Multimedia Tools and Applications
https://doi.org/10.1007/s11042-024-18532-1

A novel approach to perform linear discriminant analyses for a 4-way alzheimer's disease diagnosis based on an integration of pearson's correlation coefficients and empirical cumulative distribution function

Besma Mabrouk[1] · Ahmed Ben Hamida[2] · Noura Mabrouki[3] · Nouha Bouzidi[4] · Chokri Mhiri[4]

Received: 24 July 2023 / Revised: 25 September 2023 / Accepted: 29 January 2024

Abstract
Diagnosing Alzheimer's disease (AD) remains a significant challenge, particularly in effectively identifying individuals in the early (EMCI) and late (LMCI) stages of Mild Cognitive Impairment (MCI) within the normal control subjects (CN). Leveraging the Alzheimer's Disease Neuroimaging Initiative (ADNI) database and relevant datasets, our aim is to establish a 4-way framework for multi-class diagnosis. Linear Discriminant Analysis (LDA), often coupled with Principal Component Analysis (PCA), has conventionally served as a method for supervised classification. However, this paper introduces an alternative approach using Pearson's correlation coefficient (PCC) instead of PCA. We integrate the optimal LDA subspace with the PCC method, primarily to address the singularity issue that arises when dealing with an underdetermined dataset. Our methodology comprises three main steps. Firstly, we engage in the preprocessing of 237 Diffusion Tensor and Magnetic Resonance brain images to map brain connectivity and extract connections within and between hemispheres. Secondly, we calculate correlation coefficients between features and classes, subsequently constructing empirical cumulative distribution functions (ECDF). Features exceeding a predetermined percentile in the ECDF, guaranteeing the non-singularity of the within-class variance matrix, are subsequently chosen and assessed using a primary classifier. The top k features, linked to the highest classification accuracy, are then mapped into the LDA space through 100 iterations of five-fold Cross-Validation. Following each trial, we assess the performance of six machine learning algorithms, selecting the Logistic Regression classifier to gauge the reliability of our proposed method. As a result, we observed a significant improvement in average

for the Alzheimer's Disease Neuroimaging Initiative. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:(http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

✉ Besma Mabrouk
mabroukbesma@ymail.com

Extended author information available on the last page of the article

accuracy, achieving a performance of 65.46% ± 1.94%, compared to the commonly used PCA+LDA approach, which achieved 50.71% ± 2.1%. Notably, our work achieved 100% accuracy in diagnosing the LMCI class, surpassing other methods. Furthermore, in a separate experiment conducted within and between hemispheres datasets, we identified connectivity between hemispheres as a pivotal biomarker for disease diagnosis in a medical context.

Keywords Alzheimer's disease · Diffusion Weighted Imaging (DWI) · Pearson's correlation coefficients · Linear discriminant analysis (LDA) · Empirical cumulative distribution function (ECDF)

1 Introduction

The increasing proportion of the population aged over 65 has brought attention to multiple neurodegenerative diseases, including Alzheimer's, recognized as the most prevalent and affecting approximately 10 percent of this demographic. This irremediable disease imposes significant implications, both economically for the state and socially for the patient's family and relatives. These reasons have spurred the scientific community to dedicate substantial efforts to diagnosing the disease at its various stages. Structural Magnetic Resonance Imaging (MRI) has made significant contributions to Alzheimer's disease (AD) diagnosis. Notably, in [1], a hybrid classical-quantum machine learning model has effectively discriminated between demented and non-demented subjects with optimal results. Furthermore, this modality has been a focal point in numerous studies aimed at diagnosing the prodromal stage of the disease. For example, the combination of whole-brain T1-weighted MRIs and machine learning algorithms has demonstrated efficiency in discriminating between normal cognition (NC), mild cognitive impairment (MCI), and AD classes [2, 3].

To analyze brain atrophy, mainly cortical thinning, multiple neuroimaging software tools have been employed, utilizing a Voxel-Based Morphometry (VBM) approach for the automatic quantification of volumetric changes in brain regions [4–7].

Among the various regions of interest, special attention has been dedicated to the shape of the hippocampus due to its association with atrophy linked to disease progression. This has spurred researchers to diligently work on the segmentation of the left and right hippocampus [1, 8, 9].

Despite its undeniable efficiency, MRI-based approaches have exhibited two major drawbacks. The first pertains to errors in boundary detection, which is a laborious task, and these errors can substantially affect disease diagnosis. The second concerns its limitations in detecting earlier stages, particularly in distinguishing the EMCI and LMCI classes.

In the last decade, another imaging modality that has emerged is diffusion-weighted imaging (DWI), a non-invasive technique used to measure the Brownian motion of water molecules in brain tissues [10]. Advancements in computer-aided diagnosis systems have enabled the integration of DWI with tractography algorithms to reconstruct brain connectivity. The resulting connectome has been applied in various ways. For instance, one study [11] demonstrated the effectiveness of a network-based approach in predicting diagnoses for NC, MCI, and AD, as well as classifying EMCI and LMCI. In another study [12], topological properties associated with brain organization, such as the weighted clustering coefficient, weighted shortest path length, and betweenness of a node, were combined with node strength and inverse participation ratio to characterize both single and multi-subject data. These features

were subsequently used to evaluate various machine-learning algorithms, with the support vector machine (SVM) model demonstrating the highest performance. Interestingly, when combined with Voxel-Based Morphometry (VBM), these topological properties have also proven effective in a 3-way classification context. Another approach utilizing the connectome is the concept of communicability in the entire brain, providing an alternative to address the gap that arises when relying solely on shortest path-based models [13]. This graph metric has demonstrated its robustness and has also been informative in identifying key regions that play a crucial role in predicting the disease, when tested with various machine-learning models.

