, and specificity of e-MCI by RF model was all notably higher than those by DT model ($P < 0.01$). The above results demonstrated that the effective early AD prediction models were established by the volume changes in hippocampal subregions, which was based on RF in the research. The establishment of early AD prediction models offered certain reference basis to the diagnosis and treatment of AD patients.

1. Introduction

Alzheimer's disease (AD) is a common type of symptom among senile dementia. The clinical features of AD are progressive cognitive decline and behavior abnormality. The incidence of AD is positively correlated with age. The incidence of AD reaches about 13% among the population aged 65, and it is as high as 43% among the population aged 85 and above [1]. With the aggravation of current global population aging, the number of AD patients is growing obviously. In 2018, 50 million people suffered from AD all over

the world. It is predicted that the number of AD patients will increase to 82 million by 2030 [2]. Mild cognitive impairment (MCI) is the early phase of the development of AD disease. Effective intervention measures for MCI patients can restore their cognitive state back to normal levels [3]. With the change in disease status, amyloid beta deposition affects hippocampal volumes to some extent. The changes in hippocampal volumes promote the early diagnosis of AD population [4]. With the continuous development of imaging and medical technologies in recent years, structural magnetic resonance imaging (MRI), positron emission tomography

(PET), electroencephalogram (EEG), brain metabolic imaging, and cerebrospinal fluid biomarker examination are all adopted in AD diagnosis. Among the above methods, brain metabolic imaging and cerebrospinal fluid biomarker examination are both traumatic for human body, especially for advanced AD patients [5]. The recall rate of AD by EEG is influenced by onset age and dementia severity, and it shows some defects in the diagnosis of early AD [6]. PET can display the statuses of cerebrospinal fluid in the hippocampus and sulcus, but its definition and diagnostic capacity are poorer than those of MRI [7], which is the effective auxiliary method of the diagnosis of early AD [8].

The performance of machine learning algorithm is improved and enhanced by the constantly accumulated data to seek the optimal model. The algorithm is widely applied in data analysis and mining, mode recognition, bioinformatics, and medical treatment [9]. As the method that represents the technology level of integrated learning, random forest (RF) algorithm demonstrates unique advantages in processing high-dimension data. The results of the current studies show that RF can predict multiple diseases very accurately as a combined classifier, and it possesses high tolerance of abnormal values and noises. Besides, the overfitting of RF is rare [10]. RF consists of many decision trees (DT). Therefore, DT also shows potential values in the diagnosis of early AD [11]. In addition, some researchers point out that the classification abilities of MCI and AD by support vector machines (SVM) models are both excellent with the classification accuracy reaching as high as 82.0% [12]. However, current machine learning algorithm-based models are all adopted in AD diagnosis, while early AD prediction models are seldom studied. In addition, it is unknown which algorithm shows more advantages in early AD prediction.

To conclude, machine learning algorithm-based models showed significant advantages in AD diagnosis. However, there were few studies on the prediction of early AD by these models, and it was still unknown which algorithm demonstrated more advantages in early AD prediction. Hence, the MRI images of AD patients were included as the research objects. The AD prediction models of RF, DT, and SVM were established based on the analysis of hippocampal volume changes, and the application values of different models in early AD prediction were discussed to provide a referable basis for the diagnosis and treatment of AD patients.

2. Materials and Methods

2.1. Experimental Data and Grouping. American Alzheimer's disease neuroimaging initiative (ADNI) database was selected as the research objects. A total of 84 qualified research objects were selected randomly from serial numbers between 4801 and 5315. According to clinical dementia rating (CDR) [13] score and the diagnostic standards of *American Diagnostic and Statistical Manual of Mental Disorders* [14], all the patients were rolled into three groups, including cognitive function normal group (normal group), early mild cognitive impairment group (e-MCI group), and later mild cognitive impairment group (l-MCI group). Each group included 28 cases.

The inclusion standards of normal group were as follows: no memory loss; CDR scores were 0 points; barrier-free daily activities; normal cognitive function; and no dementia occurs.

The inclusion standards of e-MCI group were as follows: CDR scores were 0.5 points; no other cognitive disorders occur; daily activities were generally barrier-free; no dementia occurs; and delayed recall (DR) scores 20 minutes after logical memory test (LMT) were as follows. The cases educated for 16 years or longer were scored between 9 and 11 points. Those educated for 8 to 15 years were scored between 5 and 8 points. Those educated for 0 to 7 years were scored 3 or 4 points.

The inclusion standards of l-MCI group were similar to those of e-MCI group. Because cognitive function level was closely related to education level, LMT-DR scores were different. LMT-DR scores of l-MCI group were shown below. The cases educated for 16 years and longer were scored 8 points or less. Those educated for 8 to 15 years were scored 4 points or less. Those educated for 0 to 7 years were scored 2 points or less.

2.2. Experimental Data Preprocessing. Before the feature extraction and classification of MRI images, they needed to be preprocessed. The preprocessing of MRI images mainly included the following steps. The first step was the removal of skulls from MRI images by brain surface extractors. The second step was spatial standardization by FreeSurfer. The third step was the segmentation of axial MRI images into multiple subimages, the search for the maximum peak of subimage grayscale value histograms as the reference grayscale value of subimage white matters, and the smoothness processing of MRI images. The fourth step was the segmentation of brain tissues according to grayscale. The fifth step was the feature extraction according to segmentation results.

2.3. SVM Prediction Model Algorithm. SVM transformed the input space into a high-dimension space by nonlinear transformation and then obtained the optimal linear classification surface in the new space [15]. The given linear separable training set A was expressed by

$$A = \{(x_i, y_i), i = 1, 2, \cdots l\}, x_i \in R_n, y_i \in \{-1, 1\}. \quad (1)$$

In equation (1), x_i referred to the ith training sample, y_i represented the actual label of the ith training sample, and l denoted the number of training samples.

SVM was aimed at searching for a decision function $f(x)$ to separate training sets. The linear separable issue was expressed by classification hyperplane in

$$(\omega \cdot x) + b = 0. \quad (2)$$

In equation (2), ω denoted weight vectors and b referred to deviation items.

