



Research on Voxel-Based Features Detection and Analysis of Alzheimer's Disease Using Random Survey Support Vector Machine

Xianglian Meng¹, Yue Wu¹, Wenjie Liu¹, Ying Wang², Zhe Xu¹ and Zhuqing Jiao^{3*}

¹ School of Computer Information and Engineering, Changzhou Institute of Technology, Changzhou, China, ² School of Computer Science and Engineering, Changshu Institute of Technology, Changshu, China, ³ School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China

OPEN ACCESS

Edited by:

Yu-Dong Zhang,
University of Leicester,
United Kingdom

Reviewed by:

Xia-an Bi,
Hunan Normal University, China
Bingrong Xu,
Wuhan University of Technology,
China
Simin Li,
Southern Medical University, China

*Correspondence:

Zhuqing Jiao
jzq@cczu.edu.cn

Received: 17 January 2022

Accepted: 08 February 2022

Published: 28 March 2022

Citation:

Meng X, Wu Y, Liu W, Wang Y, Xu Z and Jiao Z (2022) Research on Voxel-Based Features Detection and Analysis of Alzheimer's Disease Using Random Survey Support Vector Machine. *Front. Neuroinform.* 16:856295. doi: 10.3389/fninf.2022.856295

Alzheimer's disease (AD) is a degenerative disease of the central nervous system characterized by memory and cognitive dysfunction, as well as abnormal changes in behavior and personality. The research focused on how machine learning classified AD became a recent hotspot. In this study, we proposed a novel voxel-based feature detection framework for AD. Specifically, using 649 voxel-based morphometry (VBM) methods obtained from MRI in Alzheimer's Disease Neuroimaging Initiative (ADNI), we proposed a feature detection method according to the Random Survey Support Vector Machines (RS-SVM) and combined the research process based on image-, gene-, and pathway-level analysis for AD prediction. Particularly, we constructed 136, 141, and 113 novel voxel-based features for EMCI (early mild cognitive impairment)-HC (healthy control), LMCI (late mild cognitive impairment)-HC, and AD-HC groups, respectively. We applied linear regression model, least absolute shrinkage and selection operator (Lasso), partial least squares (PLS), SVM, and RS-SVM five methods to test and compare the accuracy of these features in these three groups. The prediction accuracy of the AD-HC group using the RS-SVM method was higher than 90%. In addition, we performed functional analysis of the features to explain the biological significance. The experimental results using five machine learning indicate that the identified features are effective for AD and HC classification, the RS-SVM framework has the best classification accuracy, and our strategy can identify important brain regions for AD.

Keywords: Alzheimer's disease, RS-SVM, voxel-based features, gene-level, pathway-level

INTRODUCTION

Due to the development of medical technology, the world population has grown steadily, and the elderly population has increased rapidly. It is expected that this trend will continue to accelerate in the next few decades, and the occurrence of senile diseases and the social cost of aging are expected to increase. Alzheimer's disease (AD) is a brain disease. It is also a progressive disease, meaning that it will get worse over time. It is believed that AD begins 20 years or more before

the onset of symptoms (Jiao et al., 2020b). The preclinical stage of AD is crucial for identifying early pathophysiological events and developing interventions for disease improvement. Given that changes in synaptic function occur early in the neurodegenerative process, functional MRI (fMRI) is particularly promising for detecting early changes in brain function (Agosta et al., 2017). MRI has aroused great interest in AD-related research due to its complete noninvasiveness, high availability, high spatial resolution, and good contrast between different soft tissues (Moradi et al., 2015).

Mild cognitive impairment (MCI), known as the early stage of AD, was a disease state of cognitive decline between normal elderly and dementia patients. MCI was divided into early mild cognitive impairment (EMCI) and late mild cognitive impairment (LMCI). Studies had pointed out that if MCI patients were not diagnosed early, the probability of developing AD could be as high as 80% after 6 years, and about two-thirds of AD patients were converted through MCI (Barnes and Yaffe, 2011; Lenhart et al., 2021; Vitali et al., 2021). Using linear mixed models, Vonk et al. (2020) analyzed 2,261 individuals with MCI and non-MCI and found that the neurodegeneration was associated with letter fluency and semantic fluency. Wang et al. (2017) introduced the linear regression classification to classify samples and obtained an accuracy of 97.51% (Zhang et al., 2014). Bi et al. (2020a; 2021b) applied the random forest to identify features associated with AD. Another study showed that the Flash Visual Evoked Potential-P2 latency had AD-specific pathological information (Arruda et al., 2020). Sabuncu et al. (2011) calculated the degree of atrophy of hippocampus and cortical areas and found that the specific cortical thinning and the reduction of hippocampal volume were accelerated in early AD. As the classic analysis methods, the machine learning algorithms brought new research sight to AD-specific biomarkers (Zhang et al., 2015a; Ji et al., 2021; Jiao et al., 2021; Wang et al., 2021). Zhuo et al. (2015) applied a group lasso support vector machine to obtain the AD-specific biomarkers. Patel et al. (2019) developed two XGBoost classification models to classify AD and healthy control (HC). Studies have proved that AD was closely related to brain atrophy and that brain atrophy was mainly reflected in the reduction of cortical surface area, thickness, and gray matter volume and, therefore, gray matter volume, cortical surface area, and average thickness contributed to the pathology of AD patients (Gullett et al., 2021; Lorenzo et al., 2021; Piersson et al., 2021; Talwar et al., 2021).

Despite many efforts, it is still challenging to determine effective AD-specific biomarkers for early diagnosis and prediction of disease progression and requires more research (Bi et al., 2020b; Zhang and Shi, 2020). In our study, we proposed a novel analysis framework based on the Random Survey Support Vector Machines (RS-SVM) for the early detection of AD conversion in MCI patients by using advanced machine learning algorithms and combining voxel-based data with standard neuropsychological test results. First, to obtain the voxel sets, we extracted the differences between AD and HC. Then, we applied the RS-SVM to identify important features that classified EMCI, LMCI, AD, and HC well. Subsequently, we applied several classical methods to construct the analysis

frameworks and evaluate the accuracy of these features to classify with EMCI-HC, LMCI-HC, and AD-HC. The experiment results demonstrate that the identified features were effective in classifying AD, the RS-SVM framework performed well, and the identified regions and genes will further our understanding of AD.

MATERIALS AND METHODS

Figure 1 illustrates the framework of a voxel-based three-level analysis for AD. The framework encompasses data processing (A), features extraction (B), RS-SVM construction (C), and the gene-level analysis using effective chi-square statistic (ECS) method (Li et al., 2019) and pathway-level analysis using the resulting genes (Bu et al., 2021) (D). The novation of this framework is to make the full use of voxel-based data.

Imaging Data

In this study, we downloaded and analyzed 1,426 participants with genotyping data and MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). These data include 353 HCs, 273 EMCI, 504 LMCI, and 296 with AD. The characteristics of these participants, including average age and years of education, are shown in **Table 1**.