One commonly used technique in machine learning that has demonstrated success in the context of multi-class classification is Linear Discriminant Analysis (LDA). LDA serves several important purposes, including dimensionality reduction and multi-class separation. However, in cases with a high feature-to-sample ratio, LDA can encounter singularity issues in the sample covariance matrices. Additionally, LDA's performance tends to degrade when applied to new, unseen data, highlighting a challenge in its ability to generalize. To address these issues and obtain meaningful results, it is crucial to consider appropriate feature selection methods that are suitable for the current dataset. Principal Component Analysis (PCA) has been widely used as a precursor to the LDA technique for dimensionality reduction; however, it has demonstrated limited performance in the context of 4-way classification for Alzheimer's disease [14]

In the current study, we propose the use of Pearson's correlation coefficients for feature selection as an alternative to the PCA method, preceding the application of the LDA method for dimensionality reduction. Our research is based on DW brain images available in the ADNI dataset, aiming to develop a robust diagnostic framework for various disease stages. During the preprocessing phase, we constructed connectivity networks comprising 84 nodes within the entire brain using MRI and DW images. Consequently, we obtained an 84 x 84 connectivity matrix for each patient. However, our focus was primarily on the connectivity patterns within the left hemisphere, within the right hemisphere, and between both hemispheres. Throughout our research, these connectivity datasets were treated as independent entities. In the feature selection stage, we calculate Pearson correlation coefficients to assess the relationship between features and classes. Subsequently, we construct empirical cumulative distribution functions for each dataset. During the search for relevant features, we experiment with different quantiles. We select corresponding percentiles to ensure the non-singularity of the within-class scatter matrices. After identifying the k-relevant features for each dataset using a select k-best module, we carried out a five-fold Cross-Validation procedure over 100 trials. In each trial, the training set was fitted and then projected onto an optimal subspace using the LDA method. Once the dataset's dimensionality was reduced to three modes, they were used to train and test six different classifiers to determine the learning algorithm that provides the best results. It's important to note that the test set was also projected onto the LDA instance specific to each trial. After identifying the optimal classifier, we move forward to assess the performance of the proposed method. To ensure objectivity, we applied all the aforementioned steps to the PCA+LDA method to achieve optimal results. Subsequently, we conducted a comparison, utilizing a statistical test to assess the robustness and reliability of the proposed method in comparison to the original method. Finally, a comparison with other methods is established, and the strengths and limitations of the proposed method are discussed.

2 Related works

The aim of LDA classification is to categorize observations into their respective classes using a set of measurements or predictors. This is achieved by determining an optimal linear transformation that maximizes the distinction between classes [15, 16]. As proposed by Fisher, LDA yields optimal results when the predictors or feature vectors exhibit a multivariate normal distribution within each class and when there is similarity in the covariance among different group classes. However, an inspection of the brain datasets reveals non-stationary characteristics, which, coupled with the issue of feature-to-sample ratio, can mitigate the risk of overfitting. Therefore, previous research based on the LDA approach has placed considerable importance on data preprocessing steps, including feature selection and dimensionality reduction. To address issues related to singularity and overfitting in an Electroencephalogram (EEG)-based Dementia Diagnosis study, the authors utilized Regularized Linear Discriminant Analysis (RLDA), also known as Shrinkage Linear Discriminant Analysis (SLDA) [17]. This technique incorporates a regularization term into the within-class scatter matrix S_W, defined as λI, where I is the identity matrix and λ is the regularization parameter (shrinkage intensity). This implies that the within-group sample covariance is adjusted to a regularized matrix. To strike a balance between reducing overfitting and preserving information in the data, the authors determined the shrinking intensity (λ) through a cross-validation technique referred to in the literature as CV-RLDA. Another regularization technique that aids in preventing overfitting and facilitating automatic feature selection is the Least Absolute Shrinkage and Selection Operator, commonly referred to as the LASSO method. It has found extensive application in the detection of Alzheimer's disease. For example, in the context of Alzheimer's disease classification based on PET brain images, researchers have introduced an adaptive Linear Discriminant Analysis (LDA) approach. This method incorporates a penalty function that combines aspects of both Lasso (L_1 norm penalty) and group Lasso (L_2 norm penalty) for variable selection [18].

For the same purpose, several other researchers have chosen to integrate the Linear Discriminant Analysis (LDA) with Principal Component Analysis (PCA) method. For instance, in a study referenced as [19], the authors initially employed a two-sample t-test to extract features from anatomical MRI images in conjunction with Mini-Mental State Examination (MMSE) scores. Following this feature selection process, the resulting subset was fed into the Kernel Principal Component Analysis (KPCA) module. Within the KPCA module, the data is projected onto the principal component coefficients within a higher-dimensional kernel space, a step that aims to enhance linear separability. Subsequently, in the same study, LDA method was applied to project the KPCA coefficients into a more effective linear discriminant space. Finally, the researchers employed a multi-kernel Support Vector Machine (SVM) to complete the 3-way classification task.

Due to its ability to capture the most variance within a dataset, the PCA method has also been employed in a context involving 4-way classification [14]. In the referenced article, multimodal data were preprocessed and subsequently scored through projection onto the first linear discriminant analysis (LDA) vector. These scores, representing the progression of pathology, were utilized to develop a multiclass Alzheimer's Disease (AD) diagnosis framework based on an extreme learning machine (ELM)-based decision tree. However, the achieved results appear to be limited, especially in terms of accuracy, which suggests the need for further improvement of the LDA method.

3 Materials and methods

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://www.adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The ADNI community provides diffusion-weighted imaging (DWI) for GE MRI sessions from ADNI2 and all MRI sessions from ADNI3. Images were deliberately chosen from various manufacturers to guarantee the inclusion of three-dimensional (3D) T1-weighted (T1W) imaging and two-dimensional echo planar DWI. In this study, we have thoroughly analyzed a total of 237 MRI and DWI brain images, and the accompanying clinical and demographic information for the subjects can be found in Table 1.

3.1 Image pre-processing

The downloaded DICOM brain images are converted into NIFTI files using the Heudiconv software and organized into structured directory layouts (ANAT, DWI) in compliance with the Brain Imaging Data Structure (BIDS) format. Diffusion and anatomical images are automatically preprocessed using a combination of software packages, including MRTRIX3 (http://www.mrtrix.org), FreeSurfer (http://surfer.nmr.mgh.harvard.edu/), Advanced Normalization Tools (ANTs) http://stnava.github.io/ANTs/), and the FMRIB Software Library (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). Following recent advancements in the field [4, 20], the preprocessing involves multiple steps:

First, denoising is performed [21]. The b0 images are extracted from the diffusion data acquired in the anterior-to-posterior (AP) direction. Eddy current corrections are applied to the phase-encoding AP direction of the b0 images using the FSL command. Subsequently, bias field correction and skull stripping are carried out using Advanced Normalization Tools (ANTs).