The core idea of SVM was the control of the promotion ability, and then, the classification interval maximization of

classification hyperplane was transformed by

$$\min \frac{1}{2} \|\omega\|^2 \qquad (3)$$

$$\text{s.t : } y_i[(\omega \cdot x_i) + b] \geq 1, \quad i = 1, 2, \cdots l \qquad (4)$$

In equation (4), x_i referred to the input vector of the dimension m and y_i meant the label of samples.

As to approximate linear separable issues, the slack variable was introduced to transform the issues into the standard classification. A nonlinear mapping was introduced for linear inseparable issues, and low-dimension linear inseparable issues were transformed into high-dimension feature spatial linear separable issues. At last, the standard classification was adopted to obtain solutions. To avoid the complex calculation in high-dimension space, SVM adopted the kernel function $K(x, y)$ to replace the inner product $\Phi(x)\Phi(y)$ in the high-dimension space. Equation (5) was expressed as follows.

$$\min \left\{ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{l} \alpha_i \right\} \qquad (5)$$

$$\text{s.t : } y_i[(\omega \cdot x_i) + b] \geq 1 - \alpha_i, \quad i = 1, 2, \cdots l \qquad (6)$$

In equations (5) and (6), C represented the penalty coefficient adopted to balance classification intervals and misclassification specific gravity and a_i denoted slack variables.

Lagrange method was adopted to transform the above issues into a quadratic programming issue, and then the calculation method was expressed by

$$\min \frac{1}{2} \sum_{i=1}^{l} \sum_{i=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i x_j) - \sum_{j=1}^{l} \alpha_j \qquad (7)$$

$$\text{s.t.} \sum_{i=1}^{l} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, \quad i = 1, 2, \cdots l \qquad (8)$$

The above two equations were solved, and then, equation (9) was obtained as follows.

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^{l} \alpha_i y_i K(x_i \cdot x) + b \right\}. \qquad (9)$$

Based on functional theory, the kernel function corresponded to the inner product in a certain transformation space if it met Mercer conditions. Then, the classification function of SVM was expressed by

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^{k} \alpha_i^* y_i K(x_i \cdot x) + b^* \right\}. \qquad (10)$$

In equation (10), α_i^* referred to Lagrange multipliers and $b*$ represented classification thresholds.

In the research, the radial basis function was selected as the kernel function. As equation (11) indicated, γ referred to Gaussian kernel parameters.

$$K(x, x_i) = \exp \left(-\gamma \|x - x_i\|^2 \right). \qquad (11)$$

If a group of training samples could be separated by an optimal classification surface, the expected classification error rate of test samples met the following conditions, which were shown in

$$E[P(\text{error})] \leq \frac{E(\text{SV})}{n}. \qquad (12)$$

In equation (12), n referred to the number of training sets, SV denoted support vectors, $E[P(\text{error})]$ represented the expectation of the classification error rate of test samples, and $E(\text{SV})$ stood for the expectation of the number of SV. The accuracy discriminant weight w_t was expressed by

$$w_t = 1 - \frac{E(\text{SV})}{n} \boxtimes K(x, x_i) = \phi(x)\phi(x_i). \qquad (13)$$

2.4. DT and RF Prediction Model Algorithms. DT classified data based on the tree structure. Data was divided into different regions based on distribution features. Measurement purity was commonly adopted to show classification effects, and the common standard of measurement purity was information entropy [16], whose calculation method was demonstrated by

$$E = -p * \log(p) - (1 - p) * \log(1 - p). \qquad (14)$$

In equation (14), p denoted the true sample proportion in current nodes. It was assumed that the reliability of the result p_i of the ith sample was expressed by

$$L_i = p^{iy_i} * (1 - p_i)^{1-y_i} y_i. \qquad (15)$$

In equation (15), y_i referred to the class label of the ith sample. It was assumed that the probability of accurate classification followed Bernoulli distribution, and then, the target function of data sets was expressed by

$$T = \coprod_{i=1}^{l} \left[p^{iy_i} * (1 - p_i)^{1-y_i} y_i \right]. \qquad (16)$$

The logarithm and inverse of the above equation were obtained, and then, the calculation method of the target function of data sets was transformed as

$$f = -\log(T). \qquad (17)$$

The minimization solution of the function was the same thing as the maximization solution of the original function, and then, the issue was transformed into the function parameter of extremum solution, whose expression was
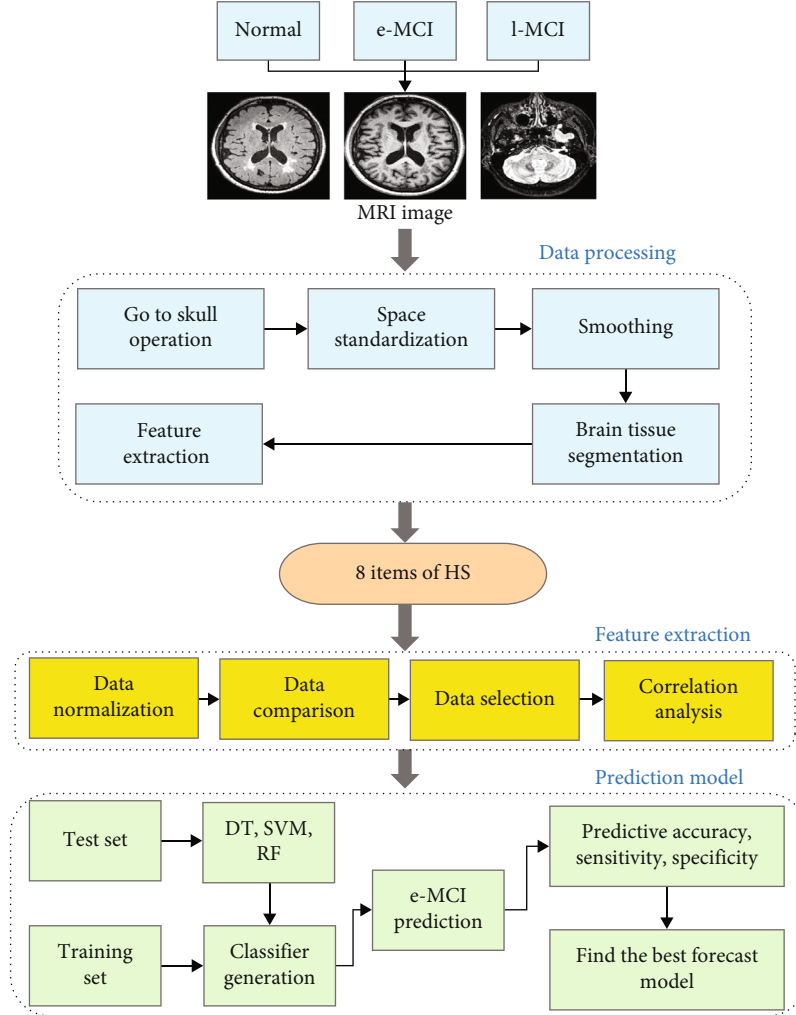
FIGURE 1: Machine learning model-based early AD prediction process.

