Discover Applied Sciences

Research

Early diagnosis of Alzheimer's disease and mild cognitive impairment using MRI analysis and machine learning algorithms

Helia Givian[1,2] · Jean-Paul Calbimonte[1,2] · and for the Alzheimer's Disease Neuroimaging Initiative

Received: 3 July 2024 / Accepted: 12 December 2024
Published online: 18 December 2024

Abstract
Early diagnosis of Alzheimer's disease (AD) and mild cognitive impairment (MCI) is crucial to prevent their progression. In this study, we proposed the analysis of magnetic resonance imaging (MRI) based on features including; hippocampus (HC) area size, HC grayscale statistics and texture features (mean, standard deviation, skewness, kurtosis, contrast, correlation, energy, homogeneity, entropy), lateral ventricle (LV) area size, gray matter area size, white matter area size, cerebrospinal fluid area size, patient age, weight, and cognitive score. Five machine learning classifiers; K-nearest neighborhood (KNN), support vector machine (SVM), random forest (RF), decision tree (DT), and multi-layer perception (MLP) were used to distinguish between groups: cognitively normal (CN) vs AD, early MCI (EMCI) vs late MCI (LMCI), CN vs EMCI, CN vs LMCI, AD vs EMCI, and AD vs LMCI. Additionally, the correlation and dependence were calculated to examine the strength and direction of association between each extracted feature and each classification of the group. The average classification accuracies in 20 trials were 95% (SVM), 71.50% (RF), 82.58% (RF), 84.91% (SVM), 85.83% (RF), and 85.08% (RF), respectively, with the best accuracies being 100% (SVM, RF, and MLP), 83.33% (RF), 91.66% (RF), 95% (SVM, and MLP), 96.66% (RF), and 93.33% (DT). Cognitive scores, HC and LV area sizes, and HC texture features demonstrated significant potential for diagnosing AD and its subtypes for all groups. RF and SVM showed better performance in distinguishing between groups. These findings highlight the importance of using 2D-MRI to identify key features containing critical information for early diagnosis of AD.

Article Highlights

- Cognitive scores, brain structure sizes, and tissue features can assist in diagnosing Alzheimer's and its early stages.
- Machine learning models classify Alzheimer's stages using optimized brain MRI features.
- MRI scans show how brain features change as Alzheimer's progresses.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://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: (https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42452-024-06440-w.

✉ Helia Givian, helia.givian@hevs.ch; Jean-Paul Calbimonte, jean-paul.calbimonte@hevs.ch | [1]Institute of Informatics, University of Applied Sciences and Arts Western Switzerland (HES-SO Valais Wallis), TechnoPole 3, 3960 Sierre, Valais, Switzerland. [2]The Sense Innovation and Research Center, Avenue de Provence 82, 1007 Lausanne, Vaud, Switzerland.

Keywords Alzheimer's disease (AD) · Mild Cognitive impairment (MCI) · Magnetic resonance imaging (MRI) · Computer-aided diagnosis (CAD) · Machine learning (ML)

1 Introduction

Dementia is a set of symptoms consisting of cognitive impairment and brain function disorders "Dementia statistics, Alzheimer's Disease International" [1]. The most common disease among adults over the age of 65 is Alzheimer's disease (AD) which is a subset of dementia [2]. Alzheimer's is a neurodegenerative disorder that causes neuronal death and brain tissue loss. The early stage of AD is mild cognitive impairment (MCI) which gradually progresses and can generally be divided into two stages; early mild cognitive impairment (EMCI) and late mild cognitive impairment (LMCI) [3]. Although not all MCI patients get converted into AD [4], MCI diagnosis is useful in predicting AD, with about 15% of MCI patients converting to AD every year [5] (refer to Appendix A for more information).

The process of diagnosing Alzheimer's demands substantial knowledge and understanding to differentiate AD patients from healthy and MCI individuals by analyzing visible variations in brain regions. There are different methods to diagnose AD. The most common method to identify biomarkers associated with AD and cognitive impairment, as well as to assess changes in the brain caused by the disease is magnetic resonance imaging (MRI) modality. MRI is a non-invasive imaging method that uses the magnetic properties of the body and produces 2D and 3D images of any part of the body with high resolution. In general, humans are not able to detect abnormal patterns correctly or recognize the special characteristics related to AD simply and with more accuracy [6]. So, Computer-aided diagnosis (CAD) systems can provide better diagnostic suggestions by analyzing these patterns or any special changes in the brain efficiently. In the CAD system in recent decades, machine learning (ML) algorithms have played a crucial role in the field of AD prediction or diagnosis [7]. Various artificial intelligence (AI) approaches based on ML and Deep Learning (DL) have been investigated in brain disorders using MRI applications. In recognition of AD biomarkers, research has demonstrated that the ML methods have performed better than DL algorithms even with smaller databases [8]. In various medical applications, methods based on ML are often preferred over DL techniques because they can better detail the factors influencing the disease and the changes created. Despite the DL algorithm's advantage in handling complex data such as MRI scans, ML remains a practical and straightforward approach to implement, providing detailed insights into the causes of Alzheimer's and its subtypes [8].

The study in [9], introduced an approach to identify AD based on the application of ML techniques. They used histograms to transform brain images into feature vectors, containing the relevant brain features, which later served as the inputs in the classification step. They suggested the random forest (RF) classifier for discriminating the AD subjects from the control subjects, which achieved an accuracy rate of 85.77%.

In the research conducted by [10], a deep learning multi-layer perceptron classification method was proposed for the diagnosis of AD, healthy persons, and MCI patients based on the texture of the hippocampus (HC). This research obtained 72.5%, 85%, and 75% for each group; AD vs. MCI, AD vs. normal controls, and MCI vs. normal controls, respectively.

The study performed by [11], proposed a transfer learning framework based on the convolutional neural network (CNN) architecture for classifying Alzheimer's images into four classes: healthy person, EMCI, LMCI, and AD. They used layer-wise transfer learning as well as tissue segmentation of brain images for their goal and achieved 98.73% for AD vs. healthy person, 83.72% for EMCI vs. LMCI, and more than 80% for other groups.

