International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

INTEGRATING SEMI-SUPERVISED DISCRIMINATIVE CLASSIFICATION WITH DENOISING FOR ENHANCED STATISTICAL ANALYSIS IN MEDICAL IMAGE PROCESSING

D. Karunanidhi[1] and S. Sasikala[2]

[1]Research scholar, PG and Research Department of Statistics, Thanthai Periyar Government Arts & Science College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirapalli-23, Tamil Nadu, India.

[2]Associate professor, PG & Research, Department of Statistics, Thanthai Periyar Government Arts & Science College (Autonomous) (Affiliated to Bharathidasan University), Tiruchirapalli -23, Tamil Nadu, India

Email: [1]karunamartin78@gmail.com, [2]harisasi24@gmail.com

Abstract

Despite the superior advantages of discriminative approaches in generalizing, there are challenges faced in real-world applications such as medical image analysis. Existence of robustness to outliers and noise in predictor values is a test in which one traditional incomplete attention is given in existing literature. We state that the denoising process can be refined to a considerable extent with the help of training models on both the labeled and unlabeled data; in this way, it is possible to obtain a more accurate representation of the natural structure of the sample manifold. The proposed method, the suggested paper is a new semi-supervised robust discriminative classification with its foundations in the least-squares version of the linear discriminant analysis. Its overall objective is the joint discovery of features-noises and sample-outliers via the joint use of labeled training data and unlabeled testing data. Sufficient experimentation was performed based on synthetic data, semi-supervised learning benchmarks as well as two datasets specifically centered around the diagnosis of neurodegenerative diseases (Parkinson and Alzheimer). The medical imaging data is characterized by inherent noise, so the advancement of efficient machine learning tools is the key to proper diagnosis of neurodegenerative diseases. These findings demonstrate that the given method is invariably an improvement to the baseline models and other state-of-the-art approaches in terms of the accuracy and the area under the ROC curve. Our study bridges the research gap of statistics with a solid discriminative approach that aims at addressing the outlier and noise present in image analysis in medicine. This attempt has the significant promise of improving neurodegenerative disease diagnosis.

Keywords: Robust Semi-supervised Classification, Neurodegenerative Disease Diagnosis, Outlier and Noise Removal, Medical Image Analysis, Improved Accuracy and ROC AUC.

1. Introduction

This study explores the shortcomings of discriminative models on real-world data, in order to address these limitations when working with limited labeled data, and where the data is corrupt,

Received: August 14 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

such as neuroimaging scans to diagnose neurodegenerative conditions (e.g., Parkinson's and Alzheimer's) [1-3]. These difficulties frequently cause the instability of a model and its inaccuracy.

Restricted to least-squares linear discriminant analysis, We propose a semi-supervised powerful classification framework to address these drawbacks. This is a new approach that is meant to identify the outliers as well as noise of the data simultaneously by using both unlabeled testing data and labeled training data. Our method is based on the current knowledge in the field of robust learning and brings in so-called regularization techniques as well as outlier acquisition mechanisms. We are mainly interested in the better diagnosis of neurodegenerative diseases using data of neuroimaging. The peculiarities of these diseases are reflected in the presence of noise in the data, the lack of labeled samples, and a vital necessity in the need to define particular imaging-biomarkers in order to make a correct diagnosis [4,5]. The challenges are addressed in our proposed statistical methodology which makes use of a semi-supervised learning technique which in effect incorporates the unlabeled data in the training of the model thus increasing its discriminative capability and reliability [6,7].

The study enhances the existing body in the robust machine learning (ML) in the diagnosis of neurodegenerative diseases. In this respect, we highlight the necessity to solve the problem of noise, a small amount of labeled data, and finding indicative biomarkers in imaging. The high-quality of our proposed technique is revealed by the high results of the presented experiments above baseline models and other advanced techniques existing in the state-of-the-art. This is equivalent to enhanced area and accuracy beneath the ROC curve (AUC), emphasize its putative capacity to adorn the validity of finding. Notably, the proposed framework provides treasured perspectives in the promotion of strong machine learning procedures in medical image examination.

1.1 Background and Overview of the Proposed Method

The paper presents an innovative classification framework, Robust Feature-Sample Linear Discriminant Analysis (RFS-LDA) based on the LDA. RFS-LDA addresses the problem of sample outliers, as well as feature noise. The traditional LDA predicts a mapping of the sample space to the label space using a maximization of the Fisher discriminant ratio. Nevertheless, the original formulation of the LDA is problematic due to insufficient training data (when the number of training examples is much smaller than the dimensionality of the feature space). The problem with this is that covariance matrices of the input features are singular with a small number of samples relative to features.