Random Survey Support Vector Machines-Based Machine Learning Method

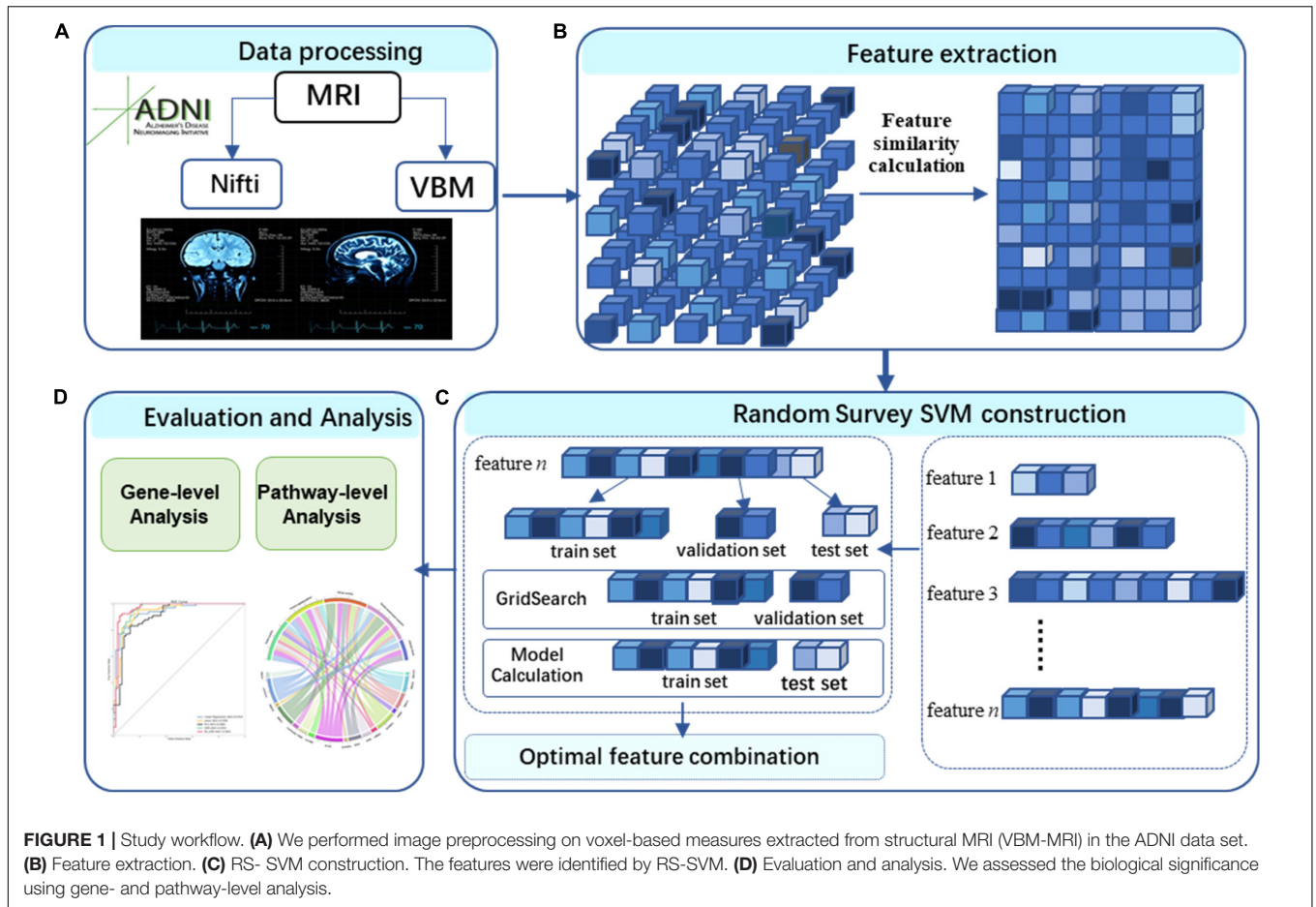
Data Processing

MRI scans, using voxel-based morphometry (VBM), were aligned and normalized to a T1-weighted template image and the Montreal Neurological Institute (MNI) space, respectively. The gray matter density (GMD) maps were segmented, extracted, and smoothed with an 8-mm FWHM (full width at the half maximum) kernel. The Automatic Anatomical Labeling (AAL) atlas was employed to define the regions of interest (ROIs) and their coordinates (whole brain) (Tzourio-Mazoyer et al., 2002). We then down-sampled the resulting maps to a dimension of $61 \times 73 \times 61$ to reduce the data size for subsequent analysis in EMCI-HC, LMCI-HC, and AD-HC groups.

To extract the differences within the three groups, i.e., A, B, and C, we first performed the weighted process of the two sets of images separately and saved them as matrices M and N (i.e., M for AD group and N for HC group). Then, let (vm'_i, vn'_i) represents the vector of two voxels ($vm'_i \in M, vn'_i \in N$), and we obtained a vector $V = \{(vm'_1, vn'_1), (vm'_2, vn'_2), \dots, (vm'_k, vn'_k)\}$ ($k = 271, 633$). Since the voxels $(vm'_i = vn'_i)$ in the two groups were meaningless for our research, we deleted these voxels and obtained 64,411 sets of different voxels. We used V' to denote the voxel set.

Feature Extraction

The data sets of features were still too many for our final binary classification in EMCI-HC, LMCI-HC, and AD-HC groups. Therefore, we estimated the number of features by calculating the



similarity between two rows (vm'_i, vn'_i) and (vm'_j, vn'_j) in V' . The similarity of the voxels is given in the following equation:

$$\rho_i = \sqrt{(vm'_i - vm'_j)^2 + (vn'_i - vn'_j)^2}, (vm'_i, vn'_i), (vm'_j, vn'_j) \in V' \quad (1)$$

where vm'_i and vm'_j ($i, j = 1, 2, \dots, 64411$) are the values of AD group. vn'_i and vn'_j ($i, j = 1, 2, \dots, 64411$) are the values of HC groups. ρ_i is the similarity between (vm'_i, vn'_i) and (vm'_j, vn'_j) .

For the convenience of calculation, we divided the V' into ten groups and obtained 55 sets of similarity matrices. On this basis, we defined the number of minimal ρ_i as C_{\min} and the number of maximal ρ_i as C_{\max} . Due to the value of C_{\min} and C_{\max} ($C_{\min} = 132, C_{\max} = 21$), we defined that the number of features should be in $[C_{\max}, C_{\min}]$. Then, we extracted 64,411 features of all subjects from the original MR images to form a $649 \times 64,411$ matrix as the initial data set.

Random Survey Support Vector Machines Construction

To extract the important feature, we proposed a single-kernel SVM model based on random survey. The goal of random survey was the establishment of a random experimental data set. Since the initial data set X was a two-dimensional matrix of $649 \times 64,411$, we selected l column from the X

randomly and constructed a single randomized experimental data set X' ($l \in [C_{\max}, C_{\min}]$). At the same time, the set of columns corresponding to each column l in each extraction was $R = \{r_1, r_2, \dots, r_l\}$, which denoted the index of brain loci coordinates. The indices are extracted as follows:

$$\begin{aligned} u(k) &= \{u_1, u_2, \dots, u_k\}, k = 64411 \\ u'(k) &= \{u_p, u_q, \dots, u_r\}, r, p, q \in [1, 64411] \\ R &= \{r_1 = u_p, r_2 = u_q, \dots, r_l = u_g\}, g \in [1, 64411] \end{aligned} \quad (2)$$

After random extraction, we defined the training set:validation set:test set as 6:2:2. The training set was used as an input for training first. The validation set was applied to obtain the optimal hyperparameters and replaced the initial parameters. The remaining 20% was introduced as the test set to calculate the accuracy of the tuned model and to evaluate whether the obtained feature set $R = \{r_1, r_2, \dots, r_l\}$ can be used as the final feature set.

In the classification process of SVM, the input data $X' = \{X'_1, X'_2, \dots, X'_N, \dots, X'_M\}$ and the learning objective $y = \{y_1, y_2, \dots, y_N, \dots, y_M\}$ were given, where N was the number of EMCI, LMCI, and AD samples, respectively, and M was the number of HC. The learning objectives were binary variables $y = \{-1, 1\}$, where -1 represents EMCI, LMCI, and AD, respectively, and 1 represents HC in the three groups. The

TABLE 1 | Participant characteristics.

