Evaluation of Feature Selection for Alzheimer's Disease Diagnosis

Feng Gu[1,2†], Songhua Ma[3,4†], Xiude Wang[1,2], Jian Zhao[5], Ying Yu[5] and Xinjian Song[6,7]*

[1] *Department of Medical Image, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China,*
[2] *Department of Medical Image, The Second People's Hospital of Nantong, Nantong, China,* [3] *Department of Neurology, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China,* [4] *Department of Neurology, The Second People's Hospital of Nantong, Nantong, China,* [5] *Nantong Center for Disease Control and Prevention, Nantong, China,* [6] *Department of Rehabilitation Medicine, Affiliated Nantong Rehabilitation Hospital of Nantong University, Nantong, China,* [7] *Department of Rehabilitation Medicine, The Second People's Hospital of Nantong, Nantong, China*

Accurate recognition of patients with Alzheimer's disease (AD) or mild cognitive impairment (MCI) is important for the subsequent treatment and rehabilitation. Recently, with the fast development of artificial intelligence (AI), AI-assisted diagnosis has been widely used. Feature selection as a key component is very important in AI-assisted diagnosis. So far, many feature selection methods have been developed. However, few studies consider the stability of a feature selection method. Therefore, in this study, we introduce a frequency-based criterion to evaluate the stability of feature selection and design a pipeline to select feature selection methods considering both stability and discriminability. There are two main contributions of this study: (1) It designs a bootstrap sampling-based workflow to simulate real-world scenario of feature selection. (2) It develops a decision graph to determine the optimal combination of supervised and unsupervised feature selection both considering feature stability and discriminability. Experimental results on the ADNI dataset have demonstrated the feasibility of our method.

Keywords: artificial intelligence, Alzheimer's disease, feature selection, stability, discriminability

INTRODUCTION

Alzheimer's disease (AD) (Xiao-Cong et al., 2018; Hou et al., 2020; Mishra and Li, 2020; Subasi, 2020; He et al., 2022) is a degenerative disease of the central nervous system, which is clinically manifested as progressive memory impairment, cognitive dysfunction, language dysfunction, and personality change, etc. AD has a serious impact on the lives of patients, but also brings a heavy economic burden to patients' families. At present, the research progress of AD is slow, and the disease factors cannot be accurately determined. It is usually found at an advanced stage, and even treatment will not produce a better therapeutic effect. Therefore, the early diagnosis of AD is very critical, which can effectively inhibit the development of the disease, and even avoid the occurrence of clinical symptoms by taking timely treatment. Mild Cognitive Impairment (MCI) is considered as an intermediate state between health and AD. In patients with MCI, the probability of progressing to AD is about 10–15% (He et al., 2022). Therefore, if patients with MCI can be effectively identified and actively intervened, it is of great significance for the control of AD.

With the rapid development of artificial intelligence (Jiang et al., 2020; Xia et al., 2020; Zhang et al., 2020a,b, 2021a,b, 2022), intelligent models are widely used in MCI or AD recognition.

Kloppel et al. (2008) input gray matter features of brain images of AD patients into linear support vector machines (SVM), so as to apply the trained SVM to clinical studies. Ashburner and Friston (2000) applied morphometric methods to the diagnosis of AD, which spatially normalized high-resolution images of all subjects into the same stereotactic space. Then, gray matter was separated from the spatially normalized images and data smoothing was performed on them. Voxel parameter test statistics were performed on the two groups of smoothed gray images to improve the uneven intensity of the brain artifact images. Hinrichs et al. (2009) also proposed an AD recognition framework based on the smoothness of three-dimensional image coordinate space. It directly integrates the spatial relations of voxels into the learning framework and does not require image preprocessing information of other modes, thus automatically classifying subjects according to structural or functional imaging features. In addition, MCI was associated with changes in cortical morphology, such as cortical thickness, sulcus depth, surface area, gray matter volume, and mean curvature in different brain regions. These features have been shown to have a specific neuropathological and genetic basis. However, most methods have focused on univariate prediction models, and cortical features are usually isolated. Therefore, Li et al. (2014) used a multivariate approach to study the abnormalities of multiple cortical features in patients with mild cognitive impairment, and identified subtle patterns of changes in cortical anatomical structure through a classification model. Liu et al. (2013) used non-linear global data structure to map multivariable MRI data such as regional brain volume and cortical thickness into a low-dimensional local linear space through local linear embedding method, and trained a disease classifier by embedding brain features to predict whether MCI would be transformed into AD in the future. Möller et al. (2016) took the voxel values extracted from the voxel data as the original feature data, and proposed a feature selection method to apply to the original feature vector, so as to reduce the dimension of the original feature vector to a low-dimensional space and carry out the next classification task. From the above-mentioned studies, we can summarize the general process of MCI/AD recognition based on intelligent model, as shown in **Figure 1**. From **Figure 1**, it can be found that the general process of MCI/AD recognition contains four components, preprocessing, feature extraction, feature selection, and prediction. Preprocessing aims to process the original images including registration, standardizing and smoothing. Feature extraction aims to extract original features from the images after preprocessing. Feature selection aims to select discriminant features from the original feature set. Prediction aims to build a classification model to recognize MCI or AD patients. In the phase of prediction, based on the selected features, a prediction model is established for MCI/AD recognition.