Second, various basis functions are estimated for each tissue type (WM, GM, and CSF) to conduct multi-shell multi-tissue constrained spherical deconvolution and generate an image of fiber orientation densities (FOD) overlaid on the estimated tissues. The FODs are normalized to enable inter-subject comparisons. It's important to note that, in our experimentation, normalization is restricted to 2-tissue (WM, CSF) for 2-shells (b=0 and b=1000) DW Images.

Table 1 Clinical and demographic information

Class	Gender	Age(mean)	Size
NC [1]	27 M/35 F	72.25	64
EMCI[2]	34 M /28 F	77.78	64
LMCI[3]	31 M/26 F	77.51	59
AD[4]	22 M/26 F	74.68	51

[1] NC normal control.
[2] EMCI Earlier mild cognitive impairment.
[3] LMCI Later mild cognitive impairment.
[4] AD Alzheimer's disease

Third, a GM/WM boundary is created for seed analysis. This is achieved by converting the anatomical image to MRTRIX3 format, segmenting it into five tissue categories (1=GM; 2=Subcortical GM; 3=WM; 4=CSF; 5=Pathological tissue), co-registering the averaged b0 diffusion images, and finally creating a boundary that separates grey matter from white matter.

Finally, streamlines are generated using MRTRIX3's default probabilistic tractography approach and are subsequently refined. These streamlines are utilized to create a weighted symmetric connectivity matrix, denoted as W (84x84), where 84 represents the number of Regions of Interest (ROIs). These ROIs consist of 42 parcellations for each hemisphere, which have been obtained through the recon-all command from FreeSurfer, following the Desikan-Killiany atlas.

The elements of the connectivity matrix, denoted as $W_{i,j}$, represent the strength of connectivity between nodes. This strength is determined through the normalization of the number of fibers connecting the i^{th} and j^{th} nodes [22].

3.2 Feature selection

The connectivity matrix provides a description of the weighted graph [4]. Specifically, each element $W_{i,j}$ of the connectome represents the normalized weight between two nodes, where $0 \leq W_{i,j} \leq 1$.

From each connectome, three sub-matrices were extracted: the first two represent interconnections within the left and right hemispheres, while the third pertains to connections between hemispheres. Since these sub-matrices are symmetric, we propose removing the upper triangle and diagonal elements and flattening the remaining part into an n-dimensional vector (Fig. 1).

In summary, we obtained three m by n matrices, where m is the number of subjects in the dataset, and n represents the number of sites in each triangle, determined as follows:

$$n = \sum_{j=0}^{42} j = \frac{j\,(j+1)}{2} \tag{1}$$

Where j represents the site of the element W in the upper triangle.

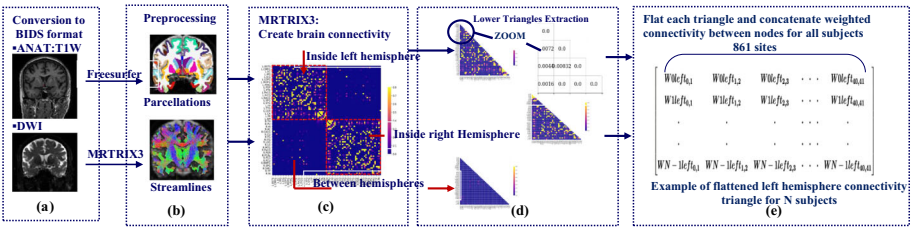

Fig. 1 The schematic overview of the proposed method: (a) Convert downloaded brain image to BIDS format, (b) Preprocessing of ANAT and DWI images, (c) Create the connectome, (d) Extract the weighted connectivity inside and between hemispheres, (e) Create three matrices, representing weighted connections in left, right and between hemispheres, by flatting triangles for each subject, then concatenating the required data of N subjects in respect of axis 0

3.2.1 Pearson's correlation coefficients

In accordance with M. Grana and co-authors [23], Pearson's correlation coefficients are computed for each vector with respect to the class labels $y_i = 0, 1, 2, 3$, representing normal control, EMCI, LMCI, and AD patients, respectively. The matrices described above are treated independently in this process, and the Pearson's correlation of the j^{th} vector is calculated as follows:

$$r_{v_{j,y}} = \frac{\sum_{i=0}^{n} v_{i,j}\, y_i - \sum_{i=0}^{n} v_{i,j} \sum_{i=0}^{n} y_i}{\sqrt{n \sum_{i=0}^{n} v_{i,j}^2 - (\sum_{i=0}^{n} v_{i,j})^2} \sqrt{n \sum_{i=0}^{n} y_i^2 - (\sum_{i=0}^{n} y_i)^2}} \tag{2}$$

Here, the vector V_j represents the weighted connections between nodes at the j^{th} site across all subjects, where $V_{i,j}$ represents the value of the j^{th} vector for the i^{th} subject, y_i represents the class to which the i^{th} subject belongs, and n is the total number of subjects.

After obtaining vectors that summarize the correlations between sites and classes, we reconstruct the empirical cumulative distribution functions. Here's how it's done:

First, we sort the absolute correlation values. Next, we scale the x-axis from the minimum to the maximum value with a step size of one per n. Finally, we construct the y-axis in such a way that each point on the x-axis is associated with the ratio of the cumulative number of immediate predecessors added by one to the set cardinality n.

As depicted in Fig. 2, we begin by randomly choosing a first percentile. Then, we select the voxel sites whose absolute correlation values have a cumulative frequency above this percentile. Following the formula in the next section, we compute the within-class variance matrix and check for singularity. This procedure is repeated until a non-singular matrix is obtained.

3.2.2 Linear discriminant analysis

Fisher Linear Discriminant Analysis (LDA) remains one of the standard methods used for supervised classification. LDA involves the linear projection of features onto an optimal subspace. This new space is designed to ensure maximum separation between classes and minimum intra-class variability. To obtain the LDA features represented by the matrix W, we must solve the generalized Rayleigh quotient:

$$W = \arg\max_{W} \frac{W^T S_B W}{W^T S_W W} \tag{3}$$

Where W represents the transformation matrix. The determination of W is detailed below.