Subjects	HC	EMCI	LMCI	AD	p
Number	353	273	504	296	–
Gender (M/F)	187/166	153/120	309/195	166/130	<0.001
Age (mean ± sd)	72.2 ± 7.6	71.3 ± 7.1	74.0 ± 7.6	75.1 ± 5.5	<0.001
Edu (mean ± sd)	16.1 ± 2.7	16.1 ± 2.6	16.0 ± 2.9	16.3 ± 2.6	<0.001

HC, healthy control; EMCI, early mild cognitive impairment; LMCI, late mild cognitive impairment; AD, Alzheimer's disease; Edu, education.

feature set of the input data was regarded as the hyperplane D in decision boundaries to separate the learning targets by positive and negative classes, making the distance ε_i between any sample and plane ≥ 1 . The hyperplane and the plane distance are defined as follows:

$$\begin{aligned} D : w^T X' + b &= 0 \\ \varepsilon_i &= y_i(w^T X' + b), \varepsilon_i \geq 1 \end{aligned} \quad (3)$$

where w denotes the normal vector of the hyperplane and b denotes the intercept of the hyperplane. The decision boundary satisfying this condition actually constructed two parallel hyperplanes D_1, D_2 as interval boundaries to classify the samples (Eq. 4).

$$\begin{aligned} w^T X'_i + b &\geq +1 \Rightarrow y_i = +1 \\ w^T X'_i + b &\leq -1 \Rightarrow y_i = -1 \end{aligned} \quad (4)$$

Based on Eq. 4, it could be derived that all samples above the upper interval boundary were positive and those below the lower interval boundary were negative. The distance between the two interval boundaries $d = \frac{2}{\|w\|}$ was defined as the margin. Since our experimental data X' was selected randomly, there was hyperboloid in the feature set to separate positive and negative classes. Using nonlinear functions, the nonlinear separable problems from the original feature set were mapped to a higher dimensional Hilbert space H . The hyperplane, using as the decision boundary, is defined as follows:

$$w^T \varphi(X') + b = 0 \quad (5)$$

where $\varphi : X' \mapsto H$ denotes the mapping function. Since the mapping function was complex, it was difficult to calculate the inner product. Therefore, the inner product of the mapping function was defined as kernel functions $k(X'_1, X'_2) = \varphi(X'_1)^T \varphi(X'_2)$ to avoid the explicit operation.

Parameter Determination

We used the original linear kernel function of the support vector machine first, and the penalty factor C and the kernel parameter gamma were set as default values ($C = 1$ and gamma = 0.5). Then, we applied the training data set and labels to train the model. Subsequently, the hyperparameters were optimized by grid search. The SVM could be transformed into an equivalent quadratic convex optimization problem to solve using the following equation:

$$\begin{aligned} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^M \varepsilon_i \\ y_i(w^T X'_i + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0 \end{aligned} \quad (6)$$

Evaluation Metrics

In this article, the samples were positive and negative, and the results classified had the following cases:

True positive (TP): the positive sample was predicted as a positive sample.

True negative (TN): the negative sample was predicted as a negative sample.

False positive (FP): the negative sample was predicted as a positive sample.

False negative (FN): the positive sample was predicted as a negative sample.

Let P denotes the positive sample and N denotes the negative sample. We then obtained the following equation:

$$\begin{aligned} TP + FN &= P \\ FP + FN &= N \end{aligned} \quad (7)$$

The evaluation metrics used in our research are as follows:

- Accuracy. Accuracy was the number of correctly classified samples divided by the total number of samples (Eq. 8).

$$ACC = \frac{TP + TN}{P + N} \quad (8)$$

- Precision. Precision was the proportion of the samples that were actually positive (or negative) divided by samples classified as positive (or negative) (Eq. 9).

$$precision = \frac{TP}{TP + FP} \quad (9)$$

- Recall. Recall was the measure of coverage (Eq. 10).

$$recall = \frac{TP}{TP + FN} \quad (10)$$

- Comprehensive evaluation indicators (F-Measure). Accuracy and sensitivity sometimes needed to be considered together as given in the following equation:

$$F = \frac{(\alpha^2 + 1) * P * R}{\alpha^2(P + R)} \quad (11)$$

When $\alpha=1$, Eq. 11 is transformed into the following equation as follows:

$$F = \frac{2 * P * R}{P + R} \quad (12)$$

Model Comparison

We used the test set to evaluate the classification ability of 5 machine learning methods, including linear regression model, least absolute shrinkage and selection operator (Lasso) model, partial least squares (PLS) model, SVM model, and RS-SVM model. First, the initial default parameters were applied to each model to train and calculate the evaluation metrics. Then, the grid search algorithm was used to optimize the hyperparameters of the five models. Finally, the hyperparameters were introduced in each model to recalculate the evaluation metrics. The results were used to evaluate the pros and cons of the five models.

Since the RS-SVM model in this article was optimized based on the traditional SVM model, the other three evaluation models were described in detail in this section.

Linear regression model was a statistical analysis method that used regression analysis in mathematical statistics to determine the quantitative relationship between the interdependence of two or more variables.

Given a data set $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$, we learned that a linear model from this data set will reflect the correspondence between x_i and y_i as accurately as possible. The linear regression model, which was a function of linear combination of attributes x , could be expressed as follows:

$$f(x) = w_1x_1 + w_2x_2 + \dots + w_ix_i + b = W^T X + b \quad (13)$$

where $W = \{w_1, w_2, \dots, w_i\}$ is column vector, indicating the weight of the corresponding attribute in the prediction result. Eq. 13 was represented as the following equation:

$$f(x_i) = w_ix_i + b, f(x_i) \approx y_i \quad (14)$$

Then it was to find a model such that $\forall i \in [1, m]$ has $f(x_i)$ as close to y_i . Therefore, the sum of the squares of the difference between the predicted value and the real value of each sample is minimized and thus gives the following equation:

$$(w^*, b^*) = \operatorname{argmin}_{(w,b)} \sum_{i=1}^m (f(x_i) - y_i)^2 \quad (15)$$

where (w^*, b^*) is the optimal parameter, and the minimum value of (w, b) is taken for the above equation.

The Lasso model was a compression estimation method with the idea of reducing the variable set (decreasing order). By constructing a penalty function, it could compress the coefficients of variables and made some regression coefficients become 0, so as to achieve the purpose of variable selection.

Given n data samples $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ where each $x_i \in R^d$ was a d -dimensional vector, i.e., each observed data point was composed of the values of d variables, and each $y_i \in R$ was a real value. What we had to do was to find a map $f: R^d \rightarrow R$ that minimized the sum of squared errors based on the observed data points. The optimization objective is given as follows:

$$\beta^* = \operatorname{argmin}_{\beta} \frac{1}{n} \sum_{i=0}^n \left((y_i - \bar{y}) - \beta^T (x_i - \bar{x}) \right)^2 \quad (16)$$

where $\beta \in R^d$ is the optimized coefficient.

If Eq. 16 is expressed in matrix form, denoted by $X = [x_1; x_2; \dots; x_n]^T$, where each data point x_i was regarded as a column vector, then $X \in R^{n \times d}$, denoted as $y = (y_1, y_2, \dots, y_n)^T$, then the optimization objective in matrix form is given as follows:

$$\beta^* = \operatorname{argmin}_{\beta} \frac{1}{n} \|y - X\beta\|_2^2 \quad (17)$$

Lasso added the L1 regularization term (see Eq. 18) to make the model avoid over-fitting.

$$\|\beta\|_1 = \sum_{j=1}^d |\beta_j|; \quad 1 \ll j \ll d \quad (18)$$

Then, the optimization objective function of Lasso is expressed as the following equation:

$$\beta^* = \operatorname{argmin}_{\beta} \frac{1}{n} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \quad (19)$$

PLS model was a many-to-many linear regression modeling method, i.e., there are multiple independent variables and multiple dependent variables. It found the best functional fit for a set of data by minimizing the sum of squared errors.

The general multivariate underlying model of PLS is given by the following equations:

$$X = TP^T + E$$

$$Y = UQ^T + F \quad (20)$$

where X is a $n \times m$ prediction matrix, Y is a $n \times p$ response matrix; T and U are $n \times l$ matrices and both of them are the projections of X and Y in the higher dimensional space; P and Q are the orthogonal loading matrices of $m \times l$ and $p \times l$, respectively, and the matrices E and F are error terms, normally distributed random variables subject to independent and identical distributions. Decompose X and Y to maximize the covariance between T and U .

Gene-Level Analyses

We analyzed the voxel-based features using gene-level analysis. First, quality control (QC) was performed using the PLINK version 1.9 software¹ (Purcell et al., 2007). We performed genome-wide association studies (GWASs) using the image data and genetic data in whole brain using the linear regression in PLINK. Age, gender, education, and the top 10 principal components from population stratification analysis were included as covariates. A total of 5,574,300 single-nucleotide polymorphisms (SNPs) were obtained by QC. We applied ECS method (Li et al., 2019) to assign SNPs' to autosomal genes. Then the significant genes was obtained by Bonferroni correction (family-wise error rate p -value < 0.05).