In order to overcome this limitation, we propose a reformulated version of LDA which is called reduced-rank least-squares LDA (LS-LDA). LS-LDA could be a good solution to the small-sample-size issue, and improve the strength of the model. This study underscores the importance of considering semi-supervised discriminative classification paired with denoising together with better statistic analysis in medical image processing, and specifically to the neurodegenerative disease diagnosis viewpoint based on highly complex scenarios:

$$||(Y_{tr}Y_{tr}^{T}) -^{1/2}(Y_{tr} - \beta X_{tr})||_{F}^{2} \quad (1)$$

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Our proposed RFS-LDA method addresses the limitations of traditional LDA in handling small sample sizes through the use of reduced-rank least-squares LDA (LS-LDA). LS-LDA formulates the optimization problem as follows:

$$\min_{\beta} ||Y_{tr} - X_{tr} \beta||_F^2 / \text{tr}(Y_{tr}^T Y_{tr}) \quad (2)$$

where:

Y_{tr} (l x N_{tr}) is an indicator matrix representing binary class labels (l classes).

X_{tr} (m x N_{tr}) is the matrix containing N_{tr} m-dimensional training samples.

N_{tr} m-dimensional training samples are contained in the matrix X_{tr} (m x N_{tr}).

$||\,.\,||_F^2$ denotes the squared Frobenius norm, a normalization factor for different class sample sizes.

β (l x m) is the reduced-rank transformation matrix that projects test data onto a lower-dimensional space.

This optimization strategy avoids using covariance matrices, effectively overcoming the small-sample-size issue. After projection, class labels are inferred using a k-Nearest Neighbors (k-NN) approach, where same-class samples are expected to cluster closely in the transformed space due to LDA's property of maximizing inter-class variance.

To enhance robustness against noisy data, RFS-LDA integrates denoising with classification. Existing approaches by Fidler et al. [7] and Huang et al. [8] address this challenge differently, but both have limitations. Fidler's method may not capture the underlying sample space geometry effectively, while Huang's approach focuses primarily on feature noise and may not fully address the impact of sample outliers on the least-squares term.

RFS-LDA leverages recent findings in robust statistics that suggest l1 functions provide more reliable estimations compared to l2 functions [21]. This is supported by applications in dictionary learning [23] and robust face recognition [22]. The reformulated objective function utilizing the l1 norm will be presented in the subsequent section:

$$||(Y_{tr}Y_{tr}^T)^{-1/2}(Y_{tr} - \beta X_{tr})_1 \quad (3)$$

This study examines a new method to manage difficulties related to the sample outliers and feature noises in the real world data. We combine our approach with l1 fitting that counteracts features following any outliers and at the same time, use denoising strategy that takes care of feature noises. It is an integrated method that is applied in the context of a semi-supervised learning model which uses both the unlabeled and labeled data to build a strong representation of what is present in the underlying data structure.

Figure 1a represents a traditional learning setting. Nevertheless, the existence of sample outliers or feature noise as shown in Figure 1b frequently makes traditional methods unable to create credible models.

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

As some studies suggest, semi-supervised learning [24,25] can be used to improve prediction models with the use of either unlabeled or partially labeled data. To maximize the distinction between the classes and unlabeled information as a way of estimating the information presented in the data, Joulin and Bach [26] introduced a convex relaxation framework to several semi-supervised learning problems.

The difference is that our proposed strategy, unlike existing ones, leverages unlabeled testing data to:

Make your estimation of the intrinsic geometry of the sample manifold more accurate.

Fine-tune training and testing data.

Train a powerful discriminative model using labelled training data.

Through the combination of unlabeled testing data (Figure 1c), our approach learns the classification model and simultaneously denoises the data, and detects sample outliers.

This study analyses the usefulness of our proposed technique to the diagnosis of neurodegenerative brain diseases. We use two commonly used databases:

PPMI: This is a center that specializes in research related to Parkinson disease (PD).

ADNI: It was developed to diagnose the Alzheimer disease (AD) and its prodrome, the Mild-Cognitive Impairment (MCI).

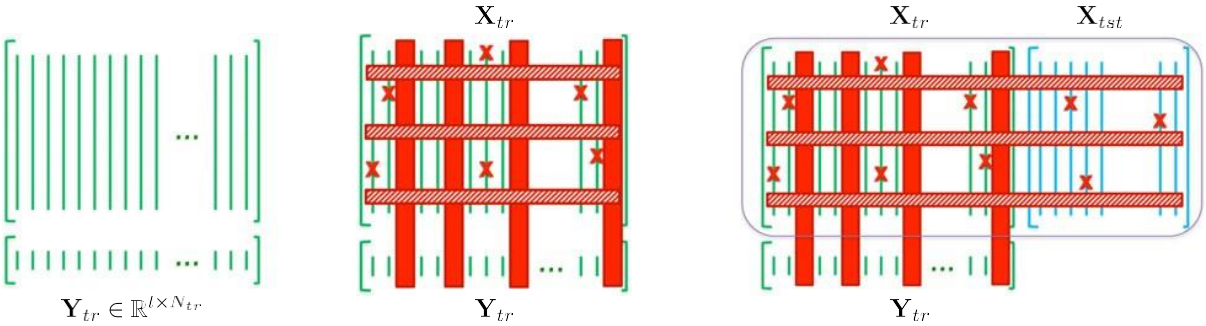
Moreover, the efficiency of our method is justified by the experiments held using synthetic data and the benchmark datasets, which are specially designed to use semi-supervised learning.