For four-class classification, the between-class and within-class scatter matrices are computed using the following mathematical formula:

$$S_B = \frac{1}{4} \sum_{i=0}^{3} (\mu - \mu_i)(\mu - \mu_i)^T \tag{4}$$

$$S_W = \sum_{i=0}^{4} \sum_{j=0}^{n} (x_{i,j} - \overline{x_{i,j}})(x_{i,j} - \overline{x_{i,j}})^T \tag{5}$$

Where μ is the overall mean, μ_i is the mean of the i^{th} class, n is the number of selected features, $x_{i,j}$ is the j^{th} sample in the i^{th} class, and $\overline{x_{i,j}}$ is its corresponding mean.

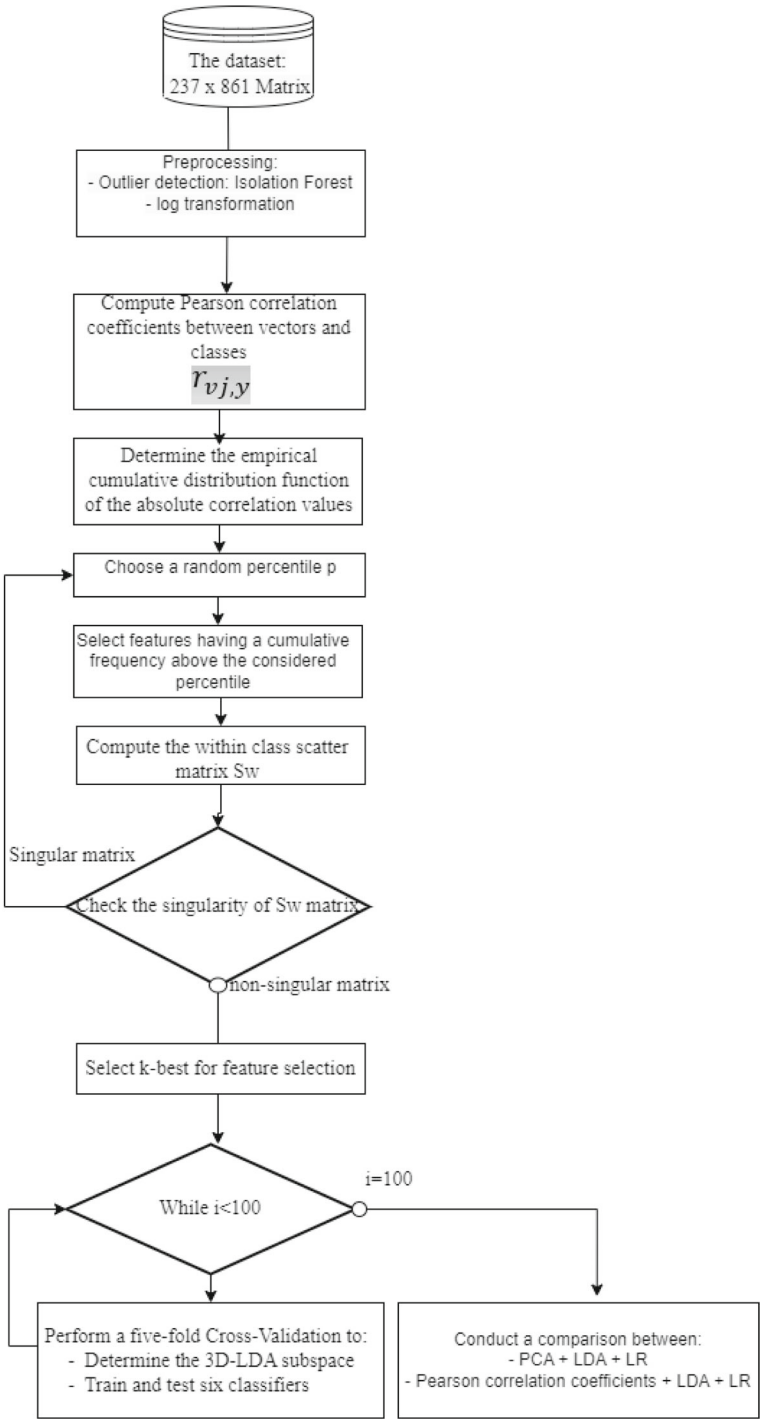

Fig. 2 Flow chart resuming the proposed methodology

The solution to the generalized Rayleigh quotient is reduced to the eigenvalue decomposition, as detailed below:

$$S_B W = \lambda S_W W \tag{6}$$

Where λ is the eigenvalue. Assuming that S_W is a non-singular matrix, we transpose the within-class variance matrix, simplifying equation (6) to:

$$S_W^{-1} S_B W = \lambda W \tag{7}$$

Since the matrix $S_W^{-1} S_B$ has no more than n -1 non-zero eigenvalues, we ultimately obtain three distinct eigenvectors.

Throughout 100 trials, a five-fold cross-validation is conducted. In each trial, the eigenvector matrix W, which has been previously determined, is used to transform the samples into a new subspace. Specifically, both the Xtrain and Xtest datasets are projected onto the three obtained modes of space W using the following formulas:

$$F(X_{train}) = W^T \cdot X_{train} \tag{8}$$

$$F(X_{test}) = W^T \cdot X_{test} \tag{9}$$

For each partition, a new 3-D dataset is created by combining the 3D features from the training and test sets, as determined by Equation (10):

$$F_{LDA} = W^T \cdot X_{train} \cup W^T \cdot X_{test} \tag{10}$$

It's important to note that, for each trial, classification models are trained using LDA features from the training set and then evaluated using LDA features from the test set.

In summary, The diagram illustrating the proposed methodology is presented in Fig. 2

4 Experimental results and discussion

In the conducted experimentation, we followed the chart proposed in the last section. The initial datasets underwent preprocessing primarily due to the potential influence of extreme values on the covariance matrix and, consequently, Pearson's correlation coefficient. To address this issue, we began by detecting outliers using the Isolation Forest Python module. After identifying and removing the outliers, our analysis proceeded with 229 subjects out of the original 237. Furthermore, the application of the LDA method assumes that the data adheres to a normal distribution. To verify this assumption, we applied log transformation, a widely accepted data preprocessing technique in both biomedical and statistical contexts [24].

Following that, we computed Pearson's correlation and class assignments. Next, we calculated the empirical cumulative distribution functions of the absolute correlation values. Subsequently, we randomly selected a percentile, and voxel sites whose absolute correlation values exceeded this percentile were selected. We then computed the within-class variance matrix and checked for singularity. This process was repeated iteratively until a non-singular matrix was obtained.