shown in

$$p_k = \frac{\sum_{i=1}^{m} y_i}{m}.$$ (18)

In equation (18), m denoted all samples of this node and p_k referred to the proportion of true samples. The target function of data sets was transformed into

$$f_k = \sum_{i=1}^{m} y_i \log p_k - \sum_{i=1}^{m} (1 - y_i) \log (1 - p_k).$$ (19)

The normalization of node samples was adopted to eliminate the influences of sample points on the number of samples [17], and then, the calculation method of the target function of data sets was expressed by

$$\frac{f_k}{m} = p_k \log p_k - (1 - p_k) \log (1 - p_k).$$ (20)

RF was a combined classifier algorithm synthesized by many DT classification models. RF added the differences

among classification models by constructing different training sets to improve the prediction effects of classification models [18]. The classification decision of the models was expressed by

$$R(x) = \arg_{\varepsilon}^{\max} \sum_{i=1}^{k} A[r_i(x) = Y].$$ (21)

In equation (21), $R(x)$ referred to combined classification models, k represented the number of training, r_i denoted a DT classification model, and Y stood for target variables.

2.5. Establishment of Machine Learning Prediction Models. In the research, the implementation of machine learning algorithm model transformation and data prediction was based both on Matlab2016a platform. Data preprocessed by MRI was imported, and MRI data were normalized. After that, 20 groups of test and training sets were selected, respectively, for the training and classification test on data, which was aimed at establishing prediction models. Besides, the prediction accuracy, sensitivity, and specificity parameters
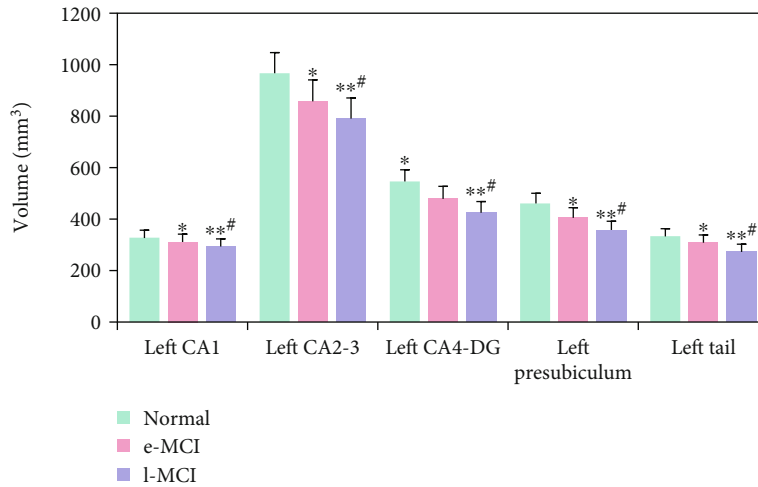
(a)
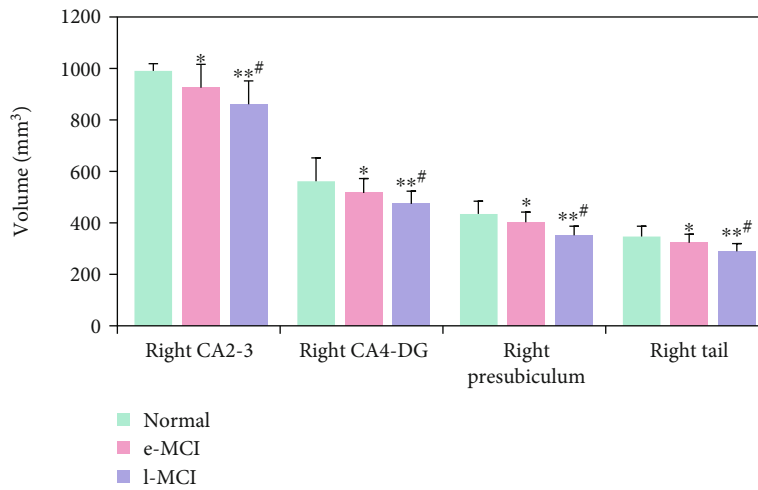
(b)

FIGURE 2: Comparison of hippocampal volumes in different groups by MRI data. (a) demonstrated the comparison of left tail hippocampal subregion volumes in different groups, and (b) showed the comparison of right tail hippocampal subregion volumes in different groups. ∗ indicated that the comparison with normal group showed significant differences, $P < 0.01$; ∗∗ meant that the comparison with normal group demonstrated extremely significant differences, $P < 0.001$; # showed that the comparison with e-MCI group revealed significant differences, $P < 0.01$.
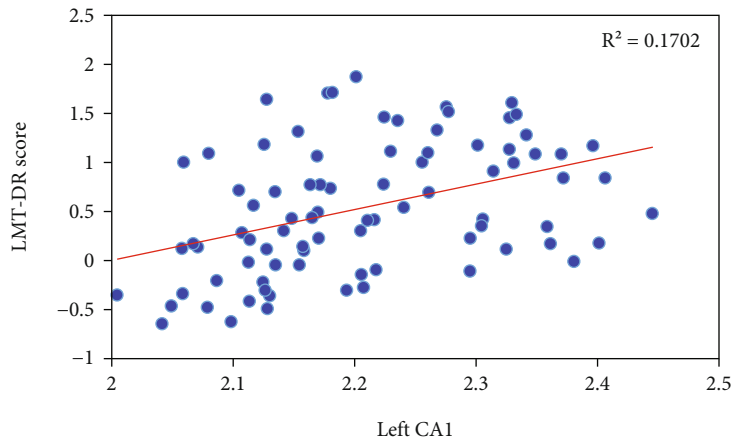