Pathway-Level Analyses

Using the resulting genes, we performed the pathway analysis to assess the biological significance of these features (Bu et al., 2021). KOBAS-I (Bu et al., 2021) pathway analysis tool (KOBAS; bioinfo.org) and the Kyoto Encyclopedia of Genes and Genomes database were applied to pathway analysis of the identified genes ($P < 0.001$).

¹<https://www.cog-genomics.org/plink/1.9/>

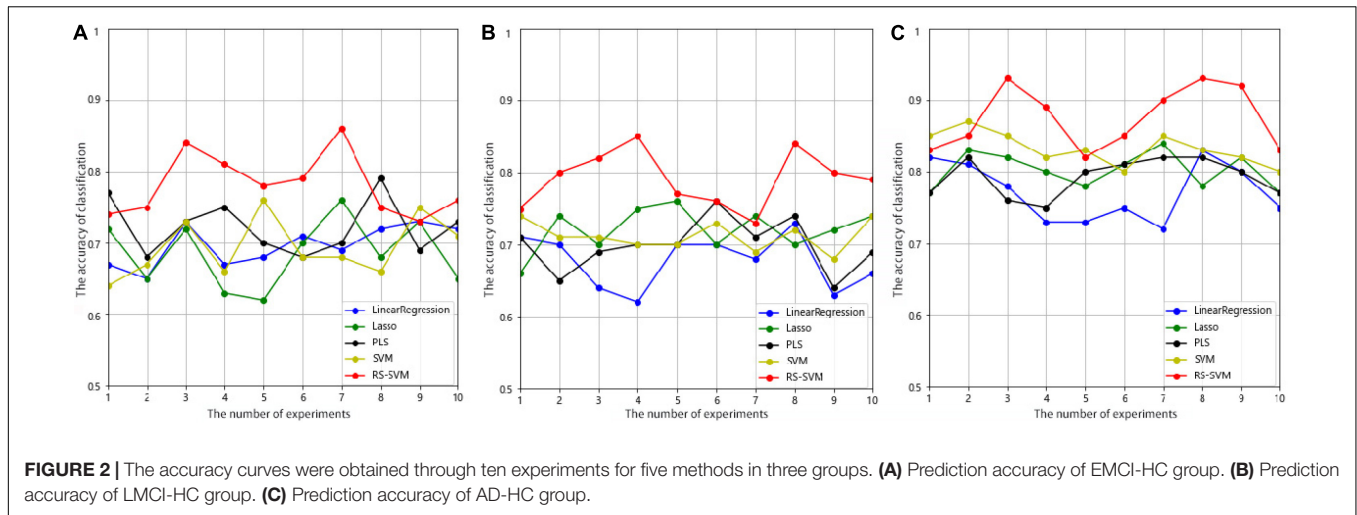


TABLE 2 | Test results of different models.

Group	Model	Validation set				Test set			
		Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
EMCI-HC	Linear regression	0.67	0.67	0.67	0.67	0.73	0.73	0.73	0.73
	Lasso	0.79	0.79	0.79	0.79	0.80	0.80	0.80	0.80
	PLS	0.8	0.8	0.8	0.8	0.82	0.81	0.81	0.81
	SVM	0.73	0.73	0.73	0.73	0.76	0.76	0.76	0.76
	RS-SVM	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
LMCI-HC	Linear regression	0.62	0.62	0.62	0.62	0.78	0.78	0.77	0.77
	Lasso	0.80	0.80	0.80	0.80	0.81	0.81	0.81	0.81
	PLS	0.65	0.64	0.65	0.64	0.66	0.65	0.66	0.65
	SVM	0.73	0.73	0.73	0.73	0.74	0.74	0.74	0.74
	RS-SVM	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
AD-HC	Linear regression	0.85	0.85	0.84	0.84	0.84	0.85	0.84	0.84
	Lasso	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
	PLS	0.91	0.92	0.91	0.91	0.91	0.92	0.91	0.91
	SVM	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
	RS-SVM	0.91	0.91	0.91	0.91	0.93	0.93	0.93	0.93

Bold fonts represented the model and experimental results in this paper.

RESULTS

In recent studies, machine learning was used to detect the subjects and brain regions of AD (Zhang et al., 2016) and the brain functional statuses of EMCI (Jiao et al., 2020a) and to identify AD and MCI (Zhang et al., 2015b; Wang et al., 2016). In this work, we applied a novel feature extraction method and SVM to obtain the features classified EMCI, LMCI, AD, and HC.

Comparison of the Five Methods

We employed the test set to evaluate the classification capability of the five methods, and the experiments were repeated 10 times with the selected parameter combination in each method. As shown in **Figure 2**, the RS-SVM model has the best prediction accuracy. The AD-HC group had more than 90% prediction accuracy, while the other four methods all peaked below 90%. The prediction accuracy of both the EMCI-HC and LMCI-HC

groups exceeded 85%, while the peak values of the other four methods were all below 80%. The curves in **Figure 2** also showed that RS-SVM had good stability. In ten replicates, the difference in accuracy was less than 10%. These analyses demonstrated the satisfactory classification ability and stability of the RS-SVM model.

Machine learning had been gradually maturing and has been applied to the classification and prediction of AD. We applied the validation set to obtain the optimal parameters and the test set to evaluate the classification capability of the five methods. The evaluation metrics of the five methods implemented in EMCI-HC, LMCI-HC, and AD-HC were shown in **Table 2**. As shown in **Table 2**, the RS-SVM has the best accuracy, precision, recall, and F-measure. Only the values of RS-SVM increase with the optimal parameters, and the values of other models are stable. In the AD-HC, EMCI-HC, and LMCI-HC groups, the F-measure of RS-SVM in the

TABLE 3 | Top 10 conditionally significant genes were obtained. Chr represents Chromosome; Gene represents the gene name; CorrectedP represents *P*-value generated by Bonferroni correction.

No.	Chr	Gene	CorrectedP
1	8	CSMD1	1.74556E-36
2	16	RBFOX1	3.18755E-23
3	16	CDH13	1.07119E-20
4	9	PTPRD	1.92988E-19
5	8	DLGAP2	3.74049E-17
6	11	CNTN5	4.81385E-16
7	7	MAGI2	5.93057E-16
8	20	MACROD2	1.50704E-14
9	16	WVOX	1.64798E-14
10	3	CNTN4	1.87567E-13

TABLE 4 | Top 10 significant pathways.

NO.	Pathways	Corrected P-value	Gene
1	Insulin secretion	1.01E-06	PLCB1, PRKCB, PRKCA, CREB5, RYR2, CHRM3, KCNMA1, RAPGEF4, CACNA1C
2	Oxytocin signaling pathway	4.80E-06	PLCB1, PRKAG2, PRKCA, CACNB2, RYR3, RYR2, PRKCB, CACNA1C, ITPR2, CACNA2D3
3	Salivary secretion	7.70E-06	PLCB1, PRKCA, RYR3, PRKCB, CHRM3, KCNMA1, PRKG1, ITPR2
4	Vascular smooth muscle contraction	7.94E-06	PLCB1, CACNA1C, PRKCH, PRKCA, PRKCB, PRKCE, KCNMA1, PRKG1, ITPR2
5	Calcium signaling pathway	1.48E-05	PLCB1, PRKCB, ERBB4, PRKCA, RYR3, RYR2, CHRM3, CACNA1C, ITPR2, PDE1A
6	Glutamatergic synapse	2.08E-05	PLCB1, CACNA1C, PRKCA, GRIK2, PRKCB, DLGAP1, ITPR2, GRM7
7	Morphine addiction	5.00E-05	PRKCA, PDE1A, PRKCB, PDE3A, GABRB3, PDE4D, PDE10A
8	Circadian entrainment	5.56E-05	PLCB1, PRKCB, PRKCA, RYR3, RYR2, CACNA1C, PRKG1
9	Pancreatic secretion	5.56E-05	PLCB1, PRKCB, PRKCA, RYR2, CHRM3, KCNMA1, ITPR2
10	Aldosterone synthesis and secretion	5.56E-05	PLCB1, CACNA1C, PRKCA, CREB5, PRKCB, PRKCE, ITPR2

validation set were 0.91, 0.86, and 0.85 from high to low. In the AD-HC, EMCI-HC, and LMCI-HC groups, the F-measure of RS-SVM in the test set were 0.93, 0.86, and 0.85 from high to low. This also indicated that the RS-SVM model was scalable, and SVM combined with other schemes have better performance than single SVM. In addition, since the same features were applied to the five models, good results were obtained for all five models (all above 0.8). This proved that the identified features were excellent in the classification of AD and HC and were meaningful for the identification of AD. Therefore, we performed GWAS of these features to analyze their biological significance.