The empirical selection of a percentile, followed by experimentation of the proposed method, effectively determined a percentile of 80% for extracting relevant features. The number of sites was reduced from n=861 to 17, representing the left hemisphere, to only six for the right hemisphere, and to 31 features concerning the connectivity between hemispheres.

It's worth noting that, to enhance our results, we utilized the Select K-Best Python module to identify the top k features associated with the highest classification accuracies when tested with a primary classifier. As shown in Fig. 3, we found that when using the logistic regression algorithm, 22 selected features from the between-hemispheres dataset contributed to achieving the highest accuracy. Similarly, we selected 8 features for the left dataset and 5 features for the right dataset.

When considering the application of the LDA technique, it's crucial to recognize that LDA is a supervised learning method, and this introduces specific considerations. The conventional approach of using an 80 to 20 training-to-testing ratio, as described in prior work [14], can potentially lead to overfitting and impact the reliability of clinical trials. Therefore, in our current study, we adopted a repeated five-fold Cross-Validation procedure to ensure the robustness of our proposed method. The essence of this procedure involves dividing the dataset into five parts, with four allocated for training and the remaining portion reserved for validation. We conducted this process over a hundred trials. During each iteration, we determined the LDA subspace based on the training set and subsequently projected the validation samples onto this subspace.

After reducing the dataset to three dimensions, we employed six distinct classifiers for to determine the most effective learning algorithm. It's worth noting that, in each trial, we projected the test set onto its respective LDA instance. The experimented classifiers included Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (NB), Extra Tree (ET), and the Multi-Layer Perceptron (MLP). The achieved accuracies are presented in Fig. 4, with the highest value obtained by the logistic regression classifier highlighted with a red star. Importantly, the dataset used for this figure pertains to between-hemispheres connectivity.

To assess the robustness of the proposed method, we tested it on multiple datasets: brain connectivity within the left hemisphere data, brain connectivity within the right hemisphere dataset, and brain connectivity between hemispheres dataset. Additionally, we conducted a comparison with PCA+LDA. The results obtained using the logistic regression classifier are presented in Fig. 5.

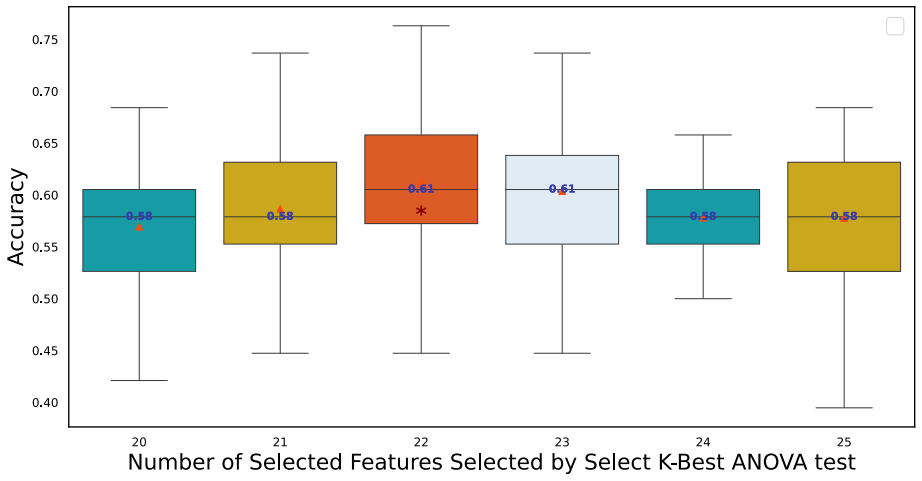

Fig. 3 Variation in accuracies using logistic regression classifier with multiple sets of top-k highest-ranked features selected by the SelectKBest Python module based on F-statistics ANOVA test

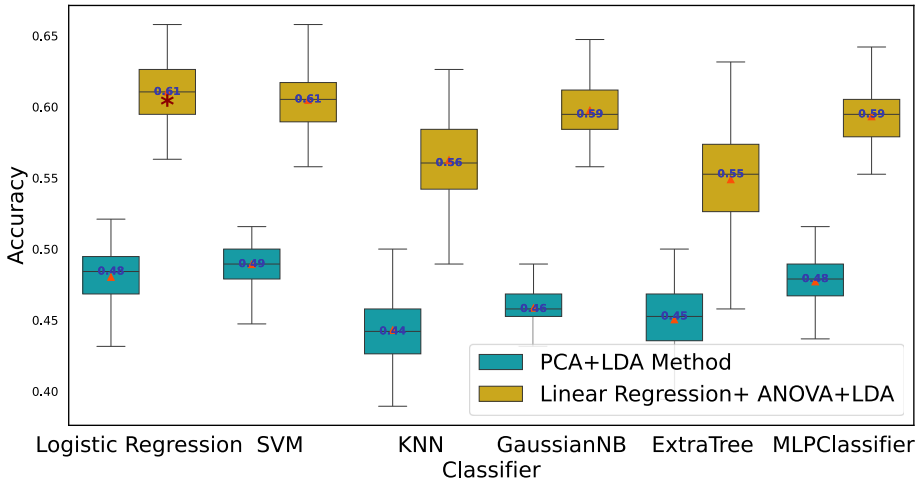

Fig. 4 A comparison of accuracies: PCA+LDA vs. pearson correlation+LDA feature selection approaches with Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (NB), Extra Tree (ET), and Multi-Layer Perceptron (MLP) Classifiers

4.1 Comparison with other methods

To the best of our knowledge, the use of Diffusion-weighted imaging (DWI) for a 4-way disease diagnosis has not been definitively established. Furthermore, a notable deficiency in research about the discrimination of all four classes of Alzheimer's has also been observed. A review of prior research [4, 12, 25] reveals a common tendency to dissect multiclass diagnosis into binary classification cases. This approach may compromise the reliability and robustness of proposed methods in clinical assessments. Therefore, we conducted experiments to compare the performance of our proposed method with state-of-the-art techniques, regardless