FIGURE 3: Relevance analysis of left CA1 region volumes and LMT-DR scores.
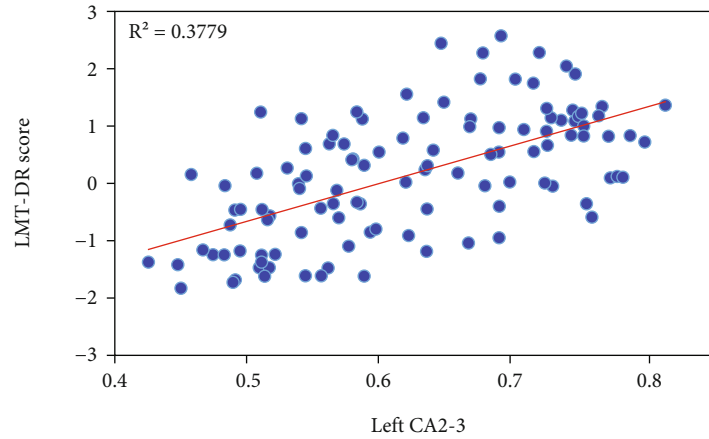
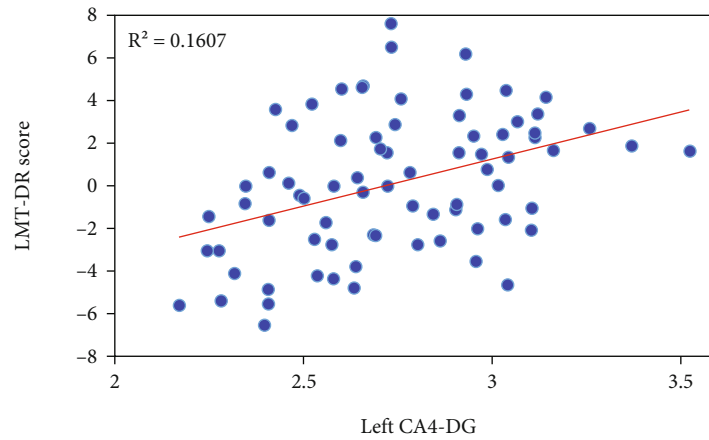FIGURE 4: Relevance analysis of left CA2-3 region volumes and LMT-DR scores.

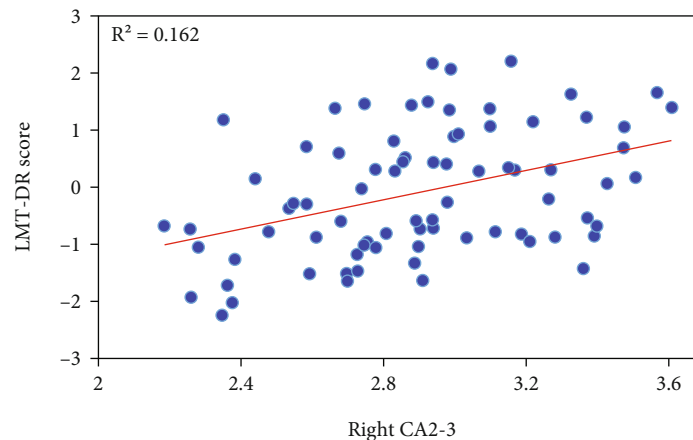FIGURE 5: Relevance analysis of left CA4-DG region volumes and LMT-DR scores.

FIGURE 6: Relevance analysis of right CA2-3 region volumes and LMT-DR scores.

of prediction models were calculated under test sets. Figure 1 demonstrated the specific process of early AD prediction based on machine learning algorithms as follows.

2.6. Observation Indexes. The volumes of hippocampal subregions, including left CA1, left CA2-3, left CA4-DG, left presubiculum, left tail, right CA2-3, right CA4-DG, right presubiculum, and right tail in different groups, were compared. The correlation between different hippocampal volumes and LMT-DR was analyzed. The predictive accuracy of early AD by prediction models established based on different algorithms and the predictive accuracy of different
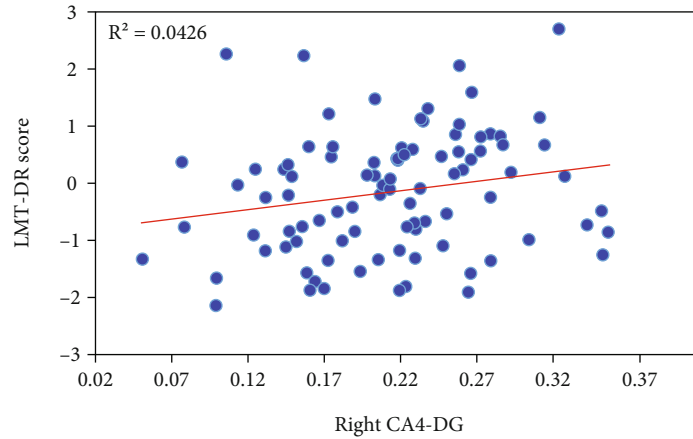
FIGURE 7: Relevance analysis of right CA4-DG region volumes and LMT-DR scores.
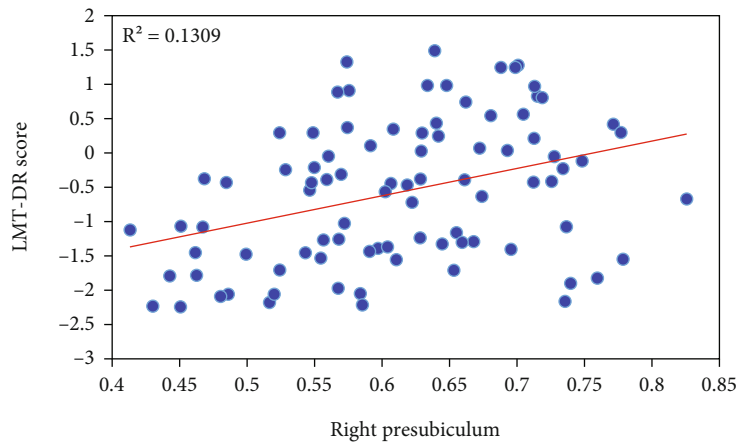

FIGURE 8: Relevance analysis of right presubiculum region volumes and LMT-DR scores.