Results of Gene-Level Genome-Wide Association Study

We performed the conditional gene-based association scans on whole genome. All genes with conditional association *P*-values passing Bonferroni correction for family-wise error rate at 0.05 were extracted. We performed the gene-based association analysis by using *P*-values of 113 novel voxel-based features for identifying susceptibility genes of AD. There are 242 genes (corrected *P*-value < 0.001) associated with AD. These top 10 conditionally significant genes are shown in **Table 3**. Studies have shown that CSMD1 (SNP: rs34464519, CorrectedP: 1.74556E-36) was related to AD (Stepanov et al., 2014; Li et al., 2020; Bi et al., 2021a). RBFOX1 (SNP: rs55642412, CorrectedP: 3.18755E-23) has been found to play a role in neuronal development (Raghavan et al., 2020). PTPRD (SNP: rs62538998, CorrectedP: 1.92988E-19) has been confirmed to be related to AD and MCI in previous studies (Huang et al., 2021). DLGAP2 (SNP: rs72507619, CorrectedP: 3.74049E-17) was found to be predominantly expressed in the brain and associated with a wide variety of neurological disorders (Linthorst et al., 2020). WVOX gene has been reported to be a potential mechanism that may be involved in the pathogenesis of AD, focusing on the cell death signaling pathway in neurons (Teng et al., 2012).

Results of Pathway-Level Genome-Wide Association Study

Detecting pathways may provide useful information about the pathogenic molecular mechanism underlying AD. In our work, 70 enriched pathways were identified. The top 10 significant pathways are shown in **Table 4**. Impaired insulin secretion was associated with higher risk of any dementia and cognitive impairment (Rönnemaa et al., 2008). Oxytocin signaling pathway was neuroprotective to many neurological disorders, such as AD (Almansoub et al., 2020). Vascular smooth muscle contraction was associated with the development of neurodegeneration in AD (Hald et al., 2016).

DISCUSSION

We proposed a voxel-based three-level analysis framework for AD that extracted the voxel-based ROI, which included the whole brain from MRI. Although the voxel-based research could solve the limitations of the research method based on the ROI, it was more easily affected by the characteristics of high-dimensional data. Feature selection in RS-SVM model solved the dimensional disaster caused by too many attributes. In this work, we identified 136, 141, and 113 MRI features for EMCI-HC, LMCI-HC, and AD-HC groups, respectively.

We performed RS-SVM model to identify important brain regions such as hippocampus, amygdala, angular gyrus, and calcarine sulcus for AD. The hippocampus was located in the midlimbic system of the brain and had an important impact on memory and cognitive function. Many studies had shown that abnormalities in hippocampal volume and function were closely linked to AD. Although many patients had not shown symptoms

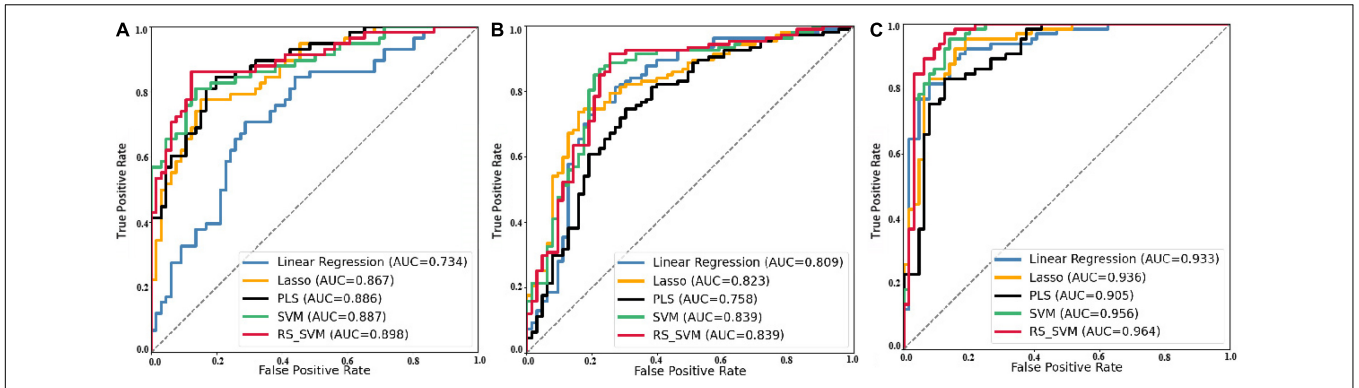


FIGURE 3 | ROC curve of five classification methods for three groups. (A) Prediction ROC of EMCI-HC group. (B) Prediction ROC of LMCI-HC group. (C) Prediction ROC of AD-HC group.

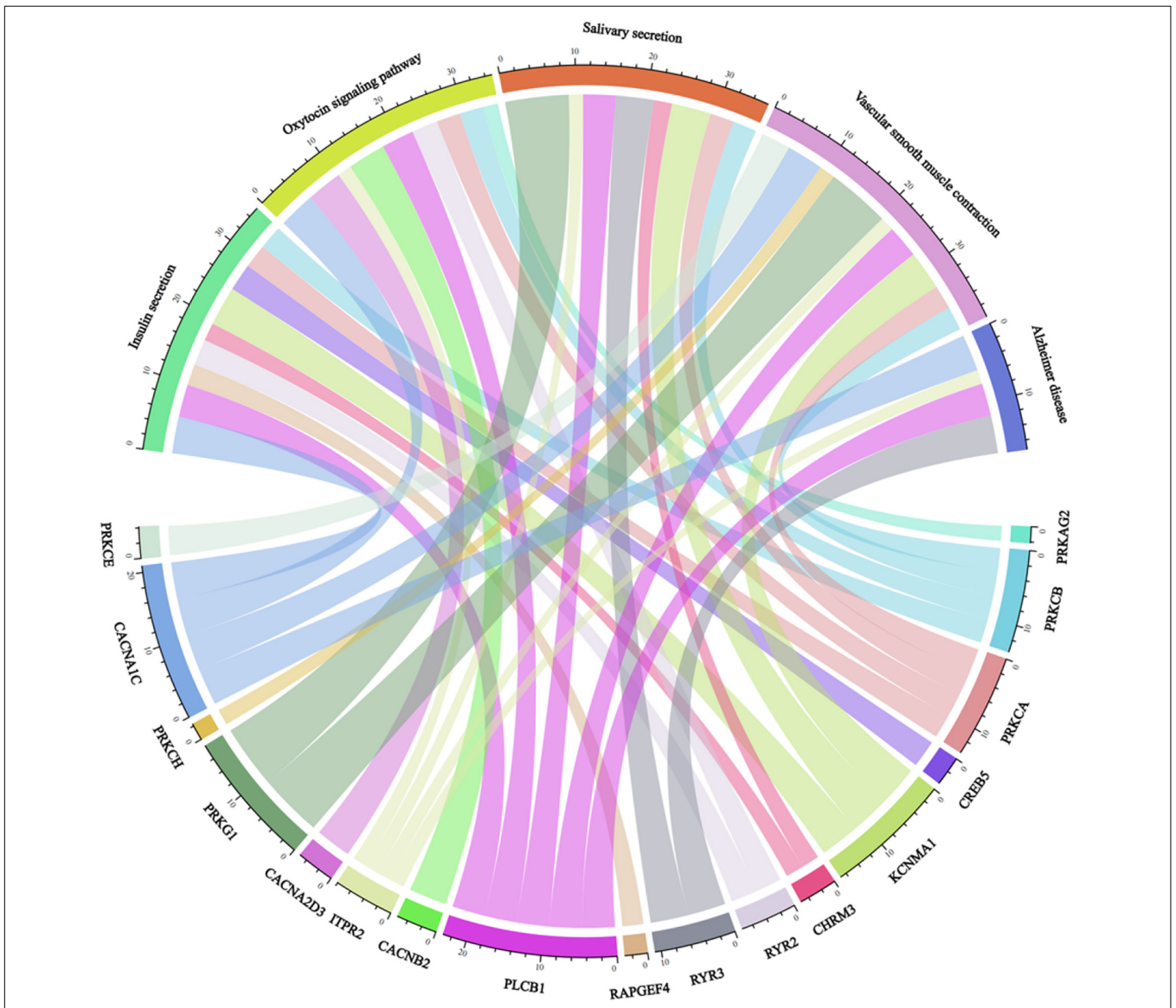


FIGURE 4 | An image showing the relation of genes and pathways.