From **Figure 1**, it can be found that feature selection is a key phase in the process of MCI/AD recognition. The goal of feature selection is to select discriminant features with low relevance between each other and high relevance to the outcome. In recent 2 years, some excellent feature selection work has emerged in the field of medical images. For example, Demir and Akbulut (2022) proposed a new residual- convolutional neural network to extract deep features from MRI images. Mainenti et al. (2022) proposed a radiomics-based pipeline to enhance MRI-based risk stratification in patients with endometrial cancer. Although previous studies have achieved great success in feature selection, feature discriminability is often the first important factor and feature stability is always omitted. In this study, first of all, feature stability, variance, and pairwise correlation were analyzed. Then, the least absolute shrinkage and selection operator (LASSO) and recursive feature elimination (RFE) were employed to search for the optimal feature set (Mainenti et al., 2022).

In this study, we focus on feature selection because few studies consider both the stability and performance of feature selection so far, which are two key factors for the classification phase. The main contributions cover two aspects. The first one is that we introduce a frequency-based criterion to evaluate the stability of a feature selection method. The second is that we propose a bootstrap-based flow chart and a decision graph to select the best combination of supervised and unsupervised feature selection methods. The following sections are organized as follows. Section "Data and Methods" presents the data we used and the methods we proposed. Section "Results" reports the experimental results, section "Discussion" discusses the experimental results and the last section concludes the whole study.

DATA AND METHODS

Data

In this study, we select 103 patients with MRI and PET from the Alzheimer's Disease Neuroimaging Initiative (ADNI) as our datasets. ADNI is a 5-year public partnership sponsored by several institutes, companies, and non-profit organizations (Zhang et al., 2021b). Owning to the original images cannot be directly used for our study, we set up a data preprocessing pipeline, which contains three main steps. Firstly, each subject in ADNI contains 96 PET images. Statistical parametric mapping (SPM) (Muzik et al., 2000) is used to fuse these PET images to construct a 3-D one which has brain spatial information and the feature information between tissue structures are also retained. In addition, motion correction is performed due to head motion. Secondly, the MRI image and PET image of each subject are registered, and affinely aligned. In the third step, the average template data generated is used to spatially normalize all PET images to the standard MNI space. PET images are also smoothed (8 mm Gaussian) to avoid the influences caused by noises. The AAL (automated anatomical atlas) (Rolls et al., 2020) which is available as a toolbox[1] for SPM is used as a template to extract original features from PET images. Based on AAL, the brain is segmented into 116 regions, and we select 90 regions from the cerebrum for feature extraction. To be specific, firstly, the PET images are resampled to the same size as the AAL template so that each region is in correspondence spatially. The size of AAL template is $61 \times 73 \times 61$. Then we extract average intensity
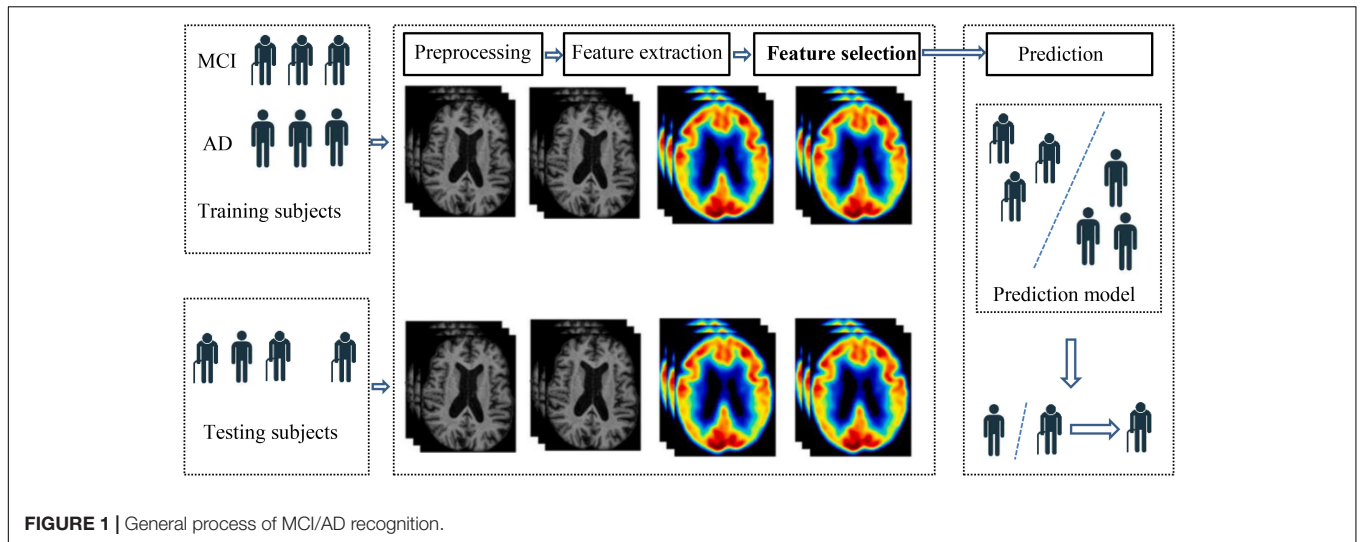
---