MRI feature indexes were analyzed. Besides, the predictive accuracy, sensitivity, and specificity of early AD by different models were compared.

2.7. Statistical Methods. After being preprocessed by MRI image data, a total of 16 volume indexes in hippocampal subregions were obtained. All data were analyzed by statistical product and service solutions (SPSS) software. In addition, a one-way variance analysis was performed on the data complying with normality and homoscedasticity. The least significant difference (LSD) method was adopted in the comparison among groups. In contrast, the data that did not conform to normality and homoscedasticity test was processed by Kruskal-Wallis H nonparametric test. Pairwise comparisons among multiple independent samples were tested by Nemenyi method. Besides, MRI feature data with significant differences among groups were processed by multiple linear regression analysis. $\alpha = 0.05$ was set as the test level, and $P < 0.05$ indicated that the differences showed statistical meaning.

3. Results

3.1. Analysis of Results of MRI Data Preprocessing. After MRI data preprocessing and feature extraction, a total of 16 vol-

ume indexes in hippocampal subregions. After the multiple comparisons by a one-way variance analysis and LSD method, pairwise comparisons among groups demonstrated that 8 volume indexes in hippocampal subregions showed differences. The volumes of hippocampi in e-MCI group, including left CA1, left CA2-3, left CA4-DG, left presubiculum, left tail, right CA2-3, right CA4-DG, right presubiculum, and right tail, were all smaller than those in normal group, and the comparison between two groups demonstrated significant differences ($P < 0.01$). The volumes of hippocampi in I-MCI group, including left CA1, left CA2-3, left CA4-DG, left presubiculum, left tail, right CA2-3, right CA4-DG, right presubiculum, and right tail, were all decreased compared with those in e-MCI group, and the comparison between two groups revealed obvious differences ($P < 0.01$). Besides, the corresponding volumes of hippocampi in hippocampal subregions in I-MCI group were significantly reduced compared with those in normal group, and the comparison between the two groups indicated extremely significant differences ($P < 0.01$). All the results of the above comparisons were illustrated in Figure 2.

3.2. Relevance Analysis of Volume Changes in Hippocampal Subregions and LMT-DR. The relevance between hippocampal volumes and LMT-DR in behavioristics was analyzed. In
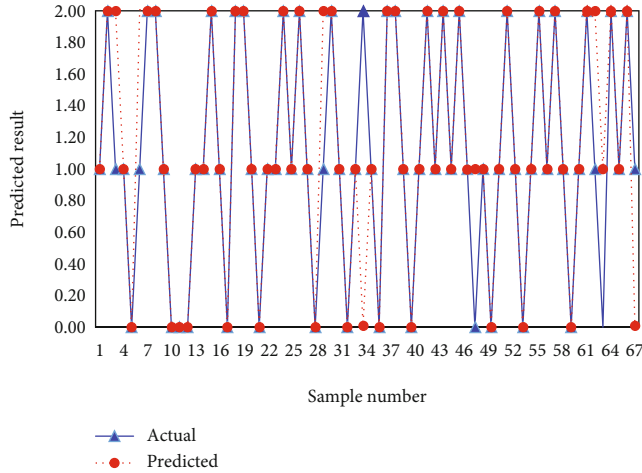
FIGURE 9: Analysis of prediction results by DT models.

FIGURE 10: Analysis of prediction results by SVM models.

FIGURE 11: Analysis of prediction results by RF models.

left hippocampal subregions, the volumes of 3 subregions, including left CA1, left CA2-3, and left CA4-DG, were all correlated with LMT-DR scores, which were shown in Figures 3, 4, and 5. The volumes of left CA1 regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.1702$, $P < 0.001$), the volumes of left CA2-3 regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.3779$, $P < 0.001$), and the volumes of left CA4-DG regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.1607$, $P < 0.001$).

The volumes of 3 subregions in right hippocampal subregions, including right CA2-3, right CA4-DG, and right presubiculum, were all correlated with LMT-DR scores, which were demonstrated in Figures 6, 7, and 8. The volumes of right CA2-3 regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.1620$, $P < 0.001$), the volumes of right CA4-DG regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.0426$, $P < 0.001$), and the volumes of right presubiculum regions were positively correlated with LMT-DR scores significantly ($R^2 = 0.1309$, $P < 0.001$).

3.3. Analysis of AD Prediction Results by Different Models. Different models established in the research were adopted to analyze the prediction results of normal, e-MCI, and l-MCI groups, respectively. The numbers including 0, 1, and 2 were adopted to symbolize normal, e-MCI, and l-MCI groups, respectively. Figure 9 displayed the prediction of different types of results by DT models. The prediction results of 8 data were inconsistent with the actual results, and the prediction accuracy of training set samples reached 86.67%.

Figure 10 presented the prediction of different types of results by SVM models. The prediction results of 4 data were inconsistent with the actual results, and the prediction accuracy of training set samples amounted to 93.33%.