Early diagnosis of Alzheimer's and different stages of cognitive impairment are important to prevent their progression. In the previous papers, the diagnosis of various stages of AD and MCI has been less discussed. This study has tried to diagnose different stages of MCI with significant accuracy and distinguish it from Alzheimer's and healthy people with MRI image processing methods. We used a unique combination of different brain area sizes and hippocampus grayscale statistical and texture features derived from 2D-MRI images to diagnose AD, CN, EMCI, and LMCI. In this study, the extraction of the specific brain features, such as the hippocampus and lateral ventricle size, based on established clinical knowledge has been utilized to provide detailed insights about the structural and textural changes associated with AD. This approach enhances early diagnostic accuracy and aids in the development of diagnostic biomarkers for different stages of MCI and AD. By focusing on transparency and specific feature analysis, this can be a valuable complement to existing DL research.

The proposed method provides a more comprehensive understanding of how the disease progresses compared to studies focusing on a single type of feature. Also, our analysis of each extracted feature's strength and direction of influence on the classification of each group, supported by correlation and p-value calculations, provides new insight into the relationship between brain structure changes and AD. Furthermore, by demonstrating the effectiveness of 2D-MRI

imaging data, our study extends the application of advanced diagnostic techniques using ML algorithms in clinical environments where 3D-MRI imaging may not be feasible, and it facilitates the diagnosis and intervention in a shorter time with significant accuracy. The use of ML algorithms instead of DL algorithms is recommended in the proposed method because it provides more interpretability and allows us to understand which brain features in MRI images have the most impact on the classification of AD and how these features change with the disease progression. Our dataset's relatively small size makes ML more suitable due to its lower risk of overfitting compared to DL models. In addition, ML models are more computationally intensive, which can be practical for various clinical applications.

2 Methods

The goal of this study is to present the method for distinguishing the six groups: CN vs. AD, EMCI vs. LMCI, CN vs. EMCI, CN vs. LMCI, AD vs. EMCI, and AD vs. LMCI, using two-dimensional T1-weighted coronal MRI analysis of the brain based on combined MRI features and ML algorithms. The proposed method endeavored to provide the ML model by extracting efficient and optimized features based on brain MRI images to enhance diagnostic accuracy which can potentially be used clinically. Furthermore, we assessed the correlation and dependence metrics to examine the strength and direction of association between each extracted feature and the classification of each group, these techniques are well-established and provide valid results. One of our main goals was to confirm the clinical relevance of the extracted features based on the analysis of MRI images and available neurobiological knowledge, and also we have tried to show how the extracted features change with the AD and MCI progression.

In this study, each label; 0, 1, 2 and 3 were used to represent CN, AD, EMCI, and LMCI groups, respectively. The feature and target matrices throughout the machine learning procedure have also been completed with the information of each group in the same order. The dataset was randomly split into 70% for training and 30% for testing. This process was repeated 20 times to ensure the performance of the model and the obtained results. Each trial used a different random seed to generate unique training and testing splits. The classifier's performance was reported as both the average result and the best result across these 20 trials.

This section describes the proposed MRI image processing based on ML algorithms. For the implementation of the proposed model to classify the patients and healthy persons, ImageJ and MATLAB (R2023a) programming for preprocessing, segmentation, and feature extraction steps and Spyder (Python−3.10 programming) for the classification process has been used. In Fig 1, the proposed ML algorithms for each step in MRI image analysis have been shown.

2.1 Database

In this paper, the Alzheimer's Disease Neuroimaging Initiative (ADNI) database has been utilized. The data resources from the North American ADNI include medical images of MRI and PET, clinical information, cognitive test scores, and biomarkers related to people with Alzheimer's, cognitive impairment in different stages, and healthy persons (https://adni.loni.usc.edu). This database has collected the validated data and provided a platform for researchers to work on Alzheimer's disease. The dataset used in this study includes two-dimensional T1-weighted-MRI images in the coronal section from all individuals of 4 categories: cognitively normal (CN), EMCI, LMCI, and AD obtained from ADNI1, ADNI2, ADNI3, and ADNI-GO data. These images consist of 100 subjects in each category with gender equality and age between 18 and 96. The MRI protocol and clinical information of subjects who participated in this study database are shown in Fig. 2 and Table 1, respectively.

2.2 Preprocessing and segmentation

After collecting the data, a series of preprocessing steps were applied to transform the raw MRI images into meaningful and usable information for subsequent analysis. In this paper, we applied a 3x3 median filter to remove the noise and the contrast enhancement technique to adjust the gray levels and improve the image's visibility to distinguish better the objects in the images. In this step, improving the quality of the images while preserving the edges and texture information has significant importance. Then, the Skull-stripping method is applied to remove the skull and non-brain tissues from the images. Skull stripping typically is based on thresholding techniques. The quality of applying the skull-stripping