[1] http://www.gin.cnrs.fr/AAL

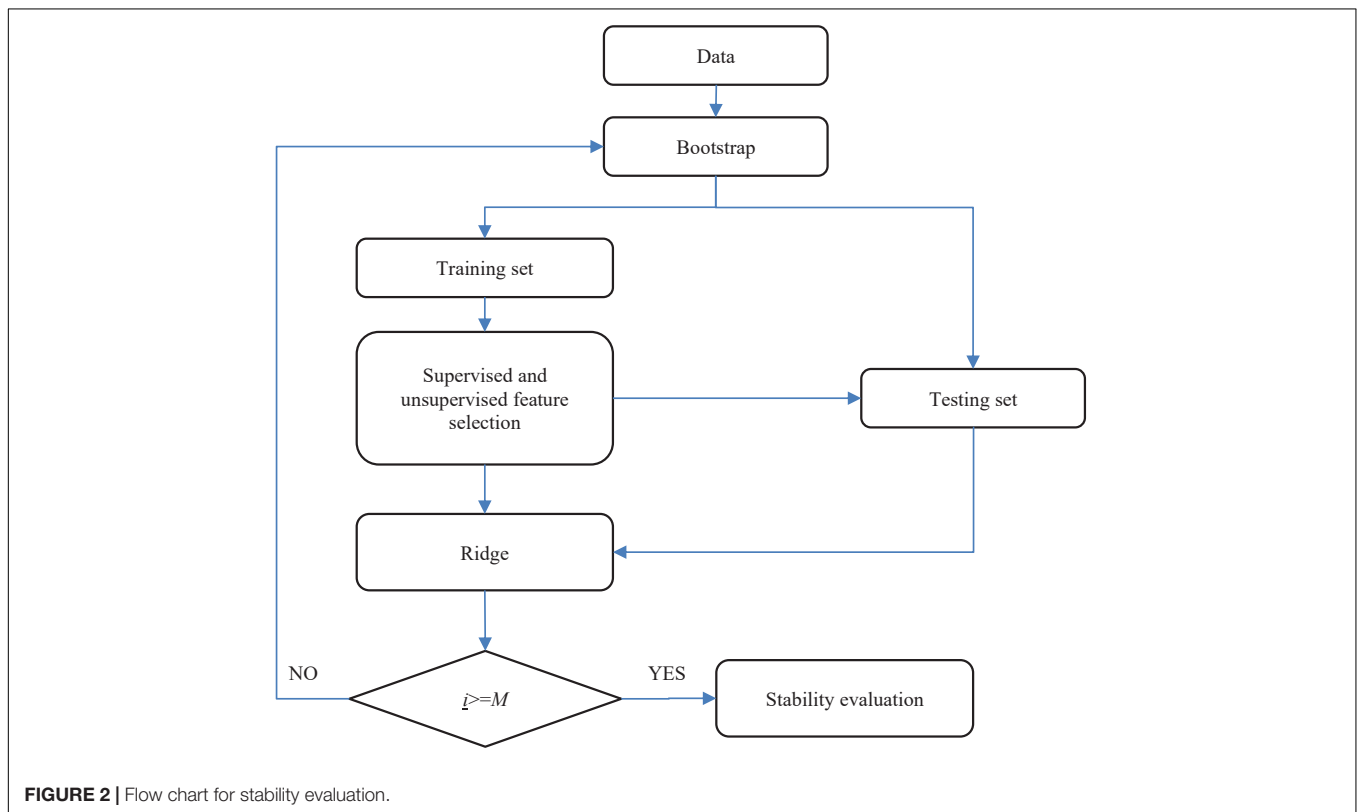**FIGURE 1 |** General process of MCI/AD recognition.

FIGURE 2 | Flow chart for stability evaluation.

values from all regions of PET images as original features for our proposed classification model.