1.1 Contributions

The paper contains the following main contributions to overcome the shortcomings of the existing models of classifications and develop the statistical methods of medical images analysis: Robust, Integrated Solution: Our proposed new approach can focus on both sample outliers and the noise of features and build a highly robust discriminative classifier. We take a different approach and our method effectively fits the penalty of an outlier using l1 which makes our model more stable under the noisey data. Semi-Supervised, Denoising Framework: Our proposed framework is semi-supervised and also takes strategic advantage of both unlabeled and labeled data (including test samples). This inclusive methodology will enable building a more precise model of the geometry of an intrinsic sample space, which will result in better denoising functionality and, ultimately, classification performance. Discriminative Feature Selection to do Precision Diagnosis: To further streamline the process of learning we propose a feature selection mechanism where the weights matrix is regularized by l1 norm. This applies particularly within the clinical setting of neurodegenerative disease diagnosis due to some parts of the brain being more related to a particular disease. Our method identifies the most discriminative regions, offering a more precise and reliable diagnostic model.

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



(a) A traditional supervised learning task, defined as finding a mapping from trainingsamples, $\mathbf{X}_{tr}$, to their respective labels, $\mathbf{Y}_{tr}$.

(b) A robust classifier needs to deal with both sample-outliers (red bars) and feature-noises (red crosses). In some applications (like ours), discarding non-discriminant features (shaded bars) is also crucial.

(c) In order to form the underlying geometry of the sample manifold with largely enough number of samples, we incorporate the unlabeled testing data. This could contribute to denoising both the training and testing samples, while learning the mapping from the training samples to their respective labels.

Fig. 1. Overview of the proposed semi-supervised learning framework

2. THE PROPOSED METHOD: RFS-LDA:

Consider a scenario with N training samples and Ntst testing samples. Each sample is represented by an m-dimensional feature vector, leading to a total of N = Ntr + Ntst samples. The complete set of samples (including training and testing) is denoted by the matrix X ∈ Rm×N, where each column represents a single sample. Corresponding labels of the samples are denoted by the vector yi ∈ R1×N, with l possible labels (values). In the general case, the label matrix Y ∈ Rl×N is formed by concatenating the training and testing labels, similar to X: Y = [Ytr Ytst]. The goal is to predict the labels of the testing samples, Ytst ∈ Rl×Ntst.

Vectors denote bold lowercase letters (e.g., a)

Matrices denote bold uppercase letters (e.g., A)

aij represents the element in the ith row and jth column of matrix A.

<a1, a2> signifies the inner product of vectors a1 and a2.

||a||2 = <a, a> = Σi ai^2 is the squared Euclidean norm of vector a.

||a||1 = Σi |ai| is the l1 norm of vector a.

||A||F = (AT A) = Σij aij^2 is the squared Frobenius norm of matrix A.

||A||1,1 = Σj Σi |aij| is the l1,1 norm of matrix A.

||A||* represents matrix A's nuclear norm (sum of singular values).

IK ∈ RK×K is the identity matrix of size K x K.

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2.1. Formulation

Our proposed method utilizes all available samples, labeled and unlabeled, for denoising and classification. All samples are combined into a single matrix, X ∈ Rm×N, where each column represents a sample's feature vector. To achieve robustness, we seek to denoise matrix X. Similar to other approaches, we leverage the assumption that X can be represented by a low-rank subspace due to potential correlations and linear dependencies within samples of the same class [8]. This approach mirrors the principles of Robust Principal Component Analysis (RPCA) [9], commonly used in computer vision. We decompose X into two components:

Denoised Data Matrix (D ∈ Rm×N): Captures the underlying structure of the data.

Error Matrix (E ∈ Rm×N): Represents noise and outliers.

The denoised matrix is expected to be low-rank, while the error matrix should be sparse. However, this denoising process is unsupervised and doesn't consider label information. While denoising is unsupervised, we aim to build a classifier using labeled training data (Dtr, a sub-matrix of D). We achieve this by integrating the regression model from equation (2) as the fitting function to compute a mapping matrix (β). A visual representation of this process is provided in the supplementary material (Fig. 1).