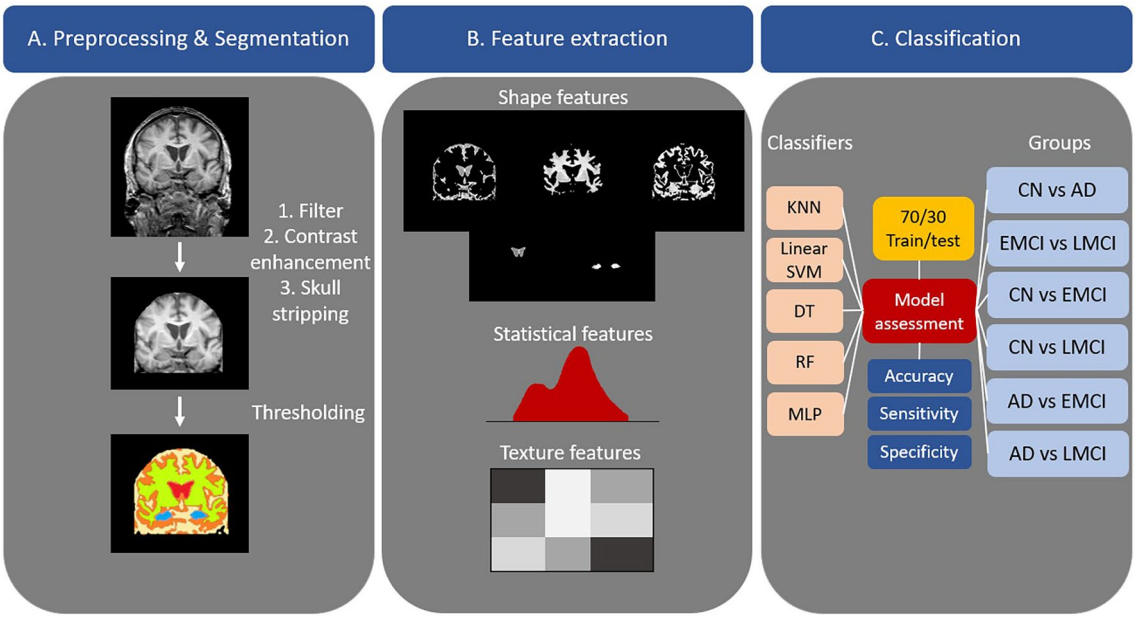

Fig. 1 The proposed image processing procedure based on the machine learning algorithms

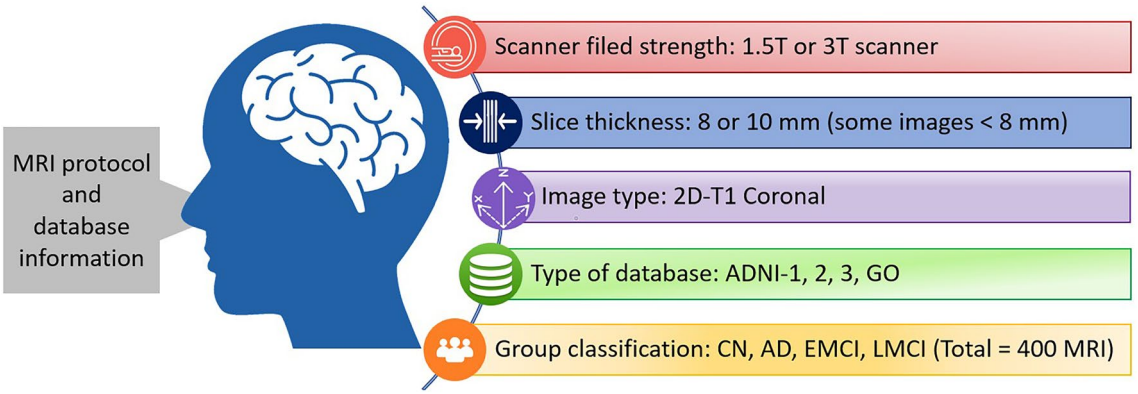

Fig. 2 The database imaging protocol

Table 1 The clinical information of subjects

S. No	Groups	Age (Mean±SD)	Weight (Mean±SD)	Gender	Quantity	Property
1	CN	77.01±6.65	74.47±16.39	F/M	50/50	Raw and Non-Filtered
2	AD	75.96±7.72	74.84±16.96	F/M	50/50	Raw and Non-Filtered
3	EMCI	73.55±7.44	79.72±17.12	F/M	50/50	Raw and Non-Filtered
4	LMCI	73.52±8.04	81.87±28.78	F/M	50/50	Raw and Non-Filtered

F Female, *M* Male

technique and removing the unwanted tissues can be affected by various factors, including the imaging protocol and MRI scanners, etc [12]. Furthermore, the variability of anatomy, age, and the extent of brain atrophy especially in different stages of MCI and AD, have an impact on skull stripping as well [13]. So, this step was performed with high accuracy, and each MRI image of the patient was evaluated.

After applying preprocessing techniques to the images, the regions of interest including; the hippocampus, lateral ventricles (LV), gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), have been segmented. The two highlighted regions in the proposed method; the hippocampus and the lateral ventricles, are presented in Fig. 3.

The hippocampus is a region of the brain that has a curved shape. Throughout the progression of MCI and AD, this region gets atrophy and its structure shape changes. This change can happen differently for each patient. So, the segmentation of this region faces challenges, especially in 2D MRI images. In this study, we used the ImageJ tool to draw polygons around the hippocampus, which can provide high accuracy in the segmentation of this region for all groups. Then, we used the multi-thresholding method based on the Otsu technique to segment the lateral ventricle, gray matter, white matter, and cerebrospinal fluid. Otsu's technique is a multi-thresholding approach that effectively can separate pixels in an image into distinct groups based on the histogram of the image [14, 15]. This technique is particularly effective in complex medical imaging applications [16, 17]. In this study, Otsu's thresholding was implemented by using MATLAB functions. This function calculates the global threshold level by maximizing the between-class variance $\sigma_B^2(T)$, which is defined as:

$$\sigma_B^2(T) = \omega_1(T)\omega_2(T)\left[\mu_1(T) - \mu_2(T)\right]^2 \quad (1)$$

where T is the selected threshold that divides the two classes, $\mu_1(T)$ and $\mu_2(T)$ are the mean intensities of each class, and $\omega_1(T)$ and $\omega_2(T)$ indicate the probabilities of each class being separated by threshold T. This function determines the threshold T that maximizes $\sigma_B^2(T)$ and yields optimal segmentation of the image containing the relevant regions.

2.3 Feature extraction

This study employed a comprehensive set of features related to brain structural/tissue changes, and clinical evaluations to maximize diagnostic accuracy for distinguishing different stages of MCI and AD by using MATLAB functions. After segmentation of regions of interest, morphological features including area sizes of the hippocampus, lateral ventricles, gray matter, white matter, and cerebrospinal fluid were extracted to assess brain structural changes. First-order statistical features were computed using means, standard deviations, skewness, and kurtosis of the intensity values of the hippocampus for evaluating tissue changes in this region. Additionally, texture features were derived using the gray-level co-occurrence matrix (GLCM) to quantify contrast, correlation, homogeneity, energy, and entropy of the hippocampus region by providing tissue heterogeneity. Invariant moments (IM with 7 components) were also computed to describe shape changes and pattern recognition in this region. Furthermore, a wavelet transform (WT-Haar wavelet with 65536 components) was extracted to obtain multi-resolution features from the entire brain. Clinical information such as age, weight, and mini-mental state examination (MMSE) scores of the patients were also used.