Methods
Stability Evaluation Metrics
In this study, we use a frequency-based criterion to measure the stability of a feature select method (Nogueira et al., 2017). For clarity, suppose we have a feature selection method Φ and a d-dimensional dataset X. The feature selection method is performed on the d-dimensional dataset X to select discriminant features. The feature selection process is repeated M times by a bootstrap strategy. Then we can define a binary matrix Z, as shown in (1) to indicate the feature selection results of M tries,

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1d} \\ z_{21} & z_{22} & \dots & z_{2d} \\ \dots & \dots & \dots & \dots \\ z_{M1} & z_{M2} & \dots & z_{Md} \end{bmatrix} \quad (1)$$

In Z, each row represents one try of feature selection. In each row, $z_{ij} = 1(i = 1, 2, ..., M, j = 1, 2, ..., d)$ represents that the j-th feature is selected in the i-th try; otherwise, the j-th feature is not selected. Based on the binary matrix Z, the stability of feature selection method Φ in terms of the frequency-based criterion can be defined as:

$$Stability(Z) = 1 - \frac{\frac{1}{d}\sum_{f=1}^{d}\left[\frac{M}{M-1}(\frac{1}{M}\sum_{i=1}^{M}z_{if})(1 - \frac{1}{M}\sum_{i=1}^{M}z_{if})\right]}{\frac{\frac{1}{M}\sum_{i=1}^{M}\sum_{f=1}^{d}z_{if}}{d}(1 - \frac{\frac{1}{M}\sum_{i=1}^{M}\sum_{f=1}^{d}z_{if}}{d})} \quad (2)$$

From (2), we can see that $Stability(Z)$ ranges from 0 to 1, the greater the value, the better the stability.

Stability Evaluation Workflow

In this study, we use a supervised feature selection method to reduce features irrelative to the outcome, and an unsupervised feature selection method to reduce redundant features. To evaluate the stability of feature selection, a bootstrap sampling-based flow chart is established, which is shown in **Figure 2**. Firstly, the AD dataset is split into the training set (70%) and the testing set (30%) by bootstrap sampling. Then supervised and unsupervised feature selection is performed on the training set to select discriminant features. The testing set is updated with the selected features. Finally, a Ridge regression model is trained based on the selected features. The bootstrap sampling is repeated M times so that the matrix **Z** in (1) can be obtained. Based on **Z**, we can use (2) to evaluate the stability of the supervised and unsupervised feature selection methods we used.

Decision Graph for Feature Selection

In Li et al. (2017), a feature selection package was shared which contains 33 different kinds of supervised and unsupervised feature selection methods. In this study, we aim to choose a best supervised and unsupervised combination from this package for AD diagnosis. First of all, we set up an initial exclusion criterion to select a part of supervised and unsupervised feature select methods from the package provided by Li et al. (2017). The exclusion criterion states: (1) if prediction performance in terms of AUC of a feature selection method is lower than 0.5, the method is excluded. (2) If the running time of one try of a feature selection method is more than 30 min, the method is excluded. These exclusion criteria are defined for two reasons. The first is that if the prediction performance of the feature selection method is lower than 0.5, it indicates that the prediction performance of the method is close to the randomness level. Second, if the running time of a feature selection method exceeds 30 min, it will exceed the normal tolerance range when the training set size is not large. With the exclusion criterion, we finally select F score (denoted as S1:), T Score (denoted as S2), ReliefF (denoted as S3), and Fish Score (denoted as S4) as supervised feature selection methods, and Lap_score (denoted as U1), spectral feature selection (SPEC, denoted as U2), Monte Carlo feature selection (MCFS, denoted as U3), non-negative discriminative feature selection (NDFS, denoted as U4), unsupervised discriminative feature selection (UDFS, denoted as U5), and Person_score (denoted as U6) as unsupervised feature

TABLE 1 | All combinations of supervised and unsupervised feature selection methods.

Combination name	Name of supervised method	Name of unsupervised method
S1U1	F score	Lap_score
S1U2		SPEC
S1U3		MCFS
S1U4		NDFS
S1U5		UDFS
S1U6		Person score
S2U1	T score	Lap_score
S2U2		SPEC
S2U3		MCFS
S2U4		NDFS
S2U5		UDFS
S2U6		Person score
S3U1	ReliefF	Lap_score
S3U2		SPEC
S3U3		MCFS
S3U4		NDFS
S3U5		UDFS
S3U6		Person score
S4U1	Fish score	Lap_score
S4U2		SPEC
S4U3		MCFS
S4U4		NDFS
S4U5		UDFS
S4U6		Person score

selection methods. Therefore, we have 24 combinations, i.e., S1U1, S1U2,..., S4U6, as shown in **Table 1**. Secondly, as we stated before that both performance and stability are important for Alzheimer's disease diagnosis.

Based on **Figure 2**, we can generate the matrix **Z.** Thus, we can use (2) to evaluate the stability of the supervised and unsupervised feature selection methods we used. Therefore, we design a decision graph, as shown in **Figure 3**, to determine the best combination of the supervised and unsupervised feature selection methods.