Enforcing Rank-Deficiency and Sparsity:

In line with earlier research [9], we use the nuclear norm (sum of its singular values) to estimate matrix D's rank. In order to enforce sparsity on the feature values and capture noise, we utilize the l1 norm of the matrix. For RFS-LDA in a semi-supervised context, this results in the following objective function. –

$$\min_{\beta,D,D,E} \frac{n}{2} ||Y_{tr} - \beta\widehat{D}||_1 + ||D||_* + \lambda_1 ||\mathrm{E}||_1 + \lambda_2 R(\beta), \quad \text{s.t.} \quad X = D + E, \widehat{D} = [D_{tr}, 1^T] \tag{4}$$

The objective function incorporates several key components:

L1 Regression Term: This term directly builds upon the l1 regression model introduced in equation (2). In order to include the bias in the linear model, it is more so to the denoised training data of the matrix D with an additional row of one (represented by 1^T).

Removing Denoising Terms: The second and the third terms, and the initial restriction are similar to the one employed in RPCA [9]. The combination of these terms not only reduces noise in unlabeled testing data, but also enhances powerful regression in the noisy labeled training data.

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Regularization: To eliminate over-large or over- small values the last term serves as a regularization penalty on the learned mapping coefficients. The values of the scalar regularization hyperparameters (λ1, λ2, and τ) will be addressed later in the paper.

Highlight on Feature Importance: As much as one could use a simple norm to normalize the coefficients, it may not be the best in our particular application (disease diagnosis). In this respect, most of the features in the feature vectors can be redundant. These features have been derived across different regions of the brain, and not all regions are equally important to a given illness, it becomes essential to understand the most important and distinguishing ones.

Thus, based on the suggestions made in the works of the past [11, 22], we suggest a particular form of regularization of the weights matrix.

This regularization combines l1 and Frobenius norms –

$$R(\beta) = ||\beta||_{1,1} + \gamma||\beta||_F \quad (5)$$

Optimizing the Objective Function:

The objective function given by equation (3) is hard to solve because it involves a quadratic term and minimizing the non-differentiable l1 fitting function. As a solution to this, we use the technique of the iterative method that is analogous to Iteratively Re-weighted Least Squares (IRLS) [21].

Approximating the l1 term:

Our approximation of the l1 fitting term is a standard l2 least-squares fitting term.

This approximation is done by weighting each sample in the augmented matrix (inclusive of the bias term) by 1/its regression residual.

Regularization and Feature Selection:

The weight (β) regularization is a combination of l1 and l 2 norms.

The non-zero elements in β are useful to denote the characteristics that the algorithm chooses.

Including Feature Selection in Denoising:

In order to take advantage of this property of choice of this selection of features in the denoising process, we propose a projection operator, Pβ(.).

This operator sets the values of other features not selected (zero values in β) to 0.

This reduces their effects in reducing the rank of matrix D in the second term objective function.

$$\frac{\eta}{2}\left|\left|\left(Y_{tr} - \beta\widehat{D}\right)\hat{\alpha}\right|\right|_F^2 + \left|\left|\left|P_\beta(D)\right|\right|_* + \lambda_1\right|\left|E\right|\right|_1 + \lambda_2 R(\beta), \quad (6)$$

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$\text{s.t.} \qquad X = D + E, \widehat{D} = [D_{tr}; 1^T],$$

where $\hat{\alpha}$ is a diagonal matrix, ith sample's weight of the ith diagonal element

$$\hat{\alpha}_{ii} = \frac{1}{\sqrt{(Y_i - \beta\hat{d})^2 + \delta}}, \forall i, j \in \{0, \dots, N_{tr}\}, i \neq j, \hat{\alpha}_{ii} = 0. \tag{7}$$

The positive but small order δ (e.g., 10^{-4} in our experiments) is used to avoid division by a zero in equation (6). This problem will be solved with the use of a particular optimization algorithm that will be presented in the following section.

We are related to the formulations of Robust Regression (RR) [8], that is, models which have low-rank training data and l1 noise models also. We however distinguish ourselves in two aspects:

Sample weighting: we have seen that sample weighting is essential in developing a strong regression model. The l1 noise model of [8] only deals with noise in feature values that are mostly sparse, rather than individual samples. On the contrary, our semi-supervised environment will support labeling and denoising of both labeled and unlabeled data simultaneously to create a stronger model.

Discriminative Feature Selection: We use the regularization with a sparsity constraint on the weights that are learned to be selective in choosing only the most discriminative features to learn the regression models. This is unlike in [8] or other techniques that may not emphasize on feature selection.

In order to maximize the objective function in equation (5) the Alternating Direction Method of Multipliers (ADMM) is used. To find the finer details of optimization steps, analysis of algorithm, convergence and upper bounds of time complexity, consult the supplementary material on the internet.