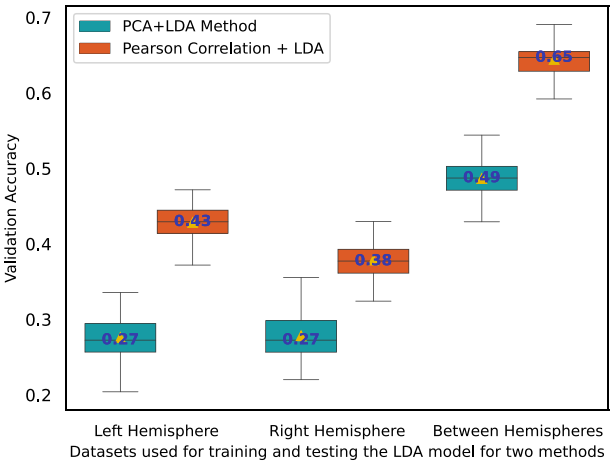

Fig. 5 An evaluation of accuracies: proposed method vs. original method on three datasets (left hemisphere, right hemisphere, and between hemispheres) using logistic regression classifier

Table 2 Comparison of classification performances with five pertaining studies conducted among the ADNI dataset for 4-way classification

Multiclass performances	Accuracy (%)	F1-score (%)	Sensitivity (%)	Specificity (%)
Ruiz et al. [29] 3D DenseNets Ensemble (4-way)	83.33%	_	_	_
Ghazal TM, Issa G [30] MRI+Transfer learning (4-way)	91.7%	93.7%	96%	_
Yao et al. [26] MRI+freesurfer5.3 + Ensemble methods (4-way)	_	54,38%	55.25%	_
Liu et al. [21] MRI and PET + SAE-ZEROMASK + Deep Learning (4-way)	53.79 ± 4.76%	_	52.14 ± 11.81	
Lin et al. [14] Mmltimodal data + LDA + ELM (4-way)	57.3± 0.9%	55.7± 1%	_	_
The proposed method (4-way)	65.98±1.81%	65.99±1.89%	66.03±1.86%	66.06±1.85%

of the used imaging modality. Table 2 provides an overview of the performance recorded in various relevant studies conducted within the ADNI dataset, with a focus on multiclass differentiation.

A careful examination of the results presented in Table 2 reveals that our proposed method has consistently achieved the highest levels of performance, outperforming traditional classification algorithms [21, 26, 27] It is worth noting that machine learning-based approaches [28–30] have demonstrated superior accuracy in the context of multiclass classification. However, we maintain that our method remains both competitive and promising, especially in terms of its predictive accuracy for the LMCI class. This holds significant implications for enabling earlier diagnoses in clinical settings, where timely intervention can yield substantial benefits.

4.2 Discussion

The observation of the results presented in Fig. 4 reveals that when utilizing the between-hemisphere dataset, the proposed feature selection and dimensionality reduction method significantly enhances accuracy compared to the PCA + LDA method, regardless of the classifier employed. Notably, when employing the Pearson correlation + LDA method for feature selection and dimensionality reduction, it becomes evident that the accuracies achieved by Logistic Regression (LR), Support Vector Machine (SVM), Gaussian Naïve Bayes (NB), and the Multi-Layer Perceptron (MLP) far surpass those attained by K-Nearest Neighbors (KNN) and Extra Tree (ET), which generally do not exceed 60%. This difference in performance can be attributed to the strength of Pearson correlation in detecting linear relationships between features and the target variable. When Pearson correlation is coupled with the effective linear separability offered by LDA, especially when used alongside linear classifiers such as Logistic Regression (LR), Support Vector Machine (SVM) with linear kernels, and Gaussian Naïve Bayes (NB), the resulting enhancement in accuracy becomes apparent.

In contrast, when comparing with K-Nearest Neighbors (KNN) and Extra Trees (ET), both of which are adept at modeling nonlinear relationships in data, the linear classifiers prove to be more effective in this specific context.

From Fig. 5, it becomes evident that the proposed method significantly enhances LDA results across various datasets. To assess the reliability of the mean differences between the compared approaches, we conducted an independent samples t-test. A p-value less than 0.05 is considered statistically significant, leading to the rejection of the null hypothesis of equal means and indicating robustness and reliability. The results are presented in Fig. 6, where it is apparent that the proposed approach exhibits significant improvements in terms of accuracy, precision, sensitivity, F1_score, and specificity when compared to the commonly used original method.

Importantly, from the same figure, we observe that the connectivity between hemispheres' nodes has proven to be effective for multi-class prediction This raises the question of whether it can be considered a potential biomarker for disease diagnosis. To delve deeper into this, we conducted a statistical test to examine potential relationships among various disease classes and specific datasets. The results of this analysis are displayed in Fig. 7.

Upon visual inspection, it becomes apparent that, except for the EMCI class, there is already a statistically significant difference, indicating the effectiveness of discrimination among various disease stages based on the connections between hemispheres (B.H) when compared to those within the left (L.H) and the right (R.H) hemispheres. The lack of significance (p−value>0.05), as revealed when comparing the provided accuracy based on R.H vs. B.H for the diagnosis of the early-stage EMCI, suggests that changes have likely occurred earlier within the right hemisphere.

Concerning the comparison of left vs. right hemisphere performance, a highly significant difference (p-value$\leq 1.00e^{-04}$) is observed between the CN and EMCI cohorts. In contrast, no significant difference has been found concerning the LMCI and AD classes. Consequently, in alignment with the findings from a region-of-interest volumetry-based anatomical study [4], we conclude that there is a notable association between changes in brain connectivity within the left hemisphere and disease progression.

Significantly, the dataset depicting the connections between hemispheres' nodes played a crucial role in diagnosing the LMCI class, where an impressive accuracy of 100% was attained. This underscores the substantial contribution of our work to this field.