of AD during MCI period, the temporal lobe located in the inner part of the brain had obvious symptoms. The atrophy of the hippocampus was the most obvious (Adachi et al., 2021; Guo et al., 2021; Zhao et al., 2021). Located at the bottom of the brain, the amygdala was shaped like an almond and was the center of the brain to control and manage emotions. Pathological protein changed in the amygdala affected the occurrence, development, and evolution of AD and led to nerve damage or tissue cell aging (Biechele et al., 2021; Ni, 2021). The angular gyrus was the portion surrounding the end of the superior temporal sulcus in the temporal lobe. Gray matter atrophy gradually spread from the basal ganglia to the angular gyrus, temporal regions, and eventually to the subcortical-cortical network as neurological disease progresses. The angular gyrus was key AD-risk region (Brown dyke et al., 2021; Zenke et al., 2021). The calcarine sulcus was located posterior to the medial surface of the hemisphere. Less activation of the bilateral anterior calcarine sulcus was associated with better delayed recall in amnesic MCI patients (Chen et al., 2021; Lejko et al., 2022).

In our work, the EMCI-HC, LMCI-HC, and AD-HC groups were classified in order to increase the specificity and detail of classification and to enable early diagnosis of the disease. In this study, five classification models of linear regression, Lasso, PLS, SVM, and RS-SVM were applied. The prediction accuracy of AD-HC was the highest. EMCI and LMCI represented the middle stage of disease progression. Therefore, the prediction accuracies of EMCI-HC and LMCI-HC were lower than that of AD-HC.

The machine learning was applied to classify AD and HC in previous studies (Zhang and Wang, 2015; Wang et al., 2018; Zhang et al., 2018; Ji et al., 2021; Jiao et al., 2022). For further classification validation, we plot the ROC curves (Figure 3) of five machine learning classification methods for EMCI-HC, LMCI-HC, and AD-HC groups. The AUCs of the RS-SVM for EMCI-HC, LMCI-HC, and AD-HC were 0.898, 0.839, and 0.964, respectively. As can be seen in Table 2, the proposed classification method, based on RS-SVM, consistently outperforms other methods (i.e., linear regression, Lasso, PLS, and SVM). This proved that the framework based on RS-SVM was optimal compared with the other four methods.

With the replication of statistical gene-level GWAS, we obtained *PLCB1*, *PTPRD*, *RYR3*, *CACNA1C*, *KCNMA1*, and *PRKCA* genes. The *PLCB1* gene was implicated in AD pathogenesis (Bastrup et al., 2021). The *PTPRD* gene has been found on the short arm of human chromosome 9. Recent report identified *PTPRD* association with the extent of neurofibrillary pathology in AD brain specimens (Chibnik et al., 2018; Uhl and Martinez, 2019), and also *RYR3* gene for the RYR which functions to release the stored endoplasmic reticulum calcium ions (Ca^{2+}) to increase intracellular Ca^{2+} concentration. The studies demonstrate that altered levels of intracellular Ca^{2+} affect neurodegeneration (Gong et al., 2018; Nilipour et al., 2018). It has been shown that the expression of *CACNA1C* inhibits the hyperphosphorylation of Tau protein (Jiang et al., 2018).

Some pathways include insulin secretion, oxytocin signaling pathway, salivary secretion, vascular smooth muscle contraction, and AD closely related to genes (Figure 4).

In summary, our proposed framework based on RS-SVM performed well in features constructed, and the framework had good classification performance for EMCI-HC, LMCI-HC, and AD-HC groups. In particular, AD-HC was the best in terms of classification accuracy. Several pathogenic genes and abnormal subregions identified using this framework are related to AD. Therefore, we speculate that the remaining genes identified could be regarded as the candidate genes for AD. The discoveries in this study provide new candidate genes for AD, and the constructed features can be regarded as a new indicator to distinguish AD from HC.

DATA AVAILABILITY STATEMENT

Rs-fMRI data were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>). Application for access to the ADNI data can be submitted at <http://adni.loni.usc.edu/data-samples/access-data/>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI). We applied the access from ADNI. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

XM, YWu, WL, and ZJ led and supervised research. XM, YWu, and WL designed the research and wrote the article. XM, YWu, and ZX performed features extraction and selection and random survey support vector machines. WL, YWu, and ZX performed data preprocessing and quality control. XM and WL did gene- and pathway-level analysis. All authors reviewed, commented, edited, and approved the manuscript.

FUNDING

This research was funded by the National Natural Science Foundation of China (61901063 and 51877013), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (19YJCZH120 and 21YJAZH091), National Statistical Science Research Project (2020LY074), and the Science and Technology Plan Project of Changzhou (CE20205042 and CJ20210155), Jiangsu Provincial Key Research and Development Program (BE2021636), and Key Lab of Intelligent Optimization and Information Processing Minnan Normal University (ZNYH202007). This work was also sponsored by Qing Lan Project of Jiangsu Province. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01

AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

ACKNOWLEDGMENTS

ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer

Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

- Adachi, Y., Ota, K., Minami, I., Yamada, T., and Watanabe, T. (2021). Lower insulin secretion is associated with hippocampal and parahippocampal gyrus atrophy in elderly patients with type 2 diabetes mellitus. *J. Diab. Investig.* 12, 1908–1913. doi: 10.1111/jdi.13554
- Agosta, F., Galantucci, S., and Filippi, M. (2017). Advanced magnetic resonance imaging of neurodegenerative diseases. *Neurol. Sci.* 38, 41–51. doi: 10.1007/s10072-016-2764-x
- Almansoub, H., Tang, H., Wu, Y., Wang, D.-Q., Abdoul Razak, M., Mahamane Salissou, M., et al. (2020). Oxytocin Alleviates MPTP-Induced Neurotoxicity in Mice by Targeting MicroRNA-26a/Death-Associated Protein Kinase 1 Pathway. *J. Alzheimer's Dis.* 74, 1–19. doi: 10.3233/JAD-191091
- Arruda, J. E., McInnis, M. C., and Steele, J. (2020). The flash visual evoked potential-P2 and the detection of amnesic mild cognitive impairment: A review of empirical literature. *Int. J. Psychophysiol.* 155, 162–167. doi: 10.1016/j.jpsycho.2020.05.012
- Barnes, D. E., and Yaffe, K. (2011). The projected effect of risk factor reduction on Alzheimer's disease prevalence. *Lancet Neurol.* 10, 819–828. doi: 10.1016/S1474-4422(11)70072-2
- Bastrup, J., Hansen, K. H., Poulsen, T. B. G., Kastaniegaard, K., Asuni, A. A., Christensen, S., et al. (2021). Anti-A β Antibody Aducanumab Regulates the Proteome of Senile Plaques and Closely Surrounding Tissue in a Transgenic Mouse Model of Alzheimer's Disease. *J. Alzheimer's Dis.* 79, 249–265. doi: 10.3233/JAD-200715
- Bi, X. A., Hu, X., Wu, H., and Wang, Y. (2020a). Multimodal Data Analysis of Alzheimer's Disease Based on Clustering Evolutionary Random Forest. *IEEE J. Biomed. Health Inform.* 24, 2973–2983. doi: 10.1109/JBHI.2020.2973324
- Bi, X.-A., Liu, Y., Xie, Y., Hu, X., and Jiang, Q. (2020b). Morbigenous brain region and gene detection with a genetically evolved random neural network cluster approach in late mild cognitive impairment. *Bioinformatics* 36, 2561–2568. doi: 10.1093/bioinformatics/btz967
- Bi, X.-A., Hu, X., Xie, Y., and Wu, H. (2021a). A novel CERNNE approach for predicting Parkinson's Disease-associated genes and brain regions based on multimodal imaging genetics data. *Med. Image Analysis* 67:101830. doi: 10.1016/j.media.2020.101830
- Bi, X. A., Zhou, W., Li, L., and Xing, Z. (2021b). Detecting Risk Gene and Pathogenic Brain Region in EMCI Using a Novel GERF Algorithm Based on Brain Imaging and Genetic Data. *IEEE J. Biomed. Health Inform.* 25, 3019–3028. doi: 10.1109/JBHI.2021.3067798
- Biechele, G., Wind, K., Blume, T., Sacher, C., Beyer, L., Eckenweber, F., et al. (2021). Microglial activation in the right amygdala-entorhinal-hippocampal complex is associated with preserved spatial learning in App(NL-G-F) mice. *Neuroimage* 230:117707. doi: 10.1016/j.neuroimage.2020.117707
- Browndyke, J. N., Wright, M. C., Yang, R., Syed, A., Park, J., Hall, A., et al. (2021). Perioperative neurocognitive and functional neuroimaging trajectories in older APOE4 carriers compared with non-carriers: secondary analysis of a prospective cohort study. *Br. J. Anaesthesia* 127, 917–928. doi: 10.1016/j.bja.2021.08.012
- Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., et al. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* 49, W317–W325. doi: 10.1093/nar/gkab447
- Chen, N., Liu, G., Guo, M., Li, Y., Yao, Z., and Hu, B. (2021). Calcarine as a bridge between brain function and structure in irritable bowel syndrome: A multiplex network analysis. *J. Gastroenterol. Hepatol.* 36, 2408–2415. doi: 10.1111/jgh.15382
- Chibnik, L. B., White, C. C., Mukherjee, S., Raj, T., Yu, L., Larson, E. B., et al. (2018). Susceptibility to neurofibrillary tangles: role of the PTPRD locus and limited pleiotropy with other neuropathologies. *Mol. Psychiatry* 23, 1521–1529. doi: 10.1038/mp.2017.20
- Gong, S., Su, B. B., Tovar, H., Mao, C., Gonzalez, V., Liu, Y., et al. (2018). Polymorphisms Within RYR3 Gene Are Associated With Risk and Age at Onset of Hypertension, Diabetes, and Alzheimer's Disease. *Am. J. Hypertens.* 31, 818–826. doi: 10.1093/ajh/hpy046
- Gullett, J. M., Albizu, A., Fang, R., Loewenstein, D. A., Duara, R., Rosselli, M., et al. (2021). Baseline Neuroimaging Predicts Decline to Dementia From Amnesic Mild Cognitive Impairment. *Front. Aging Neurosci.* 13:758298. doi: 10.3389/fnagi.2021.758298
- Guo, X. Y., Chang, Y., Kim, Y., Rhee, H. Y., Cho, A. R., Park, S., et al. (2021). Development and evaluation of a T1 standard brain template for Alzheimer disease. *Quant. Imag. Med. Surg.* 11, 2224–2244. doi: 10.21037/qims-20-710
- Hald, E. S., Timm, C. D., and Alford, P. W. (2016). Amyloid Beta Influences Vascular Smooth Muscle Contractility and Mechanoadaptation. *J. Biomechan. Engin.* 11:138. doi: 10.1115/1.4034560
- Huang, Y., Sun, X., Jiang, H., Yu, S., Robins, C., Armstrong, M. J., et al. (2021). A machine learning approach to brain epigenetic analysis reveals kinases associated with Alzheimer's disease. *Nat. Commun.* 12:4472. doi: 10.1038/s41467-021-24710-8
- Ji, Y., Zhang, Y., Shi, H., Jiao, Z., Wang, S.-H., and Wang, C. (2021). Constructing Dynamic Brain Functional Networks via Hyper-Graph Manifold Regularization for Mild Cognitive Impairment Classification. *Front. Neurosci.* 15:669345. doi: 10.3389/fnins.2021.669345
- Jiang, Y., Xu, B., Chen, J., Sui, Y., Ren, L., Li, J., et al. (2018). Micro-RNA-137 Inhibits Tau Hyperphosphorylation in Alzheimer's Disease and Targets the CACNA1C Gene in Transgenic Mice and Human Neuroblastoma SH-SY5Y Cells. *Med. Sci. Monit.* 24, 5635–5644. doi: 10.12659/MSM.908765
- Jiao, Z., Chen, S., Shi, H., and Xu, J. (2022). Multi-Modal Feature Selection with Feature Correlation and Feature Structure Fusion for MCI and AD Classification. *Brain Sci.* 12:80. doi: 10.3390/brainsci12010080
- Jiao, Z., Gao, P., Ji, Y., and Shi, H. (2021). Integration and Segregation of Dynamic Functional Connectivity States for Mild Cognitive Impairment Revealed by Graph Theory Indicators. *Contrast Med. Mol. Imag.* 2021:6890024. doi: 10.1155/2021/6890024