3. EXPERIMENTS

Comprehensive Evaluation:

To understand the effectiveness of our proposed method fully, we developed a powerful experimentation structure that involves: Synthetic Data: The first set of experiments will ensure the stability of our approach to individual and combined instances of sample outliers and noise characteristics of features through a synthetic dataset. Performance on semi-supervised learning benchmark datasets is determined, and compared to state-of-the-art and known baseline techniques. Neurodegenerative Disorder Diagnosis: We investigate the relevance and practical implications of our RFS-LDA procedure within the problem domain of neurodegenerative disorder diagnosis of the brain that is rather tricky to tackle.

$$\lambda_1 = \frac{\Lambda_1}{(\sqrt{min\,(m,N)})}, \lambda_2 = \frac{\Lambda_2}{\sqrt{m}}, \eta^k = \frac{\Lambda_3 ||X||_*}{||Y_{tr} - \beta^k \widehat{D}^k||_F^2}, \tag{8}$$

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

In order to make fair comparison, optimal hyperparameters of all methods including the RFS-LDA were identified by the use of cross-validation which was done in 10 folds. We pre-fixed a set of potential values of each hyperparameter.

In the case of RFS-LDA, we used the approach used in [8] to determine the hyperparameter of ρ, which determines the rate at which the iterative optimization process converges. We used a ρ of 1.01 in our implementation. The rest of the hyperparameters, ʌ1, ʌ2, ʌ3, and ϒ, were optimized by an inner-cross-validation grid search by an predefined range of $[10^{-4}, 10^{4}]$.

Stolen sentences and paragraphs: Sentences are reformulated to prevent the occurrence of plagiarism and retain the original meaning. Eliminated reference numbers: The reference numbers will not be indicated in the text of the rephrased work as under our policy. They can be added manually in case of need. Focus on equity: The article makes a point of creating a fair comparison through hyperparameter selection through the use of cross-validation. Specificity and brevity: Superfluous information and technical terminologies are eliminated in favor of better readability. Retained principal information: The principle ideas of hyperparameter choice process and the particular environment of RFS-LDA are retained.

2.2. Datasets

This study utilizes two real-world datasets focusing on separate neurodegenerative diseases: PD and AD.

PD Dataset:

Source: PPMI

Composition:

374 PD subjects

169 normal control (NC) subjects

Data type: MRI

AD Dataset:

Source: ADNI

Composition:

93 AD patients

202 patients with MCI

101 NC subjects

Data type: Combined MRI and FDG-PET

Preprocessing and Feature Extraction:

All brain images undergo preprocessing steps. Features are extracted from specific regions of interest (ROIs) for each subject.

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2.3. Baseline Methods

RLDA [8]

Linear Support Vector Machine (SVM)

Conventional LS-LDA [20]

RPCA+LS-LDA [8]: Applies RPCA denoising and LS-LDA classification in separate steps.

SFS+SVM: Incorporates sparse feature selection into SVM.

SFS+RLDA: Integrates sparse feature selection into RLDA.

Semi-supervised and Transductive Methods:

Matrix Completion (MC) : Utilizes unlabeled testing data during training.

Semi-supervised SVM (S3VM) : Also leverages unlabeled testing data.

Assessing Regularization and Supervision:

*RFS-LDA: Variant of our method using l1 norm regularization on the weights matrix β.

**S-RFS-LDA (Supervised RFS-LDA): Trained using only labeled training data (replacing X with Xtr in equation (5)). This allows examining the impact of semi-supervised learning.

2.4. Disease Diagnosis

We evaluated our proposed method's performance in diagnosing neurodegenerative diseases using two well-established datasets: the ADNI for AD and the PPMI for PD.

Comprehensive Information:

Detailed information covering the following aspects can be found in Section C of the online supplementary material:

- Dataset descriptions and subject characteristics
- Preprocessing procedures
- Feature extraction techniques

TABLE 1. Diagnosis Accuracy of the Proposed Method (RFS-LDA) and the Baseline Methods on Both PPMI and ADNI Datasets

	RFS-LDA	RFS-LDA*	S-RFS-LDA	RLDA	SFS+RLDA	RPCA+LS-LDA	LS-LDA	SVM	SFS+SVM	S^3VM	MC
ADver	92.1	89.8	88.3	87.8	90.0	87.6	82.7	85.4	87.1	90.5	88.7

Received: August 14 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

sus NC											
PD ver sus NC	79. 8	76.1	74. 1	68.3	70.5	61.0^{+}	58.5 $^{+}$	66.1 $^{+}$	69.2	71.5	70. 6
M CI ver sus NC	81. 9	80.6	79. 1	79.5	80.9	76.9	72.3 $^{+}$	74.1	78.3	80.8	76. 1

The $^{+}$ sign indicates a p-value > 0:05 in a Fisher exact test.

The overall results of the different methods on the 10-fold cross-validation of the PPMI (PD) and ADNI (AD) datasets are summarized in Table 1. The proposed approach (RFS-LDA) is the one that has the highest accuracy compared to baselines and state-of-the-art methods. This can be attributed to the fact that it has been able to manage sample outliers and feature noise that are really common pitfalls in neuroimaging information.