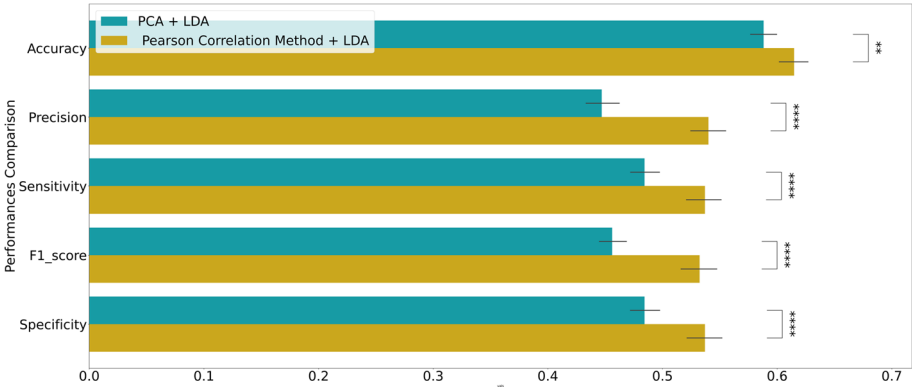

Fig. 6 Assessing the Robustness and Reliability of the Proposed Approach: A Paired t-Test Between PCA + LDA and Pearson Correlation + LDA in Evaluating Accuracy, Precision, Sensitivity, F1 Score, and Specificity Performances where; ns: no significance p-value $\leq 1.00\ e^{00}$, *: $1.00\ e^{-02} <$ p−value $\leq 5.00e^{-02}$, **: $1.00\ e^{-03} <$ p−value$\leq 1.00e^{-02}$,***: $1.00\ e^{-04} <$ p−value$\leq 1.00e-03$

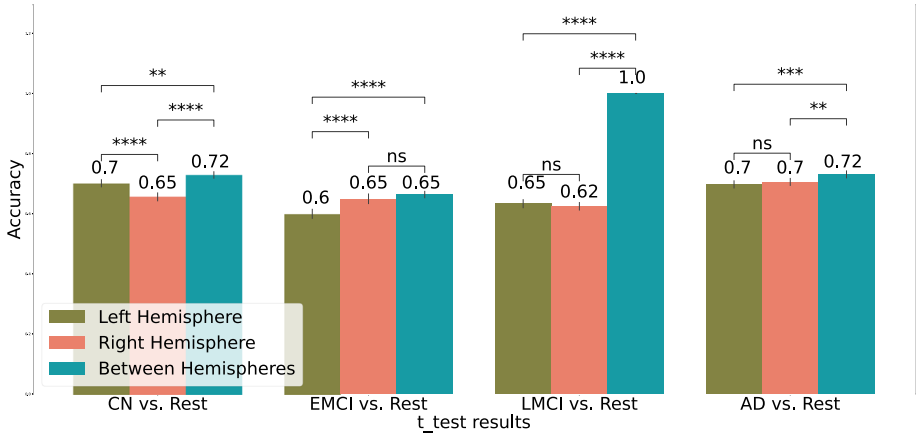

Fig. 7 Registered accuracies for each class CN, EMCI, LMCI, and AD, along with statistical significance in three-sample t-tests between left, right, and between hemispheres feature groups where; ns: no significance p-value $\leq 1.00\, e^{00}$, *: $1.00\, e^{-02} <$ p−value $\leq 5.00e^{-02}$, **: $1.00\, e^{-03} <$ p−value$\leq 1.00e^{-02}$,***: $1.00\, e^{-04} <$ p−value$\leq 1.00e{-}03$

Despite the undeniable role played by the proposed approach in improving LDA results, the achieved results are still considered limited. Several factors contribute to these limitations.

One key limitation is related to the sensitivity of Pearson correlation coefficients to outliers or extreme values in the data. When extreme values are present, they can unduly influence the correlation coefficient and potentially lead to inaccurate feature selection. While we attempted to address this drawback through outlier detection, our efforts were constrained by the small dataset size.

Additionally, the relatively small dataset presents challenges for the generalization of the model. The limited amount of data may restrict the model's ability to perform well on unseen examples. The absence of external validation data further limits the model's capacity to generalize effectively. These limitations have a noticeable impact on both precision and recall, ultimately affecting the F1 score.

Furthermore, the preprocessing steps required for diffusion-weighted images to create the connectome are exceptionally time-consuming. This time-intensive process hinders our ability to enrich the dataset with a larger volume of data, which is often beneficial for deep learning methods. Consequently, the feasibility of employing deep learning techniques is constrained by the dataset's size and the time required for preprocessing.

5 Conclusion

In this paper, we present a novel approach for diagnosing Alzheimer's disease using the Linear Discriminant Analysis (LDA) method within a 4-way classification framework. Our method integrates Pearson's correlation coefficient and empirical cumulative distribution function to perform feature selection and dimensionality reduction. The selection of an appropriate percentile is crucial to ensure the non-singularity of the within-class variance matrix. To evaluate the model's performance robustly, we employ repeated five-fold cross-validation. This approach allows us to compute 3D-LDA features and train multiple classifiers, ensuring that the model performs effectively on various subsets of the data while avoiding overfitting

of the training data, thereby enhancing its credibility for clinical trials. Experimental results demonstrate that our proposed method outperforms the traditional PCA+LDA approach, despite the use of smaller datasets. While our method achieves slightly lower average accuracy compared to machine learning methods employing extensive datasets, it remains promising for multi-class diagnosis. Importantly, our approach is not limited to Alzheimer's disease diagnosis and can be applied to various multi-class classification scenarios. Furthermore, our work highlights the potential of interhemispheric node connectivity as a valuable biomarker for Alzheimer's disease diagnosis.

Acknowledgements 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). 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

Author Contributions Author contributions included conception and study design (BM,ABH), statistical analysis (BM, NM, and ABH), interpretation of results (BM, NM,NB, and CM), drafting the manuscript work or revising it critically for important intellectual content (BM, NM) and approval of final version to be published and agreement to be accountable for the integrity and accuracy of all aspects of the work (All authors).

Funding Not applicable

Availability of data and materials Not applicable

Code Availibility Not applicable

Declarations

Conflict of interest/Competing interests The authors declare no competing interests.

Ethics approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

References

1. Shahwar T, Zafar J, Almogren A, Zafar H, Rehman AU, Shafiq M, Hamam H (2022) Automated detection of alzheimer's via hybrid classical quantum neural networks. Electronics 11(5):721
2. Hazarika RA, Kandar D, Maji AK (2023) A novel machine learning based technique for classification of early-stage alzheimer's disease using brain images. Multimedia Tools and Applications:1–23