Figure 11 demonstrated the prediction of different types of results by RF models. The prediction result of 1 data was inconsistent with the actual result, and the prediction accuracy of training set samples amounted to 98.33%.
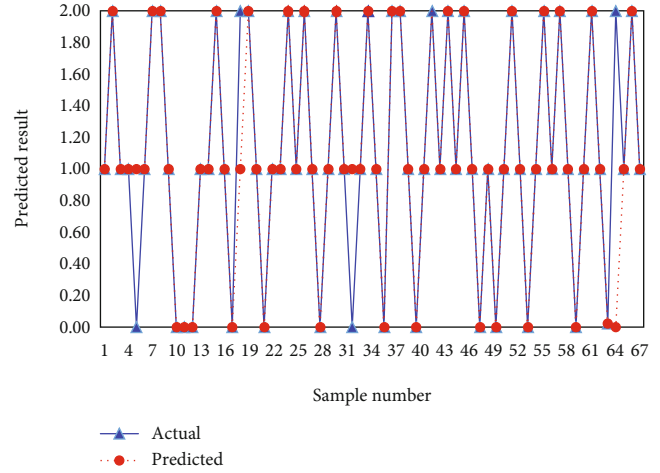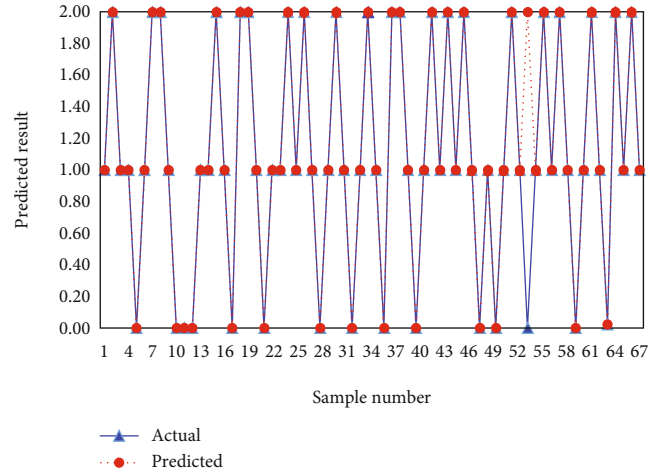
3.4. Comparison of Prediction Accuracy of Different MRI Feature Indexes. The e-MCI and l-MCI were predicted by RF, SVM, and DT models based on the volume change parameters of different hippocampal subregions, and the prediction results were compared and analyzed (Figure 12). In the prediction of the volume changes in different hippocampal subregions, the prediction accuracy by RF models was higher than that by SVM and DT models. Based on left CA1 and left CA2-3 volume changes, the prediction accuracy of e-MCI by RF models was higher than that by DT models ($P < 0.05$). Besides, the prediction accuracy of e-MCI based on left CA4-DG, right CA2-3, and right CA4-DG volume changes by RF models was significantly higher than that by DT models ($P < 0.01$). Based on left CA2-3 volume changes, the prediction accuracy of l-MCI by RF models was higher than that by DT models ($P < 0.05$), and the prediction accuracy of l-MCI was obviously higher than that by DT models, which was based on left CA4-DG, right CA2-3, and right CA4-DG volume changes ($P < 0.01$). What is more, it was observed that the prediction accuracy of e-MCI and l-MCI by different models based on left CA4-DG, right CA2-3, and right CA4-DG volume changes was
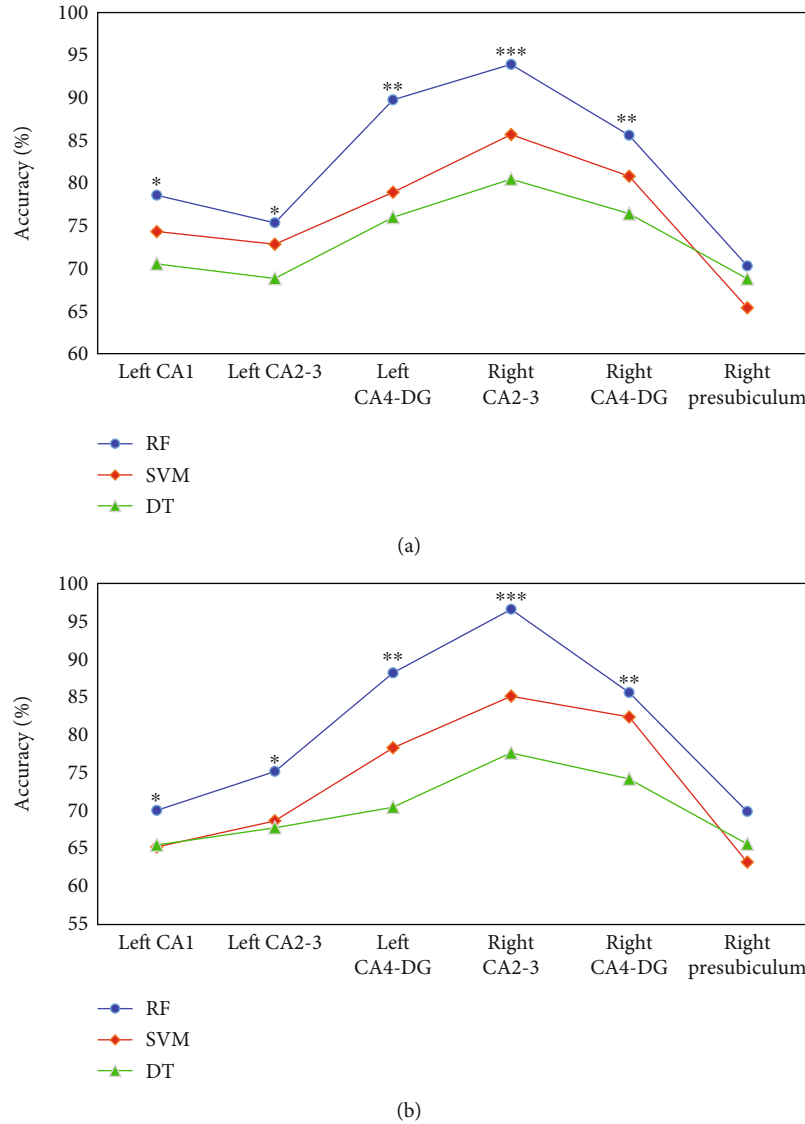
(a)

(b)

FIGURE 12: Comparison of prediction accuracy by different algorithms. (a) showed the comparison of the prediction accuracy of e-MCI by different models based on different hippocampal subregions volume changes, and (b) demonstrated the comparison of the prediction accuracy of I-MCI by different models based on different hippocampal subregions volume changes. ∗ indicated that the comparison with DT models showed statistical differences, $P < 0.05$; ∗∗ showed that the comparison with DT models revealed significant differences, $P < 0.01$; ∗∗∗ suggested that the comparison with DT model demonstrated extremely significant differences, $P < 0.001$.

remarkably higher than that by right presubiculum, and the comparison between two values showed significant differences ($P < 0.01$).