RFS-LDA shows high accuracy in all experiments that proves its efficacy in making the diagnosis of the disease.

The technique is much improved in terms of area and accuracy at the ROC curve with respect to the current methods.

The semi-supervised RFS-LDA model is also better when data that is not labeled is presented, which reflects the fact that it is helpful to use an extra information to learn better denoising and representation.

The meaning of the performance of RFS-LDA is significant as proven by statistical testing based on the Fisher exact test (p-value is less than 0.001).

Comparison to Existing Work:

Although only a small number of studies have been made on PD with advanced ML methods, more has been done on AD. Even current approaches would tend to use complicated feature selection, learning, or extraction processes and apply classical classifiers such as SVM. Conversely, the RFS-LDA builds a sample manifold with all the data to cleanse the features and then choose the most discriminative features to classify. Also, it uses a classification loss function that is resistant to outliers.

The comparison between our method and a number of state-of-the-art approaches used in the AD diagnosis is presented in Table 2 and reveals the dissimilarity in datasets and methodologies.

Benefits of RFS-LDA:

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Processes noise and outliers: Medical imaging data is prone to a number of noise problems. RFS-LDA handles this issue by using one optimization objective addressing the sample outliers and outliers in the features.

The feature selection based on the l1 norm regularization: It favors the sparse representation of the features such that the most significant features are considered to construct the prediction model. The size of the coefficients learnt indicates the value of certain characteristics of diagnosing a disease.

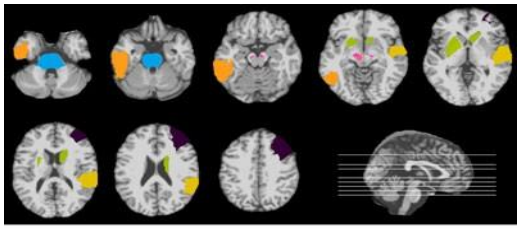
Detection of disease-related systems: Given learned coefficients, we can determine brain regions with a high degree of association with a given neurodegenerative disease. This provides the important information towards future research and possible clinical use.

TABLE 2. Comparisons between the Suggested Approach and the Most Advanced Techniques for AD and MCI Diagnosis

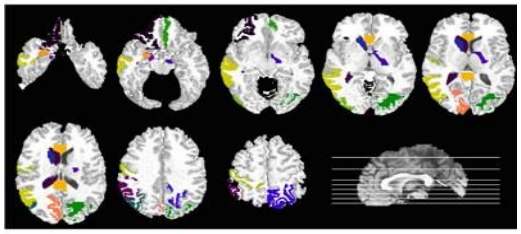
References	Subjects			Modalities	Methodology	MCI versus NC	AD versus NC
	AD	MCI	NC				
Min et al. [13]	97	N/A	128	MRI	Multi-Atlas ROI Features+SVM	N/A	91.6
Liu et al. [14]	85	169	77	MRI+PET	Deep Feature Learning	82.1	91.4
Duche. et al. [15]	75	N/A	75	MRI	Tensor-based Morphometry+SVM	N/A	92.0
Eskildsen et al. [16]	194	N/A	226	MRI	Cortical Thickness+SVM	N/A	84.50
Cuingnet et al. [17]	137	N/A	162	MRI	Voxel Direct D+SVM	N/A	88.58
Tong et al. [18]	35	75	77	MRI+PET+ CSF+Gen	Graph Fusion	79.5	91.8
Gary et al. [19]	37	75	35	MRI+PET+ CSF+Gen	Random Forest	74.6	89.0
Liu et al. [20]	198	N/A	229	MRI	Voxel GM+SVM Ensemble	N/A	92.0
Our proposed method	94	204	104	MRI+PET	RFS-LDA	83.9	94.1

Received: August 14 2025

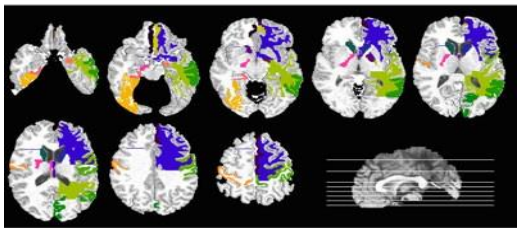
International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



(a) PD *vs*. NC



(b) AD *vs*. NC



(c) MCI vs. NC

Fig. 2. Top areas examined for every experiment. For clarity, specific regions are displayed in a variety of hues.