Subsequently, feature normalization was implemented to enhance the machine learning performance speed and its efficiency [18]. Feature normalization refers to rescaling the input features by the minimum and range to make all values lie between 0 and 1 [19].

High-feature matrix dimensional analysis and overfitting pose challenges for researchers and engineers in the fields of machine learning and data mining. Feature selection serves as an effective solution to address these issues by eliminating irrelevant and redundant data, thereby reducing computation time and enhancing the learning model procedure [20]. In this study, the obtained feature matrix dimensions have been large especially, due to the extraction of WT features.

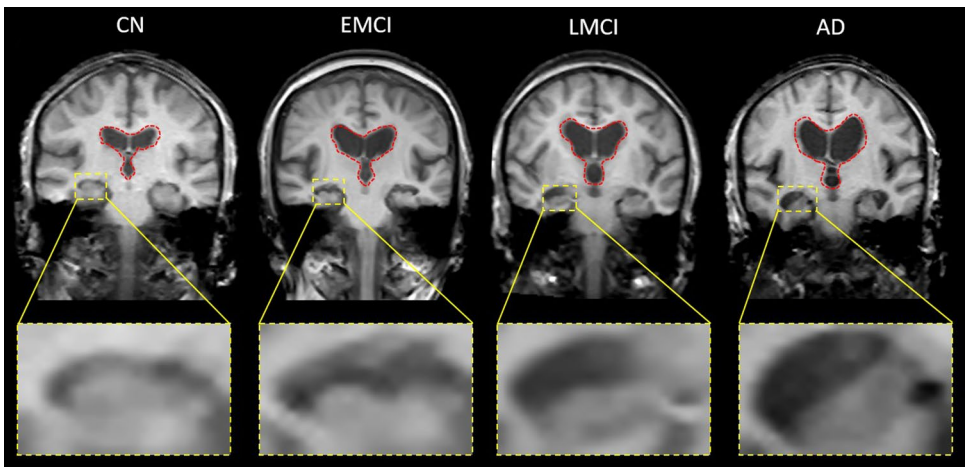

Fig. 3 The two-dimensional T1-weighted coronal MRI images for each category; CN, EMCI, LMCI, and AD (from left to right). The hippocampus regions and their magnifications are shown in the rectangles with the yellow dashed line, and the lateral ventricle regions are shown with the red dashed line

So, we examined two feature selection approaches; principal component analysis (PCA) as a first approach and making feature matrices based on different combinations of extracted features as a second one.

PCA is a statistical feature selection algorithm that transforms the large feature matrix into a smaller matrix by recognizing the efficient features [21]. Our feature matrix was reduced to 100 features by the implementation of PCA. We evaluated our classifiers using the first 20 features of this matrix. Additionally, as mentioned before, to find the best combination of 65560 features, we tried to combine the extracted features manually to find the best feature matrix for the classification of all groups.

2.4 Classification

Classification is the predicting or identifying process of new data relying on a training dataset consisting of extracted useful information [22]. In this study, we investigated several ML methods for classifying AD and different stages of MCI. Here, recent methods have been used for distinguishing 6 binary groups; CN vs. AD, EMCI vs. LMCI, CN vs. EMCI, CN vs. LMCI, AD vs. EMCI, and AD vs. LMCI. The KNN is an algorithm that uses the calculation of the distance between new data and all training data and makes a decision relying on the majority vote to categorize the new data [8]. Support vector machine (SVM) is a method that categorizes the new data by finding the maximum margin between classes and the optimal hyperplane [8, 23]. Decision tree (DT) is a recursively selecting model that divides the data based on the most significant feature [24]. Random forest (RF) is the combination of several decision tree models that can provide more accurate predictions [8, 25]. Lastly, a multi-layer perceptron (MLP) is a feed-forward neural network consisting of activation functions and neuron layers to explore the complex patterns in data [26, 27].

In the proposed method, KNN uses $k = 7$ for the number of neighbors (we assessed this classifier with different numbers of K from 3 to 10, that $K = 7$ having the best result) and *euclidean* metric for distance computation. SVM uses the *Linear* kernel, and $C = 1$for the regularization. DT uses the *Gini* criterion for measuring the quality of a split, the default *best* strategy for choosing the split at each node, and the unlimited depth and unlimited depth until all leaves are purged or until all leaves have less than the minimum number of samples required to split an internal node. RF consists of 100 trees with the same parameters which is defined for DT algorithm. MLP includes 100 hidden layers, *ReLU* function activation, and the *Adam* optimizer. These classifiers with described parameters are used to recognize 6 target binary groups.

2.5 Statistical analyses

To investigate the performance of the classifiers and assess the impact of each extracted feature on the diagnosis of six target groups, the statistical metrics such as; accuracy, sensitivity, and specificity have been calculated and their formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

where the patient is determined with a positive label and the healthy subject is determined with a negative label. Also, TP expresses true positive (persons that identified correctly as patients), FN indicates false negative (persons that recognized incorrectly as healthy), FP is used for false positive (persons that determined incorrectly as patients), and TN is used for true negative (persons that identified correctly as healthy). The accuracy measures the ability of a model to differentiate the patient and healthy cases correctly. The sensitivity measures the ability of a model to determine the patient cases correctly. The specificity measures the ability of a model to determine the healthy cases correctly [28].

To examine the strength and direction of association between each extracted feature and classification of each group, the Spearman and Pearson coefficients (correlation coefficients and P-value), and mutual information were calculated. This exploration aimed to identify which features contain crucial information for diagnosing different stages of MCI and AD (refer to Appendix B for more information).