3. Fouladi S, Safaei AA, Arshad NI, Ebadi M, Ahmadian A (2022) The use of artificial neural networks to diagnose alzheimer's disease from brain images. Multimedia Tools and Applications 81(26):37681–37721
4. Mabrouk B, BenHamida A, Drissi N, Bouzidi N, Mhiri C (2023) Contribution of brain regions asymmetry scores combined with random forest classifier in the diagnosis of alzheimer's disease in his earlier stage. Journal of Medical and Biological Engineering 43(1):74–82
5. Hirata Y, Matsuda H, Nemoto K, Ohnishi T, Hirao K, Yamashita F, Asada T, Iwabuchi S, Samejima H (2005) Voxel-based morphometry to discriminate early alzheimer's disease from controls. Neurosci Lett 382(3):269–274
6. Busatto GF, Diniz BS, Zanetti MV (2008) Voxel-based morphometry in alzheimer's disease. Expert Rev Neurother 8(11):1691–1702
7. Schmitter D, Roche A, Maréchal B, Ribes D, Abdulkadir A, Bach Cuadra M, Daducci A, Granziera C, Klöppel S, Maeder P et al (2015) An evaluation of volume-based morphometry for prediction of mild cognitive impairment and alzheimer's disease. NeuroImage: Clinical 7:7–17
8. Liang C, Lao H, Wei T, Zhang X (2022) Alzheimer's disease classification from hippocampal atrophy based on pcanet-bls. Multimedia Tools and Applications 81(8):11187–11203
9. Baskar D, Jayanthi V, Jayanthi A (2019) An efficient classification approach for detection of alzheimer's disease from biomedical imaging modalities. Multimedia Tools and Applications 78:12883–12915
10. Takahashi H, Ishii K, Kashiwagi N, Watanabe Y, Tanaka H, Murakami T, Tomiyama N (2017) Clinical application of apparent diffusion coefficient mapping in voxel-based morphometry in the diagnosis of alzheimer's disease. Clin Radiol 72(2):108–115
11. La Rocca M, Amoroso N, Monaco A, Bellotti R, Tangaro S, Initiative ADN et al (2018) A novel approach to brain connectivity reveals early structural changes in alzheimer's disease. Physiol Meas 39(7):074005
12. Mabrouk B, Bouzidi N, Mhiri C, Hamida AB (2022) Combination of volumetric and topologic brain characteristics towards a diagnosis of alzheimer's disease in his earlier stage. In: 2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp 1–4. IEEE
13. Lella E, Lombardi A, Amoroso N, Diacono D, Maggipinto T, Monaco A, Bellotti R, Tangaro S (2020) Machine learning and dwi brain communicability networks for alzheimer's disease detection. Appl Sci 10(3):934
14. Lin W, Gao Q, Du M, Chen W, Tong T (2021) Multiclass diagnosis of stages of alzheimer's disease using linear discriminant analysis scoring for multimodal data. Comput Biol Med 134:104478
15. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7(2):179–188
16. Omar T, Wassim Z, Mohamed BM (2014) Brain-computer interface: frequency domain approach using the linear and the quadratic discriminant analysis. In: 2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp 346–349. IEEE
17. Neto E, Biessmann F, Aurlien H, Nordby H, Eichele T (2016) Regularized linear discriminant analysis of eeg features in dementia patients. Frontiers in aging neuroscience 8:273
18. Le KT, Chaux C, Richard FJ, Guedj E (2020) An adapted linear discriminant analysis with variable selection for the classification in high-dimension, and an application to medical data. Computational Statistics & Data Analysis 152:107031
19. Alam S, Kwon GR, Initiative ADN (2017) Alzheimer disease classification using kpca, lda, and multi-kernel learning svm. Int J Imaging Syst Technol 27(2):133–143
20. Wu Z, Gao Y, Potter T, Benoit J, Shen J, Schulz PE, Zhang Y, Initiative ADN (2021) Interactions between aging and alzheimer's disease on structural brain networks. Frontiers in aging neuroscience 13:639795
21. Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ et al (2014) Multimodal neuroimaging feature learning for multiclass diagnosis of alzheimer's disease. IEEE Trans Biomed Eng 62(4):1132–1140
22. Amoroso N, Monaco A, Tangaro S (2017) Topological measurements of dwi tractography for the alzheimer's disease detection. Computational and Mathematical Methods in Medicine in press. https://doi.org/10.1155/2017/5271627
23. Graña M, Termenon M, Savio A, Gonzalez-Pinto A, Echeveste J, Pérez J, Besga A (2011) Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by pearson's correlation. Neurosci Lett 502(3):225–229
24. Feng C, Wang H, Lu N, Chen T, He H, Lu Y, Tu X (2014) Log transformation and its implications for data analysis. Shanghai Arch Psychiatry 26:105–109
25. Tufail AB, Anwar N, Othman MTB, Ullah I, Khan RA, Ma YK, Adhikari D, Rehman AU, Shafiq M, Hamam H (2022) Early-stage alzheimer's disease categorization using pet neuroimaging modality and convolutional neural networks in the 2d and 3d domains. Sensors 22(12):4609
26. Yao D, Calhoun VD, Fu Z, Du Y, Sui J (2018) An ensemble learning system for a 4-way classification of alzheimer's disease and mild cognitive impairment. J Neurosci Methods 302:75–81
27. Liu M, Zhang J, Adeli E, Shen D (2018) Joint classification and regression via deep multi-task multi-channel learning for alzheimer's disease diagnosis. IEEE Trans Biomed Eng 66(5):1195–1206

28. Oktavian MW, Yudistira N, Ridok A (2022) Classification of alzheimer's disease using the convolutional neural network (cnn) with transfer learning and weighted loss. arXiv:2207.01584
29. Ruiz J, Mahmud M, Modasshir M, Shamim Kaiser M, Alzheimer's Disease Neuroimaging Initiative ft (2020) 3d densenet ensemble in 4-way classification of alzheimer's disease. In: Brain Informatics: 13th International Conference, BI 2020, Padua, Italy, September 19, 2020, Proceedings 13,pp. 85–96. Springer
30. Ghazal TM, Issa G (2022) Alzheimer disease detection empowered with transfer learning. Computers, Materials & Continua 70(3):5005–5019

Authors and Affiliations

Besma Mabrouk[1] · Ahmed Ben Hamida[2] · Noura Mabrouki[3] · Nouha Bouzidi[4] · Chokri Mhiri[4]

1 Advanced Technologies for Medicine and Signals ATMS, Department of Electrical and Computer Engineering, National Engineers School, Sfax, Tunisia

2 Department IS, College of Computer Science, King Khaled Universiy, Abha, Saudi Arabia

3 Department of Psychiatry, University Hospital Taher Sfar, Mahdia, Tunisia

4 Department of Neurology, University Hospital Habib Bourguiba, Sfax, Tunisia