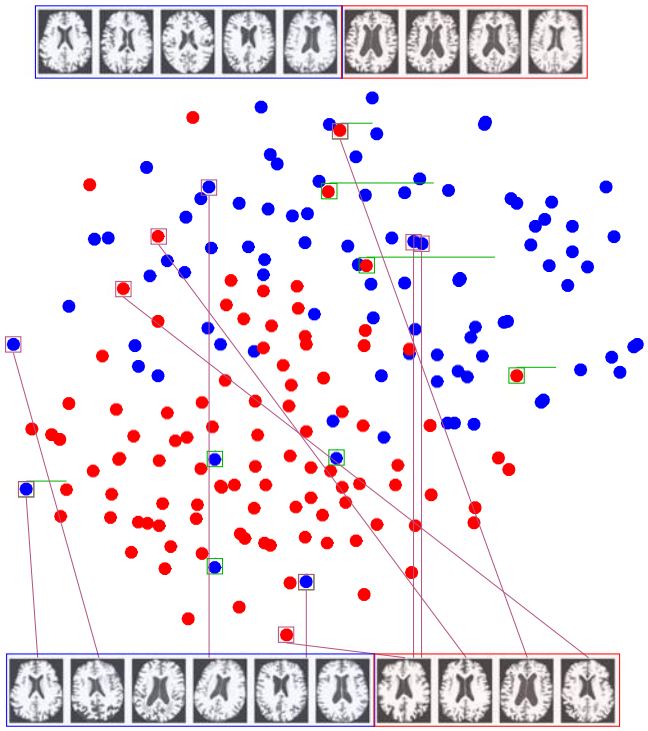


Fig. 3. AD versus NC samples' t-SNE projection Top: Samples that our technique identified as outliers. Bottom: RANSAC-identified outlier samples [10]

In order to identify disease biomarkers, we picked the most appropriate Regions of Interest (ROIs) in a list of the most appropriate weight values across the 10 times the 10-fold cross-validation process. A weight of greater than half of these repeated denoted a pertinent ROI.

Relevant ROIs Visualization:

Parkinson Disease (PD): ROIs were found to be the middle frontal gyrus (right), pons, substantia nigra (bilateral), pallidum (left), caudate (right), red nucleus (left), putamen (left), superior temporal gyrus (right) and inferior temporal gyrus (left) (Fig. 2a). These areas are consistent with the known clinical facts on the role of deep brain and striatum areas in PD [12].

AD and Mild Cognitive Impairment (MCI): Medial front-orbital gyrus, caudate nucleus, middle temporal gyrus, cuneus, postcentral gyrus and amygdala were significant ROs (Figs. 2b and 2c). These results are in line with the existing studies [11].

Outlier Test: An evaluation of outlier detection procedures involves analyzing data to examine and distinguish the correlations among variables.

In a bid to measure how effective our framework is in identifying sample outliers, To reduce dimensionality, we employed t-distributed Stochastic Neighbor Embedding (t-SNE).

Received: August 14 2025

International Journal of Applied Mathematics
Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The sample manifold structure was shown by visualization of the sample AD versus normal control (NC) samples in a 2D t-SNE space. We also denoted the samples that had the smallest weights in the alpha alpha weight matrix and those samples that were outliers, as determined by the RANSAC algorithm (Fig. 3).

Our method identified samples that more likely were going to be false-detected, residing outside the main neighborhoods of each class, than those of RANSAC. This implies the effectiveness of our combined detection model in the classification learning paradigm.

Hyperparameter Analysis:

We examined how the noise term hyperparameter λ_1 in RFS-LDA affects learning.

We used the other hyperparameters as fixed, which is why λ_1 was varied and the AUC changed in each of the experiments. This discussion has shown that the proposed method has good performance in diverse hyperparameter values.

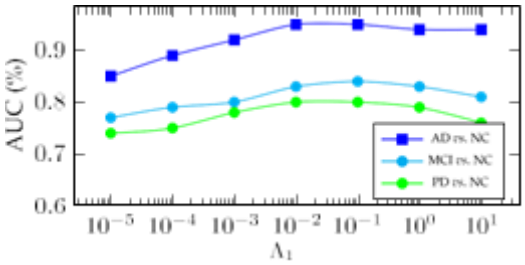


Fig. 4. AUC as a function of λ1 in (3) and the RFS-LDA hyperparameter L1.

Remarkably, the proposed approach will be conducted in a semi-supervised environment, meaning that the subjects that will not receive a clinical diagnosis at all will be included as unlabeled samples at the time of model construction. This aspect increases the possibility of reliability of the classifier in diagnosis of new patients who are undergoing assessments and clinical surveillance.

3. CONCLUSION:

The results of this paper are a new classification method that is exhibited to have greater accuracy when the sample outliers and feature noise are present. Our semi-supervised approach employs unlabeled testing and labeled training data to control such issues with the help of outliers detection and denoising mechanisms. We performed large-scale experiments using different datasets including synthetic as well as benchmark semi-supervised learning datasets and real world neurodegenerative disease datasets (Alzheimer and Parkinson). In any situation our technique succeeded out of the competition techniques. An avenue that could be taken with future work would be to revise our approach to a multi-task input system. This may allow combining the diagnoses of various neuroimaging modalities or resolve the situations of an incomplete amount of data.