3 Result

As mentioned before, we extracted 65,560 features in this study. As a first approach, we used PCA to reduce the feature dimension to 100 features. The first 20 features of this new 100 features were evaluated. As a second approach, to find the best combination of 65560 features, we tried to combine the extracted features manually to find the best feature matrix for the classification of all groups. We found out that the feature matrix combining all extracted features excluding WT and IM features has a significant performance (65560 (all) - 65536 (WT) - 7 (IM) = 17). The WT and IM did not have an acceptable result, so they were removed from the feature matrix. This 17-feature includes: HC area size, HC grayscale statistics (mean, standard deviation, skewness, kurtosis, contrast, correlation, energy, homogeneity, entropy), lateral ventricle (LV) area size, gray matter area size, white matter area size, cerebrospinal fluid area size, patient age, weight, and cognitive score. Then, the obtained 20 features from PCA were compared with the 17 features obtained from the manual combination. Among these two approaches, the 17 features had a significant performance. Consequently, this study focused on the 17 features mentioned for further analysis.

To identify which features among these 17 features contain more crucial information for diagnosing different stages of MCI and AD, and how these features change with AD progression, the strength and direction of association between each extracted feature and classification of each group were examined by calculating the Spearman and Pearson coefficients (correlation coefficients and *P*-value), and mutual information. The results of these metrics have been shown in Figs. S1–S5 in the supplementary file.

The five proposed classifiers were compared with each other based on 17 selected features. The obtained results for each binary classification; CN vs. AD, EMCI vs. LMCI, CN vs. MCI, CN vs. LMCI, AD vs. EMCI, and AD vs. LMCI, have been shown in Table 2. This table presents the average result of each performance metric in 20 trials and the best result that has been obtained during these trials.

According to the result table, the proposed method in this study obtained these best average accuracies of the classification in 20 trials; 95% (SVM), 71.50% (RF), 82.58% (RF), 84.91% (SVM), 85.83% (RF), and 85.08% (RF) for each group; CN vs AD, EMCI vs LMCI, CN vs EMCI, CN vs LMCI, AD vs EMCI, and AD vs LMCI, respectively. In addition, the best accuracies derived in 20 trials are; 100% (SVM, RF, and MLP), 83.33% (RF), 91.66% (RF), 95% (SVM, and MLP), 96.66% (RF), and 93.33% (DT), for each above-mentioned group, respectively.

4 Discussion

In this study, we proposed a medical image processing technique based on ML algorithms that describes what features and classifiers have good performance for distinguishing six groups of individuals: CN vs AD, EMCI vs LMCI, CN vs EMCI, CN vs LMCI, AD vs EMCI, and AD vs LMCI. This discrimination is achieved through the analysis of the two-dimensional T1-weighted coronal MRI. In addition, we conducted correlation, dependence, and mutual information analysis to examine the strength and direction of association between each extracted feature and the classification of each group. Corresponding to the statistical figures, both Spearman and Pearson correlations showed similar performance on average. Our analysis of extracted features revealed several key insights into their effectiveness. The results (coefficients and P-values) indicate that the MMSE score feature had the most correlation relationship with the classifications of CN vs AD, CN vs. LMCI, AD vs. EMCI, and AD vs. LMCI. Its high correlation across these multiple classifications emphasizes its role as an important indicator of cognitive impairment. Additionally, based on these coefficients and *P*-values, the hippocampus grayscale entropy feature for the EMCI vs. LMCI classification and the hippocampus grayscale correlation feature for the CN vs. EMCI classification had the most dependent relationship. These features reflect changes in hippocampal tissue properties relevant to disease progression. Furthermore, the MI figure shows that the MMSE score feature includes high mutual information for the classifications of CN vs AD, AD vs. EMCI, and AD vs. LMCI. Additionally, the WM area size feature for the EMCI vs. LMCI classification and the hippocampus grayscale entropy feature for the CN vs. EMCI and the CN vs LMCI classifications contain powerful information. It highlights structural changes in white matter associated with cognitive decline. Also, in general, the hippocampus and lateral ventricle size demonstrated significant potential for the classification of all groups. This result aligns with the established medical report and supports them as reliable biomarkers for AD and different stages of MCI. These findings highlight the importance of considering both correlation and mutual information analyses to identify features containing crucial information for diagnosing AD and MCI, thereby aiding in the development of impressive diagnostic biomarkers.

Table 2 The classifier's performance comparison for each group; CN vs AD, EMCI vs LMCI, CN vs EMCI, CN vs LMCI, AD vs EMCI, and AD vs LMCI

		CN vs AD		EMCI vs LMCI		CN vs EMCI		CN vs LMCI		AD vs EMCI		AD vs LMCI	
		AVG±SD	Best	AVG±SD	Best	AVG±SD	Best	AVG±SD	Best	AVG±SD	Best	AVG±SD	Best
KNN	ACC	91.25±0.04	98.33	61.66±0.05	75.00	71.66±0.06	86.66	80.16±0.05	90.00	79.25±0.04	88.33	77.66±0.04	83.33
	SEN	85.91±0.07	100	53.66±0.11	81.48	61.09±0.12	86.66	70.08±0.09	95.00	76.042±0.07	100	73.84±0.09	95.45
	SPE	96.17±0.03	100	70.81±0.08	88.88	84.08±0.06	100	90.63±0.06	100	83.02±0.05	100	82.64±0.06	92.00
SVM	ACC	95.00±0.02	100	61.25±0.05	71.66	77.66±0.05	90.00	84.91±0.03	95.00	82.16±0.03	91.66	70.75±0.04	85.00
	SEN	92.70±0.04	100	62.35±0.08	88.88	75.23±0.08	100	80.57±0.05	95.83	77.69±0.05	96.00	68.56±0.08	88.88
	SPE	97.47±0.02	100	61.26±0.12	92.30	80.64±0.07	100	89.44±0.08	100	87.00±0.04	100	73.53±0.07	96.29
DT	ACC	90.91±0.03	96.66	65.66±0.06	80.00	76.66±0.05	88.33	76.33±0.05	90.00	82.08±0.05	95.00	78.33±0.04	93.33
	SEN	91.42±0.05	100	66.58±0.10	85.71	75.99±0.06	96.87	77.41±0.09	96.55	82.80±0.09	100	75.91±0.08	100
	SPE	90.33±0.04	100	64.82±0.10	88.88	77.53±0.08	96.42	75.56±0.07	92.30	81.71±0.07	100	81.46±0.08	96.87
RF	ACC	93.91±0.02	100	71.50±0.05	83.33	82.58±0.04	91.66	83.91±0.04	93.33	85.83±0.04	96.66	85.08±0.04	91.66
	SEN	92.30±0.03	100	71.37±0.10	95.23	81.87±0.06	96.87	84.34±0.07	100	91.57±0.06	100	88.09±0.06	100
	SPE	95.37±0.03	100	72.83±0.12	96.15	83.72±0.06	100	83.86±0.06	100	80.03±0.06	100	82.18±0.07	100
MLP	ACC	92.41±0.02	100	63.08±0.04	80.00	75.50±0.06	90.00	82.75±0.05	95.00	80.75±0.05	91.66	80.16±0.03	88.33
	SEN	89.39±0.04	100	64.02±0.10	91.30	77.23±0.09	93.10	80.19±0.07	96.55	77.07±0.10	100	78.43±0.07	92.85
	SPE	95.55±0.03	100	63.38±0.08	86.36	74.47±0.09	96.42	85.85±0.07	100	84.72±0.05	100	82.21±0.07	96.29