- Jiao, Z., Ji, Y., Gao, P., and Wang, S. (2020a). Extraction and analysis of brain functional statuses for early mild cognitive impairment using variational auto-encoder*. *J. Ambient Intellig. Human. Comput.* 1–12.
- Jiao, Z., Ji, Y., Jiao, T., and Wang, S. (2020b). Extracting Sub-Networks from Brain Functional Network Using Graph Regularized Nonnegative Matrix Factorization. *Comput. Model. Engin. Sci.* 123, 845–871. doi: 10.32604/cmesci.2020.08999
- Lejko, N., Tumati, S., Opmeer, E. M., Marsman, J. C., Reesink, F. E., De Deyn, P. P., et al. (2022). Planning in amnesic mild cognitive impairment: an fMRI study. *Exp. Gerontol.* 159:111673. doi: 10.1016/j.exger.2021.111673
- Lenhart, L., Seiler, S., Pirpamer, L., Goebel, G., Potrusil, T., Wagner, M., et al. (2021). Anatomically Standardized Detection of MRI Atrophy Patterns in Early-Stage Alzheimer's Disease. *Brain Sci.* 11:11. doi: 10.3390/brainsci11111491
- Li, M., Jiang, L., Mak, T. S. H., Kwan, J. S. H., Xue, C., Chen, P., et al. (2019). A powerful conditional gene-based association approach implicated functionally important genes for schizophrenia. *Bioinformatics* 35, 628–635. doi: 10.1093/bioinformatics/bty682
- Li, M., Lyu, J. H., Zhang, Y., Gao, M. L., Li, R., Mao, P. X., et al. (2020). Efficacy of Group Reminiscence Therapy on Cognition, Depression, Neuropsychiatric Symptoms, and Activities of Daily Living for Patients With Alzheimer Disease. *J. Geriatr. Psychiatry Neurol.* 33, 272–281. doi: 10.1177/0891988719882099
- Linthorst, J., Meert, W., Hestand, M. S., Korlach, J., Vermeesch, J. R., Reinders, M. J. T., et al. (2020). Extreme enrichment of VNTR-associated polymorphism in human subtelomeres: genes with most VNTRs are predominantly expressed in the brain. *Transl. Psychiatry* 10:369. doi: 10.1038/s41398-020-01060-5
- Lorenzo, T., Operto, G., Falcon, C., Cuenca, A., Forcano, L., Piera, I., et al. (2021). Cognitive function and neuroimaging correlates in a subjective cognitive decline population within a multimodal intervention supplemented with green tea extract (PENSA Study): Preliminary results of a voxel-based morphometry study. *Alzheimers Dement* 17:e055894. doi: 10.1002/alz.055894
- Moradi, E., Pepe, A., Gaser, C., Huttunen, H., and Tohka, J. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *NeuroImage* 104, 398–412. doi: 10.1016/j.neuroimage.2014.10.002
- Ni, R. (2021). Positron Emission Tomography in Animal Models of Alzheimer's Disease Amyloidosis: Translational Implications. *Pharmaceuticals* 14:11. doi: 10.3390/ph14111179
- Nilipour, Y., Nafissi, S., Tjast, A. E., Ravenscroft, G., Hossein Nejad, Nedai, H., et al. (2018). Ryanodine receptor type 3 (RyR3) as a novel gene associated with a myopathy with nemaline bodies. *Eur. J. Neurol.* 25, 841–847. doi: 10.1111/ene.13607
- Patel, H., Iniesta, R., Stahl, D., Dobson, R., and Newhouse, S. (2019). Working Towards a Blood-Derived Gene Expression Biomarker Specific for Alzheimer's Disease. *J. Alzheimers Dis.* 74, 545–561. doi: 10.3233/JAD-191163
- Pierson, A. D., Ibrahim, B., Suppiah, S., Mohamad, M., Hassan, H. A., Omar, N. F., et al. (2021). Multiparametric MRI for the improved diagnostic accuracy of Alzheimer's disease and mild cognitive impairment: Research protocol of a case-control study design. *PLoS One* 16:e0252883. doi: 10.1371/journal.pone.0252883
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Raghavan, N. S., Dumitrescu, L., Mormino, E., Mahoney, E. R., Lee, A. J., Gao, Y., et al. (2020). Association Between Common Variants in RBFOX1, an RNA-Binding Protein, and Brain Amyloidosis in Early and Preclinical Alzheimer Disease. *JAMA Neurol.* 77, 1288–1298. doi: 10.1001/jamaneurol.2020.1760
- Rönnemaa, E., Zethelius, B., Sundelöf, J., Sundström, J., Degerman-Gunnarsson, M., Berne, C., et al. (2008). Impaired insulin secretion increases the risk of Alzheimer disease. *Neurology* 71:1065. doi: 10.1212/01.wnl.0000310646.32212.3a
- Sabuncu, M. R., Desikan, R. S., Sepulcre, J., Yeo, B. T. T., Liu, H., Schmansky, N. J., et al. (2011). The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch. Neurol.* 68, 1040–1048. doi: 10.1001/archneurol.2011.167
- Stepanov, V. A., Bocharova, A. V., Marusin, A. V., Zhukova, N. G., Alifirova, V. M., and Zhukova, I. A. (2014). Replicative association analysis of genetic markers of cognitive traits with Alzheimer's disease in the Russian population. *Mole. Biol.* 48, 835–844. doi: 10.1134/S0026893314060168
- Talwar, P., Kushwaha, S., Chaturvedi, M., and Mahajan, V. (2021). Systematic Review of Different Neuroimaging Correlates in Mild Cognitive Impairment and Alzheimer's Disease. *Clin. Neuroradiol.* 31, 953–967. doi: 10.1007/s00062-021-01057-7
- Teng, C.-C., Yang, Y.-T., Chen, Y.-C., Kuo, Y.-M., and Sze, C.-I. (2012). Role of WWOX/WOX1 in Alzheimer's disease pathology and in cell death signaling. *Front. Biosci.* 4:1951–1965. doi: 10.2741/e516
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Uhl, G. R., and Martinez, M. J. (2019). PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. *Anna. N. Y. Acad. Sci.* 1451, 112–129. doi: 10.1111/nyas.14002
- Vitali, F., Branigan, G. L., and Brinton, R. D. (2021). Preventing Alzheimer's disease within reach by 2025: Targeted-risk-AD-prevention (TRAP) strategy. *Alzheimers Dement* 7:e12190. doi: 10.1002/trc2.12190
- Vonk, J. M. J., Bouteloup, V., Mangin, J.-F., Dubois, B., Blanc, F., Gabelle, A., et al. (2020). Semantic loss marks early Alzheimer's disease-related neurodegeneration in older adults without dementia. *Alzheimer's Dementia* 12:e12066. doi: 10.1002/dad2.12066
- Wang, S., Zhang, Y., Liu, G., Phillips, P., and Yuan, T.-F. (2016). Detection of Alzheimer's Disease by Three-Dimensional Displacement Field Estimation in Structural Magnetic Resonance Imaging. *J. Alzheimer's Dis.* 50, 233–248. doi: 10.3233/JAD-150848
- Wang, S., Zhou, Q., Yang, M., and Zhang, Y.-D. (2021). ADVIAN: Alzheimer's Disease VGG-Inspired Attention Network Based on Convolutional Block Attention Module and Multiple Way Data Augmentation. *Front. Aging Neurosci.* 13:687456. doi: 10.3389/fnagi.2021.687456
- Wang, S.-H., Du, S., Zhang, Y., Phillips, P., Wu, L.-N., Chen, X.-Q., et al. (2017). Alzheimer's Disease Detection by Pseudo Zernike Moment and Linear Regression Classification. *CNS Neurol. Disor. Drug Targets* 16, 11–15. doi: 10.2174/187152731566616111123024
- Wang, S.-H., Zhang, Y., Li, Y.-J., Jia, W.-J., Liu, F.-Y., Yang, M.-M., et al. (2018). Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimed. Tools Appl.* 77, 10393–10417. doi: 10.1007/s11042-016-4222-4
- Zenke, J. K., Rahman, S., Guo, Q., Leung, A. W. S., and Gomaa, N. A. (2021). Central Processing in Tinnitus: fMRI Study Outlining Patterns of Activation Using an Auditory Discrimination Task in Normal Versus Tinnitus Patients. *Otol. Neurotol.* 42, e1170–e1180. doi: 10.1097/mao.00000000000003194
- Zhang, T., and Shi, M. (2020). Multi-modal neuroimaging feature fusion for diagnosis of Alzheimer's disease. *J. Neurosci. Methods* 341:108795. doi: 10.1016/j.jneumeth.2020.108795
- Zhang, Y., Dong, Z., Phillips, P., Wang, S., Ji, G., Yang, J., et al. (2015a). Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. *Front. Comput. Neurosci.* 9:66. doi: 10.3389/fncom.2015.00066
- Zhang, Y., and Wang, S. (2015). Detection of Alzheimer's disease by displacement field and machine learning. *PeerJ* 3:e1251. doi: 10.7717/peerj.1251
- Zhang, Y., Wang, S., Phillips, P., Dong, Z., Ji, G., and Yang, J. (2015b). Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomed. Signal Proc. Control* 21, 58–73. doi: 10.1016/j.bspc.2015.05.014
- Zhang, Y., Wang, S., Phillips, P., Yang, J., and Yuan, T.-F. (2016). Three-Dimensional Eigenbrain for the Detection of Subjects and Brain Regions Related with Alzheimer's Disease. *J. Alzheimer's Dis.* 50, 1163–1179. doi: 10.3233/JAD-150988

- Zhang, Y., Wang, S., Sui, Y., Yang, M., Liu, B., Cheng, H., et al. (2018). Multivariate Approach for Alzheimer's Disease Detection Using Stationary Wavelet Entropy and Predator-Prey Particle Swarm Optimization. *J. Alzheimer's Dis.* 65, 855–869. doi: 10.3233/JAD-170069
- Zhang, Y.-D., Wang, S., and Dong, Z. (2014). Classification of Alzheimer Disease Based on Structural Magnetic Resonance Imaging by Kernel Support Vector Machine Decision Tree. *Progr. Electromagnet. Res.* 144, 185–191. doi: 10.2528/PIER13121310
- Zhao, Z., Zhang, L., Wen, Q., Luo, W., Zheng, W., Liu, T., et al. (2021). The effect of beta-amyloid and tau protein aggregations on magnetic susceptibility of anterior hippocampal laminae in Alzheimer's diseases. *Neuroimage* 244:118584. doi: 10.1016/j.neuroimage.2021.118584
- Zhuo, S., Yong, F., Boudewijn, P. F. L. M. D., Martijn, and van de, G. (eds) (2015). Detection of Alzheimer's disease using group lasso SVM-based region selection. *ProcSPIE* 2015:941414.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Meng, Wu, Liu, Wang, Xu and Jiao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.