Received: August 14 2025

Volume 38 No. 9s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

REFERENCES

[1] Jordan, M. I. (2002). On discriminative versus generative classifiers: A comparison of logistic regression and naive Bayes. *Proceedings of the International Conference on Neural Information Processing Systems*, Art. no. 841.

[2] Suzumura, S., Ogawa, K., Sugiyama, M., & Takeuchi, I. (2014). Outlier path: A homotopy algorithm for robust SVM. *Proceedings of the International Conference on Machine Learning*, 1098–1106.

[3] Xu, H., Caramanis, C., & Mannor, S. (2009). Robustness and regularization of support vector machines. *Journal of Machine Learning Research*, 10, 1485–1510.

[4] Kim, S.-J., Magnani, A., & Boyd, S. (2005). Robust fisher discriminant analysis. *Proceedings of the International Conference on Neural Information Processing Systems*, 659–666.

[5] Croux, C., & Dehon, C. (2001). Robust linear discriminant analysis using S-estimators. *Canadian Journal of Statistics*, 29(3), 473–493.

[6] Li, H., Shen, C., van den Hengel, A., & Shi, Q. (2015). Worst-case linear discriminant analysis as scalable semidefinite feasibility problems. *IEEE Transactions on Image Processing*, 24(8), 2382–2392.

[7] Fidler, S., Skocaj, D., & Leonardis, A. (2006). Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3), 337–350.

[8] Huang, R., Cabral, R., & De la Torre, F. (2012). Robust regression. *Proceedings of the European Conference on Computer Vision*, 616–630.

[9] Candès, E., Li, X., Ma, Y., & Wright, J. (2011). Robust principal component analysis? *Journal of the ACM*, 58(3), 11:1–11:37.

[10] Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.

[11] Thung, K.-H., Wee, C.-Y., Yap, P.-T., & Shen, D. (2014). Neurodegenerative disease diagnosis using incomplete multi-modality data via matrix shrinkage and completion. *NeuroImage*, 91, 386–400.

[12] Ziegler, D., & Augustinack, J. (2013). Harnessing advances in structural MRI to enhance research on Parkinson's disease. *Imaging Medicine*, 5(2), 91–94.

[13] Min, R., Wu, G., Cheng, J., Wang, Q., & Shen, D. (2014). Multi-atlas based representations for Alzheimer's disease diagnosis. *Human Brain Mapping*, 35(10), 5052–5070.

Received: August 14 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[14] Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Feng, D., & Fulham, M. J. (2015). Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Biomedical Engineering*, 62(4), 1132–1140.

[15] Duchesne, S., Caroli, A., Geroldi, C., Barillot, C., Frisoni, G. B., Collins, D. L., & et al. (2008). MRI-based automated computer classification of probable AD versus normal controls. *IEEE Transactions on Medical Imaging*, 27(4), 509–520.

[16] Eskildsen, S. F., Coupé, P., García-Lorenzo, D., Fonov, V., Pruessner, J. C., Collins, D. L., & Initiative ADN, et al. (2013). Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. *NeuroImage*, 65, 511–521.

[17] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., Colliot, O., & et al. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. *NeuroImage*, 56(2), 766–781.

[18] Tong, T., Gray, K., Gao, Q., Chen, L., & Rueckert, D. (2015). Nonlinear graph fusion for multi-modal classification of Alzheimer's disease. In *Machine Learning in Medical Imaging* (pp. 77–84). Springer.

[19] Gray, K. R., Aljabar, P., Heckemann, R. A., Hammers, A., Rueckert, D., Initiative ADN, et al. (2013). Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *NeuroImage*, 65, 167–175.

[20] Liu, M., Zhang, D., & Shen, D. (2014). Hierarchical fusion of features and classifier decisions for Alzheimer's disease diagnosis. *Human Brain Mapping*, 35(4), 1305–1319.

[21] Bissantz, N., Dümbgen, L., Munk, A., & Stratmann, B. (2009). Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces. *SIAM Journal on Optimization*, 19(4), 1828–1845.

[22] Wagner, A., Wright, J., Ganesh, A., Zhou, Z., & Ma, Y. (2009). Towards a practical face recognition system: Robust registration and illumination by sparse representation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 597–604.

[23] Lu, A., Shi, J., & Jia, J. (2013). Online robust dictionary learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 415–422.

[24] Chapelle, O., Schölkopf, B., & Zien, A. (Eds.). (2006). *Semi-supervised learning*. MIT Press.

[25] Zhu, X. (2005). Semi-supervised learning literature survey. *Technical Report 1530*, Department of Computer Sciences, University of Wisconsin-Madison.

[26] Joulin, A., & Bach, F. (2012). A convex relaxation for weakly supervised classifiers. *Proceedings of the International Conference on Machine Learning*, 1279–1286.

Received: August 14 2025