ACC Accuracy, *SEN* Sensitivity, *SPE* Specificity, *AVG* Average

In evaluating the performance of different classifiers, our results present that various models show different strengths. The SVM classifier achieved the highest average accuracy for the classification of CN vs. AD (95%) and has had a remarkable performance for other group classifications. This shows that SVM is highly effective in detecting groups that have significant differences in their features. Also, the RF classifier showed strong performance in most classifications with average accuracy ranging from 71.50% to 96.66%. The compatibility and strength of RF make this classifier appropriate for classifying different feature sets and performing complex classification tasks. Moreover, the MLP and DT classifiers showed competitive performance, especially achieving 100% accuracy in several groups. Although the KNN classifier also did not have the highest performance compared to SVM or RF, it showed that it can show strong potential in situations where understanding local data patterns is essential for accurate classification. These results emphasize the effectiveness of SVM and RF in diagnosing AD and its subtypes. The diverse performance of different classifiers indicates that the different kinds of models will have different potentials in data classification and for achieving the best results, the appropriate model should be selected based on specific classification tasks and feature properties.

In our proposed method, we employed several strategies to overcome possible overfitting or reduce it in our model. First, to reduce the data dimension, we carefully selected the most relevant features for our models to avoid overcomplicating the models and fitting noise in the data. Also, we utilized a cross-validation method by randomly splitting the dataset into training and testing data and repeating this process 20 times with different random seeds. This approach allows us to validate the performance of the model in different subsets of the data and avoids overfitting to any particular set of data. Additionally, we used different classification methods such as ensemble methods like RF, which is made by combining several DT algorithms. It is notable that in the proposed method, ML techniques have been preferred over DL models because of their better performance on smaller data sets and reducing the risk of overfitting by having fewer parameters. These approaches collectively can help us to enhance the reliability of our framework in classifying different stages of MCI and AD.

5 Conclusion

This study aimed to provide a method to use a comprehensive set of features related to brain structural/tissue changes, and clinical evaluations to maximize diagnostic accuracy for distinguishing different stages of AD and MCI. In the proposed method, several strategies have been employed to overcome possible overfitting or reduce it in the model. Also,

this study has tried to investigate the effect of each extracted feature on the classification of the target groups and to evaluate the relationship between each extracted feature and the target groups in order to better recognize how the brain regions and tissue properties change with AD and MCI. These findings highlight the significance of using 2D-MRI image processing techniques and statistical analyses to identify key features that contain crucial information for diagnosing AD and different stages of MCI. This approach can enhance the accuracy of diagnosis and be used in clinical applications to help specialists. Early diagnosis of AD and differentiation of its symptoms from normal age-related cognitive decline and various stages of MCI have long been a challenging problem. Consequently, studies like this paper can be effective and significantly contribute to the development of important biomarkers related to AD and the prediction of that.

For future work, we plan to use a larger database and automated feature selection techniques in ML models. This approach can combine the strengths of neurobiological knowledge and extracted features to lead to model robustness and recognize new insights into AD progression. Also, we aim to increase the classification accuracy of each group by integrating different 2D views, including sagittal and axial. Additionally, we are interested in combining the two modalities; MRI and PET to have more information about the brain changes related to AD and enhance the accuracy of prediction and diagnostic procedure.

Acknowledgements For conducting this study, the database has been obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (https://adni.loni.usc.edu) [47]. The investigators involved in ADNI contributed to the design and implementation of ADNI and/or provided data but they did not take part in the analysis or the writing of this paper. A comprehensive list of ADNI investigators is available at the ADNI Acknowledgement List. In addition, this study was supported by the Swiss Government Excellent Scholarship through the Federal Commission for Scholarships for Foreign Students (FCS/ESKAS-Nr: 2023.0599). We would like to express our gratitude to this federation for the facilities and support it provided for this research. Also, we are profoundly grateful to the Sense, Innovation, and Research Center which has been founded by the University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), the University of Lausanne (UNIL), and the University Hospital of Lausanne (CHUV) for their generous assistance and encouragement.

Author contributions Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://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. All corresponding authors including Helia Givian and Jean-Paul Calbimonte contributed to each part of this study such as; analysis, conception, and design. Additionally, all corresponding authors participated in the writing and editing of the manuscript, and have reviewed and approved the final manuscript.