3.5. Comparison of Prediction Performances of Different Machine Learning Algorithms. RF, SVM, and DT models were adopted to predict e-MCI, as shown in Figure 13. The predictive accuracy of e-MCI by RF, SVM, and DT models reached 95.68%, 82.58%, and 78.25%, respectively. The predictive sensitivity of e-MCI by the three models amounted to 88.16%, 68.34%, and 68.52%. The predictive specificity of e-MCI by the three algorithms was 93.05%, 79.81%, and 81.57%. The predictive accuracy, sensitivity, and specificity of e-MCI by RF model were all obviously higher than those by DT model ($P < 0.01$). Besides, the comparison of the predictive accuracy, sensitivity, and specificity of e-MCI by DT

model and by SVM model showed no statistical differences ($P > 0.05$).

4. Discussion

Early AD patients might suffer from memory impairment, depression, sleep arousal disorders, and sexual dysfunction [19, 20]. With MCI developing into AD, the decrease in hippocampal volumes was aggravated [21]. In the research, early AD prediction models were established based on multiple machine learning algorithms, and the prediction performances of different models were compared to seek the optimal prediction model. As an essential data preprocessing step, feature selection can solve dimension disaster issues in real tasks [22]. In the research, MRI image data was preprocessed, and its features were extracted to obtain 16 volume
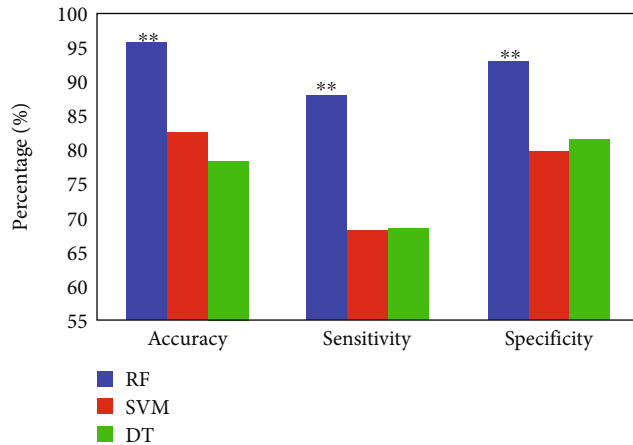
FIGURE 13: Comparison of prediction performances of different machine learning algorithms. ∗∗ indicated that the comparison with DT models showed significant differences, $P < 0.01$.

indexes in hippocampal subregions. After multiple comparisons, by one-way variance analysis and LSD method, pairwise comparisons among groups demonstrated that a total of 8 volume change indexes in hippocampal subregions showed differences. The results of the research showed that the volumes of hippocampi in e-MCI group, including left CA1, left CA2-3, left CA4-DG, left presubiculum, left tail, right CA2-3, right CA4-DG, right presubiculum, and right tail, were all smaller than those in normal group ($P < 0.01$). The volumes of hippocampal subregions corresponding to hippocampi in I-MCI group were all decreased compared with those in e-MCI group ($P < 0.01$). Besides, the corresponding volumes of hippocampal subregions in I-MCI group were significantly reduced compared with those in normal group ($P < 0.001$). The above results demonstrated that the severity of AD disease was positively correlated with the volume changes in hippocampal subregions. In addition, the further analysis of the relevance between hippocampal volumes and LMT-DR in behavioristics revealed that the volumes of subregions, including left CA, left CA2-3, left CA4-DG, right CA2-3, right CA4-DG, and right presubiculum, were all positively correlated with LMT-DR scores significantly. The results showed that the volume reduction in gray matters in CA2-3 and CA4-DG hippocampal subregions could be viewed as the potential indexes of evaluating patient memory impairment. According to some studies, the hippocampal subregions including CA2-3, CA4-DG, and subiculum also could be used to reflect or predict disease progress [23]. Besides, these hippocampal subregions might be more suitable for predicting AD [24]. There were some similarities between the research result and these findings. The left and right CA2-3 and CA4-DG subregions were both related to the development of AD. In addition, it was found out that the region of right presubiculum was also correlated with the impairment of memory function based on deep learning model. The region could be used as a potential AD prediction index.

In the training of data sets, RF could introduce two randomness to avoid overfitting [25]. Besides, it showed advantages in antinoise performance [26]. RF could process continuous and discrete data simultaneously [27] and possess significant advantages in processing high-dimension data, such as simplicity, realizability, and low computing cost [28]. The results of the research revealed that the prediction accuracy of RF training set samples reached 98.33%. Furthermore, the prediction accuracy of e-MCI based on left CA4-DG, right CA2-3, and right CA4-DG volume changes was obviously higher than that by DT models ($P < 0.01$). The prediction accuracy, sensitivity, and specificity of e-MCI by RF models were all remarkably higher than those by DT models ($P < 0.01$). The above results demonstrated that RF showed excellent performance in predicting the risk transformation from e-MCI to I-MCI.

5. Conclusion

Based on machine learning method, early AD was predicted by hippocampal subregion volume changes. The result indicated that RF model showed high predictive performance in predicting the risk of transformation from e-MCI to l-MCI among different machine learning methods. However, there were still some advantages in this research. For example, the research was implemented based only on ADNI public database without further clinical verification of RF performance. In future research, more early AD cases needed to be included, and early AD should be predicted by RF models to verify its correctness. To conclude, effective early AD prediction models were established by hippocampal subregion volume changes, which was based on RF. The establishment of early AD prediction model provided certain references for the diagnosis and treatment of AD patients.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] H. Zhang and Y. Zheng, "β amyloid hypothesis in Alzheimer's disease: pathogenesis, prevention, and management," *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*, vol. 41, no. 5, pp. 702–708, 2019.

[2] Z. Breijyeh and R. Karaman, "Comprehensive review on Alzheimer's disease: causes and treatment," *Molecules*, vol. 25, no. 24, p. 5789, 2020.

[3] S. Jongsiriyanyong and P. Limpawattana, "Mild cognitive impairment in clinical practice: a review article," *American Journal of Alzheimer's Disease and Other Dementias*, vol. 33, no. 8, pp. 500–507, 2018.