RESULTS

The decision graph of all combinations for MRI features is shown in **Figure 4**. It is observed that the combination S2U6 wins the best in terms of *AUC*Stability*, which means that the combination of T Score (supervised feature selection method) and Person Score (unsupervised feature selection method) performs better than other combinations in terms of both AUC and stability. Therefore, the supervised feature selection method *T Score* and the unsupervised feature selection method *Person Score* will be selected as the feature selection methods for modeling.

The decision graph of all combinations for PET features is shown in **Figure 5**. Similar to **Figure 4**, it is observed that the combination S1U1 and S4U6 wins the best. Therefore, the combination *F score + Lap score* or the combination *Fish Score + Person Score* will be selected for the following phase of modeling.

From **Figures 4**, **5**, it can be found that this is no combinations that always perform best. Our method is case-dependent, which means that it provides decision support for users.
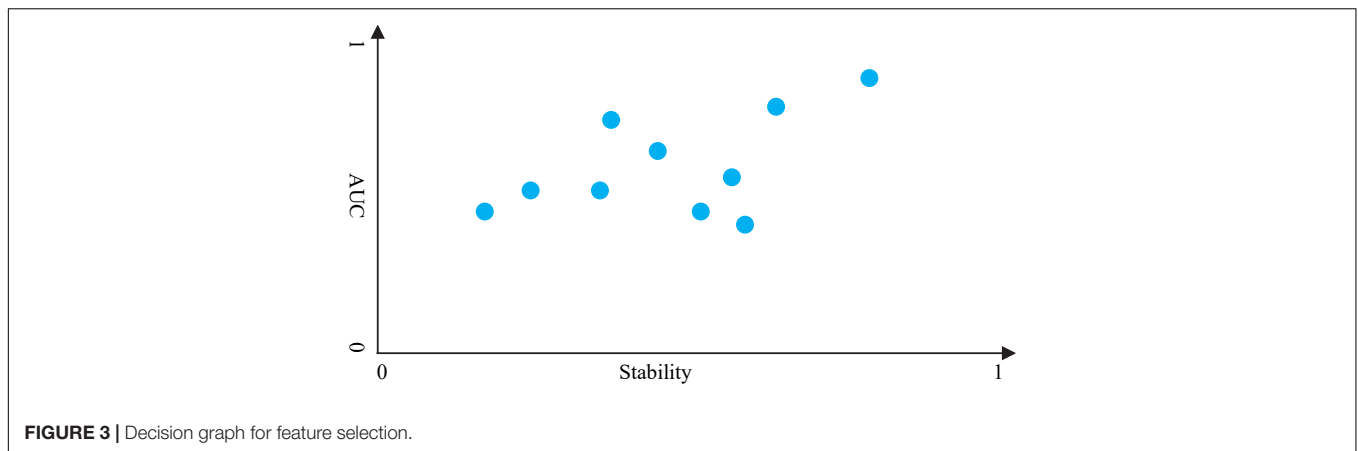
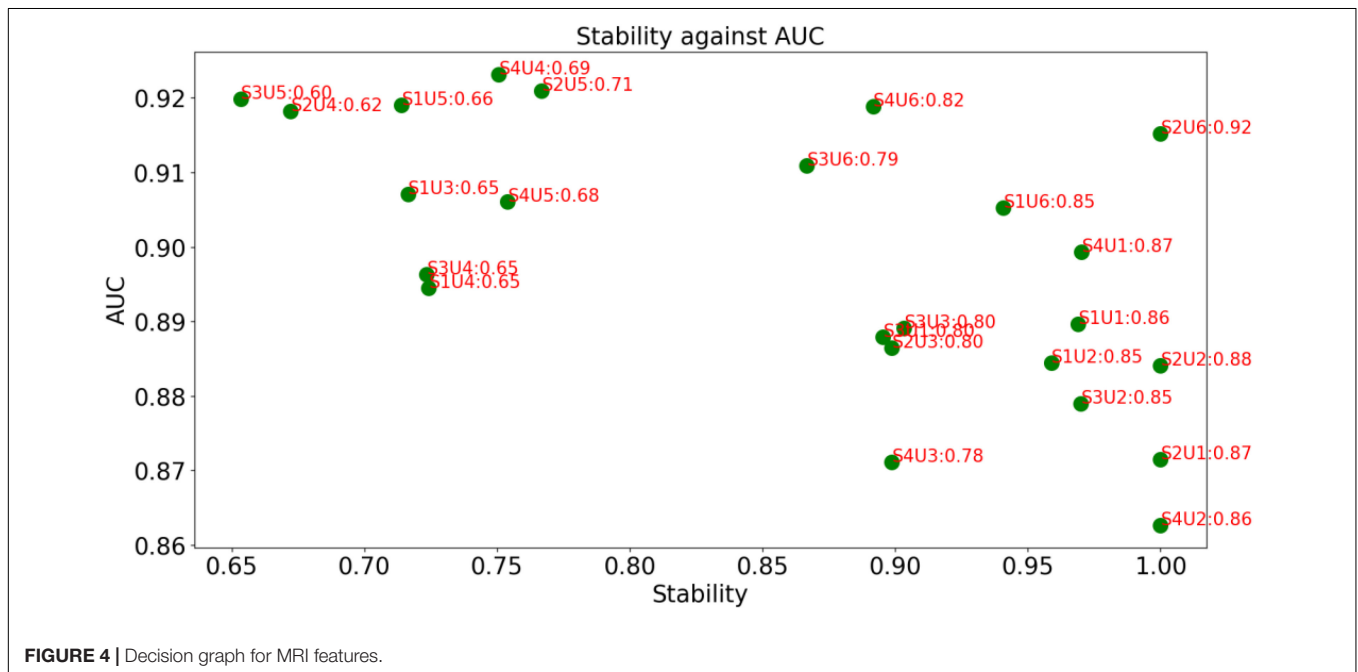**FIGURE 3 |** Decision graph for feature selection.