Funding This study was founded by the Swiss Government Excellence Scholarship via the Federal Commission for Scholarships for Foreign Students (FCS/ESKAS-Nr: 2023.0599). The Alzheimer's Disease Neuroimaging Initiative (ADNI) receives funding for its data collection and sharing from the National Institute on Aging (National Institutes of Health Grant U19 AG024904). Also, the grantee organization for this database is the Northern California Institute for Research and Education. Historically, ADNI has also been supported by the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private sector contributions via the Foundation for the National Institutes of Health (FNIH) 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.; EuroImmun; Eli Lilly and Company; F Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; IXICO Ltd.; Fujirebio; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Lumosity; NeuroRx Research; Johnson and Johnson Pharmaceutical Research and Development; Lundbeck; Merck and Co., Inc.; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Meso Scale Diagnostics, LLC.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. More details can be found at https://adni.loni.usc.edu/.

Data availability The datasets generated by the survey research during and/or analyzed during the current study are available in the Dataverse repository, (https://adni.loni.usc.edu). The primary goal of ADNI has been to examine the MRI, PET, biological markers, and clinical and neuropsychological assessment for evaluating the progression of mild cognitive impairment and early Alzheimer's disease. Michael W. Weiner, MD (email: Michael.Weiner@ucsf.edu) is the principal investigator for ADNI (This article does not include any studies involving human participants conducted by the authors).

Declarations

Ethics approval and consent to participate The entire database used in this study was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. According to ADNI protocols, all procedures conducted in studies involving human participants were based on Good Clinical Practice guidelines in accordance with the ethical standards of the institutional and/or national research committee and with the Helsinki Declaration US 21CFR Part 50 - Protection of Human Subjects, and Part 56 - Institutional Review Boards (IRBs) / Research 39 Ethics Boards (REBs), pursuant to state and federal HIPAA regulations, and its later amendments or comparable ethical standards. Written informed consent was obtained from all subjects and/or their authorized representatives and study partners before performing the protocol-specific procedures. More details are accessible at https://adni.loni.usc.edu/, ADNI-1-Protocol, ADNI-2-Protocol. The database procedures were approved by the institutional review boards of all participating centers ADNI Acknowledgement List. Ethics approval was secured from the institutional review boards of each involved institution including: the University of Southern California; Baylor College of Medicine; Johns Hopkins University; University of Michigan; Columbia University Medical Center; Oregon Health and Science University; Mayo Clinic, Rochester; Washington University, St. Louis; New York University; University of Alabama at Birmingham; University of California-San Diego; Mount Sinai School of Medicine; Rush University Medical Center; Wien Center; Duke University Medical Center; University of Kentucky;

University of Pittsburgh; University of Texas Southwestern Medical School; University of Rochester Medical Center; University of California, Irvine; University of Pennsylvania; Emory University; University of Kansas, Medical Center; University of California, Los Angeles; Mayo Clinic, Jacksonville; Indiana University; Yale University School of Medicine; McGill University, Montreal-Jewish General Hospital; Sunnybrook Health Sciences, Ontario; U.B.C. Clinic for AD and Related Disorders; Cognitive Neurology-St. Joseph's, Ontario; Cleveland Clinic Lou Ruvo Center for Brain Health; Northwestern University; Premiere Research Inst (Palm Beach Neurology); Georgetown University Medical Center; Brigham and Women's Hospital; Stanford University; Howard University; Ohio State University; Parkwood Hospital; Wake Forest University Health Sciences; Boston University; Banner Sun Health Research Institute; University of California, Davis-Sacramento; Neurological Care of CNY; University of Wisconsin; University of California, Irvine-BIC; Banner Alzheimer's Institute; Dent Neurologic Institute; Case Western Reserve University; Albany Medical College; Hartford Hospital, Olin Neuropsychiatry Research Center; Dartmouth-Hitchcock Medical Center; Rhode Island Hospital; Butler Hospital; UC San Francisco; Medical University South Carolina; University of Iowa College of Medicine; St. Joseph's Health Care Nathan Kline Institute; Cornell University and University of South Florida: USF Health Byrd Alzheimer's Institute. The investigators involved in ADNI contributed to the design and implementation of ADNI and/or provided data but they did not take part in the analysis or the writing of this paper. A comprehensive list of ADNI investigators is available at the ADNI Acknowledgement List (This article does not include any studies involving human participants conducted by the authors).

Human ethics and consent to participate The entire database used in this study was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. According to ADNI protocols, all procedures conducted in studies involving human participants were based on Good Clinical Practice guidelines in accordance with the ethical standards of the institutional and/or national research committee and with the Helsinki Declaration US 21CFR Part 50 - Protection of Human Subjects, and Part 56 - Institutional Review Boards (IRBs) / Research 39 Ethics Boards (REBs), pursuant to state and federal HIPAA regulations, and its later amendments or comparable ethical standards. Written informed consent was obtained from all subjects and/or their authorized representatives and study partners before performing the protocol-specific procedures. More details are accessible at https://adni.loni.usc.edu/, ADNI-1-Protocol, ADNI-2-Protocol. The database procedures were approved by the institutional review boards of all participating centers ADNI Acknowledgement List (This article does not include any studies involving human participants conducted by the authors).

Consent for publication Not applicable

Competing interests The authors have declared that no Conflict of interest exist

Appendix A

Research background

One of the main and important features in the development of AD is the build-up of beta-amyloid plaques and neurofibrillary tangles called tau protein in the brain [2, 29]. By increasing the number of these plaques and proteins, the message transmission and the communication between nerve cells is lost. As a result, it leads to a decrease in the function of nerve cells, and the cells gradually lose the ability to communicate with other cells. Finally, they are not able to do their tasks and die, which leads to the contraction or shrinking of tissues in the brain. The death of neurons, especially in the brain's hippocampus region, limits the ability to create new memories. The hippocampus region is responsible for memory formation which is the first region in the brain affected by this disease and shrinks as the disease progresses [29, 30]. Another area of the brain that is affected by the disease process is the lateral ventricle which contains cerebrospinal fluid and enlarges as the disease progresses [31]. As AD is an irreversible disease and there is no definitive treatment for this disease [32], early diagnosis of the disease is essential to prevent its progression. Doctors and specialists can prevent the progression of the disease or slow down its progress by accurately and correctly identifying brain changes. The current clinical diagnosis of AD relies on mental examinations such as the mini-mental state examination (MMSE). Brain imaging and biomedical research focus on finding associated biomarkers by using neuroimaging techniques [33]. These techniques are often the standard diagnosis routine process, used primarily to discard other conditions that may cause similar symptoms of mental decline such as stroke and head injury [34]. In the neuroimaging modalities, structural magnetic resonance (MRI) imaging has been extensively employed for different studies related to AD as an alternative to invasive analysis owing to its inter-tissue contrast, spatial resolution, wide accessibility, and relatively lower cost [35].