[4] R. C. Petersen, O. Lopez, M. J. Armstrong et al., "Practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology," *Neurology*, vol. 90, no. 3, pp. 126–135, 2018.

[5] E. Tönnies and E. Trushina, "Oxidative stress, synaptic dysfunction, and Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 57, no. 4, pp. 1105–1121, 2017.

[6] Z. Jafari, B. E. Kolb, and M. H. Mohajerani, "Neural oscillations and brain stimulation in Alzheimer's disease," *Progress in Neurobiology*, vol. 194, p. 101878, 2020, Epub 2020 Jun 29.

[7] A. Chandra, P. E. Valkimadi, G. Pagano et al., "Applications of amyloid, tau, and neuroinflammation PET imaging to Alzheimer's disease and mild cognitive impairment," *Human Brain Mapping*, vol. 40, no. 18, pp. 5424–5442, 2019.

[8] A. Chandra, G. Dervenoulas, and M. Politis, "Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment," *Journal of Neurology*, vol. 266, no. 6, pp. 1293–1302, 2019.

[9] K. A. Carpenter and X. Huang, "Machine learning-based virtual screening and its applications to Alzheimer's drug discovery: a review," *Current Pharmaceutical Design*, vol. 24, no. 28, pp. 3347–3358, 2018.

[10] S. I. Dimitriadis and D. Liparas, "How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database," *Neural Regeneration Research*, vol. 13, no. 6, pp. 962–970, 2018.

[11] J. Kuang, P. Zhang, T. Cai et al., "Prediction of transition from mild cognitive impairment to Alzheimer's disease based on a logistic regression-artificial neural network-decision tree model," *Geriatrics & Gerontology International*, vol. 21, no. 1, pp. 43–47, 2021, Epub 2020 Dec 1.

[12] J. Chen, N. Ma, G. Hu et al., "rTMS modulates precuneus-hippocampal subregion circuit in patients with subjective cognitive decline," *Aging*, vol. 13, no. 1, pp. 1314–1331, 2021.

[13] A. M. Wessels, S. A. Dowsett, and J. R. Sims, "Detecting treatment group differences in Alzheimer's disease clinical trials: a comparison of Alzheimer's disease assessment scale - cognitive subscale (ADAS-Cog) and the clinical dementia rating - sum of boxes (CDR-SB)," *The Journal of Prevention of Alzheimer's Disease*, vol. 5, no. 1, pp. 15–20, 2018.

[14] M. A. Mintun, A. C. Lo, C. Duggan Evans et al., "Donanemab in early Alzheimer's disease," *The New England Journal of Medicine*, vol. 384, no. 18, pp. 1691–1704, 2021.

[15] A. Shiino, "Artificial intelligence for diagnostic support of neuroimage," *Brain and Nerve*, vol. 71, no. 7, pp. 733–748, 2019.

[16] X. Tang and J. Liu, "Comparing different algorithms for the course of Alzheimer's disease using machine learning," *Annals of Palliative Medicine*, vol. 10, no. 9, pp. 9715–9724, 2021.

[17] S. El-Sappagh, T. Abuhmed, B. Alouffi, R. Sahal, N. Abdelhade, and H. Saleh, "The role of medication data to enhance the prediction of Alzheimer's progression using machine learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8439655, 8 pages, 2021.

[18] A. Sarica, A. Cerasa, and A. Quattrone, "Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review," *Frontiers in Aging Neuroscience*, vol. 9, p. 329, 2017.

[19] J. L. Cummings, G. Tong, and C. Ballard, "Treatment combinations for Alzheimer's disease: current and future pharmacotherapy options," *Journal of Alzheimer's Disease*, vol. 67, no. 3, pp. 779–794, 2019.

[20] I. Koval, A. Bône, M. Louis et al., "*AD Course Map* charts Alzheimer's disease progression," *Scientific Reports*, vol. 11, no. 1, p. 8020, 2021.

[21] Y. Li, A. Haber, C. Preuss et al., "Transfer learning-trained convolutional neural networks identify novel MRI biomarkers of Alzheimer's disease progression," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 13, no. 1, article e12140, 2021.

[22] I. Krashenyi, J. Ramírez, A. Popov, J. M. Górriz, and The Alzheimer's Disease Neuroimaging Initiative, "Fuzzy computer-aided Alzheimer's disease diagnosis based on MRI data," *Current Alzheimer Research*, vol. 13, no. 5, pp. 545–556, 2016.

[23] L. Lazli, M. Boukadoum, O. Ait Mohamed, and O. Ait Mohamed, "Computer-aided diagnosis system of Alzheimer's disease based on multimodal fusion: tissue quantification based on the hybrid fuzzy-genetic-possibilistic model and discriminative classification based on the SVDD model," *Brain Sciences*, vol. 9, no. 10, p. 289, 2019.

[24] X. Hao, Y. Bao, Y. Guo et al., "Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease," *Medical Image Analysis*, vol. 60, p. 101625, 2020.

[25] M. Song, H. Jung, S. Lee, D. Kim, and M. Ahn, "Diagnostic classification and biomarker identification of Alzheimer's disease with random forest algorithm," *Brain Sciences*, vol. 11, no. 4, p. 453, 2021.

[26] Z. Ma, B. Jing, Y. Li et al., "Identifying mild cognitive impairment with random forest by integrating multiple MRI morphological metrics," *Journal of Alzheimer's Disease*, vol. 73, no. 3, pp. 991–1002, 2020.

[27] S. I. Dimitriadis, D. Liparas, M. N. Tsolaki, and Alzheimer's Disease Neuroimaging Initiative, "Random forest feature selection, fusion and ensemble strategy: combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and Alzheimer's disease patients: from the Alzheimer's disease neuroimaging initiative (ADNI) database," *Journal of Neuroscience Methods*, vol. 302, pp. 14–23, 2018.

[28] J. Kim, M. Lee, M. K. Lee et al., "Development of random forest algorithm based prediction model of Alzheimer's disease using neurodegeneration pattern," *Psychiatry Investigation*, vol. 18, no. 1, pp. 69–79, 2021.