FIGURE 4 | Decision graph for MRI features.

DISCUSSION

In this study, we have 103 subjects, for both MRI and PET, the feature dimension of each subject is 93, which is near to the number of subjects. When classification models are applied to the high-dimensional data, a critical issue is known as the curse of dimensionality, which refers to the phenomenon that data becomes sparse in high-dimensional space may occur (Li et al., 2017). Therefore, feature selection plays a very significant role in the recognition of AD or MCI. So far, many feature selection methods have been successfully applied in the field of medical image-based diagnosis. For example, in Salvatore et al. (2015), employed PCA (principle component analysis) to select discriminant features from the density maps of WM (white matter) and GM (gray matter) as input of SVM for AD recognition. In Liu et al. (2013), employed LLE (local linear embedding) as the unsupervised feature reduction method to reduce features from the space of multivariate regional

brain volume and cortical thickness MRI to a locally low-dimensional linear space while maintaining the global non-linear data structure. Then, the reduced brain features in the low-dimensional space were used to train the prediction model. Unlike Liu et al. (2013) and Salvatore et al. (2015) in Beheshti et al. (2015) proposed a filter-based supervised feature reduction method containing three main steps. First of all, feature extraction was carried out by using the voxel clusters that are detected by the voxel-based morphometric (VBM) on sMRI and the voxel values as the volume of interest (VOI). Secondly, the probability distribution function of the VOI was employed to represent the statistical information of the respective high-dimensional structural MRI samples. Thirdly, the final selected features were employed to train a SVM classifier to perform the AD recognition task. In Nir et al. (2015) extracted DTI-based features and proposed a tractography-based model to recognize AD and MCI. First of all, the authors used tractography and clustering techniques to locate and organize fibers into 18

FIGURE 5 | Decision graph for PET features.
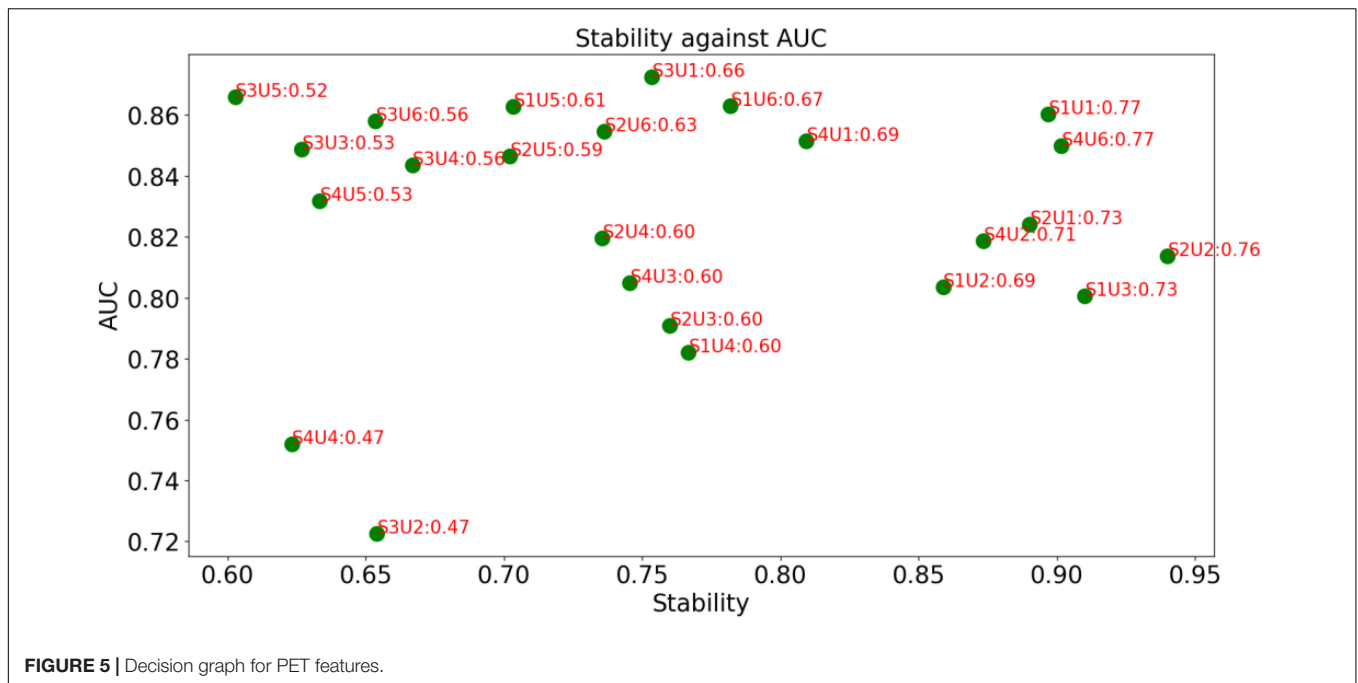
fiber bundles. Secondly, the authors computed density maps to quantify the number of fibers passing through each voxel and used the shortest path graph search to reduce the fiber bundles based on maximum density path (MDP) so that the fiber bundles can be expressed in a compact and low-dimensional space. Thirdly, the diffusivity measures of fractional anisotropy (FA) and MD computed along all the registered across subjects (MDPs) were selected as the features to train an SVM classifier. Feature selection methods in this category can be characterized as making use of the global or local statistical information. In De Martino et al. (2008) employed multivariate feature selection to select features to model functional MRI spatial patterns. To be specific, the authors employed RFE combined with an SVM classifier (REF-SVM) to reduce the irrelevant voxels recursively. Similarly, in Wee et al. (2011), based on DTI images, Wee et al. proposed a framework for MCI recognition. In this framework, the original features come from the anatomical regions, and REF-SVM was also used to reduce the original feature set.

Although different kinds of feature selection (reduction) methods have been widely used for AD and MCI recognition, an important thing that is not fully considered is the stability of the feature selection methods. In practice, we expect that the selected feature selection method can maintain robustness when training data changes slightly. Therefore, in this study, we introduce a frequency-based criterion to evaluate the stability and design a pipeline to select feature selection methods considering both stability and discriminability. Experimental results shown in **Figures 4**, **5** indicate that the proposed pipeline works well and can help us to determine the best combination of feature selection methods. That is to say, the proposed criterion *AUC**Stability can find the optimal combination of supervised and unsupervised feature selection methods.

CONCLUSION

In this study, we introduce a frequency-based criterion to evaluate the stability of feature selection and design a pipeline to select feature selection methods considering both stability and discriminability.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The data is available on http://adni.loni.usc.edu/about/.

AUTHOR CONTRIBUTIONS

FG and SM contributed to the writing and experiments. XW, JZ, and YY contributed to the data collection and preprocessing. XS supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821.

Beheshti, I., Demirel, H., and Alzheimer's Disease Neuroimaging Initiative. (2015). Probability distribution functionbased classification of structural mri for the detection of alzheimer's disease. *Comput. Biol. Med.* 64, 208–216. doi: 10.1016/j.compbiomed.2015.07.006

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., and Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fmri spatial patterns. *Neuroimage* 43, 44–58. doi: 10.1016/j.neuroimage.2008.06.037

Demir, F., and Akbulut, Y. (2022). A new deep technique using R-CNN model and L1NSR feature selection for brain MRI classification. *Biomed. Signal Process. Control* 75:103625.

He, B., Bukhari, S., Fox, E., Abid, A., Shen, J., Kawas, C., et al. (2022). AI-enabled in silico immunohistochemical characterization for Alzheimer's disease. *Cell Rep. Methods* 2:100191. doi: 10.1016/j.crmeth.2022.100191

Hinrichs, C., Singh, V., Mukherjee, L., Xu, G., Chung, M. K., and Johnson, S. C. (2009). Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* 48, 138–149. doi: 10.1016/j.neuroimage.2009.05.056

Hou, K., Zhao, J., Wang, H., Li, B., Li, K., Shi, X., et al. (2020). Chiral gold nanoparticles enantioselectively rescue memory deficits in a mouse model of Alzheimer's disease. *Nat. Commun.* 11:4790. doi: 10.1038/s41467-020-18525-2

Jiang, Y., Zhang, Y., Lin, C., Wu, D., and Lin, C. T. (2020). EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system. *IEEE Trans. Intell. Transport. Syst.* 22, 1752–1764.

Kloppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 681–689. doi: 10.1093/brain/awm319

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., et al. (2017). Feature selection: a data perspective. *ACM Comput. Surv.* 50, 1–45.

Li, S., Yuan, X., Pu, F., Li, D., Fan, Y., Wu, L., et al. (2014). Abnormal changes of multidimensional surface features using multivariate pattern classification in amnestic mild cognitive impairment patients. *J. Neurosci.* 34, 10541–10553. doi: 10.1523/JNEUROSCI.4356-13.2014

Liu, X., Tosun, D., Weiner, M. W., Schuff, N., and Alzheimer's Disease Neuroimaging Initiative. (2013). Locally linear embedding (lle) for mri based Alzheimer's disease classification. *Neuroimage* 83, 148–157. doi: 10.1016/j.neuroimage.2013.06.033

Mainenti, P. P., Stanzione, A., Cuocolo, R., Del Grosso, R., Danzi, R., Romeo, V., et al. (2022). MRI radiomics: a machine learning approach for the risk stratification of endometrial cancer patients. *Eur. J. Radiol.* 149:110226. doi: 10.1016/j.ejrad.2022.110226

Mishra, R., and Li, B. (2020). The application of artificial intelligence in the genetic study of Alzheimer's disease. *Aging Dis.* 11:1567. doi: 10.14336/AD.2020.0312

Möller, C., Pijnenburg, Y., Wm, V. D. F., Versteeg, A., Tijms, B., De Munck, J., et al. (2016). Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. *Radiology* 279, 838–848. doi: 10.1148/radiol.2015150220

Muzik, O., Chugani, D. C., Juhász, C., Shen, C., and Chugani, H. T. (2000). Statistical parametric mapping: assessment of application in children. *Neuroimage* 12, 538–549. doi: 10.1006/nimg.2000.0651

Nir, T. M., Villalon-Reina, J. E., Prasad, G., Jahanshad, N., Joshi, S. H., Toga, A. W., et al. (2015). Diffusion weighted imaging-based maximum density path analysis and classification of Alzheimer's disease. *Neurobiol. Aging* 36, S132–S140. doi: 10.1016/j.neurobiolaging.2014.05.037

Nogueira, S., Sechidis, K., and Brown, G. (2017). On the stability of feature selection algorithms. *J. Mach. Learn. Res.* 18, 6345–6398.

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., and Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage* 206:116189. doi: 10.1016/j.neuroimage.2019.116189

Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307

Subasi, A. (2020). Use of artificial intelligence in Alzheimer's disease detection. *Artif. Intell. Precis. Health* 2020, 257–278.

Wee, C.-Y., Yap, P.-T., Li, W., Denny, K., Browndyke, J. N., Potter, G. G., et al. (2011). Enriched white matter connectivity networks for accurate identification of mci patients. *Neuroimage* 54, 1812–1822. doi: 10.1016/j.neuroimage.2010.10.026

Xia, K., Zhang, Y., Jiang, Y., Qian, P., Dong, J., Yin, H., et al. (2020). TSK fuzzy system for multi-view data discovery underlying label relaxation and cross-rule & cross-view sparsity regularizations. *IEEE Trans. Indust. Inform.* 17, 3282–3291.

Xiao-Cong, P. A. N. G., De, K. A. N. G., Jian-Song, F. A. N. G., Ying, Z. H. A. O., Lv-Jie, X. U., Wen-Wen, L. I. A. N., et al. (2018). Network pharmacology-based analysis of Chinese herbal Naodesheng formula for application to Alzheimer's disease. *Chin. J. Nat. Med.* 16, 53–62. doi: 10.1016/S1875-5364(18)30029-3

Zhang, Y., Chung, F. L., and Wang, S. (2020a). Clustering by transmission learning from data density to label manifold with statistical diffusion. *Knowl. Based Syst.* 193:105330.

Zhang, Y., Lam, S., Yu, T., Teng, X., Zhang, J., Lee, F. K. H., et al. (2022). Integration of an imbalance framework with novel high-generalizable classifiers for radiomics-based distant metastases prediction of advanced nasopharyngeal carcinoma. *Knowl. Based Syst.* 235:107649.

Zhang, Y., Wang, G., Chung, F. L., and Wang, S. (2021a). Support vector machines with the known feature-evolution priors. *Knowl. Based Syst.* 223:107048.

Zhang, Y., Wang, S., Xia, K., Jiang, Y., Qian, P., and Alzheimer's Disease Neuroimaging Initiative. (2021b). Alzheimer's disease multiclass diagnosis *via* multimodal neuroimaging embedding feature selection and fusion. *Inform. Fusion* 66, 170–183.

Zhang, Y., Zhou, Z., Bai, H., Liu, W., and Wang, L. (2020b). Seizure classification from EEG signals using an online selective transfer TSK fuzzy classifier with joint distribution adaption and manifold regularization. *Front. Neurosci.* 14:496. doi: 10.3389/fnins.2020.00496