The primary advantage of MRI in AD research is providing detailed area structural changes and soft tissues [33]. Changes in the size of the lateral ventricle and hippocampus, as well as the changes in the grayscale of the hippocampus associated with different stages of Alzheimer's disease and mild cognitive impairment based on MRI, can be used as important biomarkers. These features contain critical information in the Alzheimer's diagnosis procedure and evaluation of the treatment process that can provide a better understanding of brain structural or tissue changes occurring in the brain due to Alzheimer's. Developing an automated MR image-based computer-assisted diagnosis (CAD) system for applications in the medical field helps to improve diagnosis, follow-up, and predictive ability [36]. Machine learning (ML) is an evolving branch of computational algorithms that are designed to emulate human intelligence by learning from the surrounding environment [37]. ML processing is a subset of artificial intelligence (AI) that starts with the observation of data to look for patterns and make effective decisions for the new data in the future, for instance through prediction, classification, or clustering tasks. The primary aim of these algorithms is to enable computers to *learn* without being directly programmed and without direct human intervention [38]. ML is a learning model that depends on many factors including the size, quality, and representativeness of the database, which may have a key impact on the machine learning results. A large database for research can be a limiting factor, as it is hard to access large volumes of sensitive information. Therefore, research especially in medical fields such as brain study is oftentimes based on small databases, in such cases, very complex diagnostic models tend to be overfitting. Also, as classification in ML is generally based on the training database, this overfitting model provides good results in the training data and poor results in the test data. Therefore, it can be argued that many researchers have overfitting problems [39]. The method of evaluation, type of classification method, and feature extraction are the factors that can help prevent the overfitting of the model. The application of ML in healthcare fields, especially in the early diagnosis of AD and different stages of MCI has recently improved and got attention due to the rapid progress in neuroimaging techniques that have generated large-scale multimodal neuroimaging data [40]. However, no technique has been identified as the definitive best, and most accurate method for diagnosing the various stages of MCI and AD. Consequently, distinguishing between different stages of MCI, particularly in early and late levels, alongside individuals with AD, poses a significant challenge that warrants further research endeavors.

Appendix B

Statistical analysis description

The Spearman and Pearson correlation coefficients are two different statistical parameters expressing the monotonic and linear relationship (or dependence) between variables, respectively [41, 42]. Their coefficients consist of a range from −1 to +1 with 0 value that each one indicates; perfect negative correlation, perfect positive correlation, and no correlation, respectively. A correlation around a value of 0.7 between the two variables describes the strong and positive relationship between variables [43]. The *P* value represents probability and calculates the likelihood that any recognized difference between groups is due to chance. If this probability is lower than the conventional 5% ($P< 0.05$) the correlation coefficient is called statistically significant. Another useful metric is mutual information. This metric as one of the most fundamental measurements has been calculated to quantify the dependence between extracted features and each group target [44, 45], which can measure the information gain between these variables [46]. This metric is a non-negative value that is equal to zero if only two variables are independent, and higher values mean higher dependency. The primary difference between correlation analysis and MI analysis is that correlation is a measure of monotonic and linear dependence between variables, whereas mutual information measures general dependence (including non-linear relations) which is a bit more complex. Therefore, MI can identify dependencies beyond those reliant solely on covariance.

References

1. Prince M, Wimo, A, Guerchet M, Ali G-C, Wu Y-T, Prina M. World Alzheimer report 2015. The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International (2015). https://www.alzint.org/resource/world-alzheimer-report-2015/
2. Raut A, Dalal V. A machine learning based approach for detection of Alzheimer's disease using analysis of hippocampus region from mri scan. In: 2017 International Conference on Computing Methodologies and Communication (ICCMC), 2017;236–242

3. Clarke A, Ashe C, Jenkinson J, Rowe O, Hyland P, Commins S. Predicting conversion of patients with mild cognitive impairment to Alzheimer's disease using bedside cognitive assessments. J Clin Exp Neuropsychol. 2022;44(10):703–12.
4. Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging. 2008;29(4):514–23.
5. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in mci patients, and their use in prediction of short-term conversion to ad: results from adni. Neuroimage. 2009;44(4):1415–22.
6. Raghavaiah P, Varadarajan S. A cad system design for Alzheimer's disease diagnosis using temporally consistent clustering and hybrid deep learning models. Biomed Signal Process Control. 2022;75: 103571.
7. Richhariya B, Tanveer M, Rashid AH, Initiative ADN. Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (usvm-rfe). Biomed Signal Process Control. 2020;59:101903.
8. Shukla A, Tiwari R, Tiwari S. Alzheimer's disease detection from fused pet and MRI modalities using an ensemble classifier. Mach Learn Knowl Extr. 2023;5(2):512–38.
9. Alickovic E, Subasi A, Initiative ADN. Automatic detection of Alzheimer disease based on histogram and random forest. In: CMBEBIH 2019: Proceedings of the International Conference on Medical and Biological Engineering, 16-18 May 2019, Banja Luka, Bosnia and Herzegovina, 2020;91–96. Springer, Cham
10. So J-H, Madusanka N, Choi H-K, Choi B-K, Park H-G. Deep learning for Alzheimer's disease classification using texture features. Curr Med Imaging. 2019;15(7):689–98.
11. Mehmood A, Yang S, Feng Z, Wang M, Ahmad AS, Khan R, Maqsood M, Yaqub M. A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. Neuroscience. 2021;460:43–52.
12. Pei L, Ak M, Tahon NHM, Zenkin S, Alkarawi S, Kamal A, Yilmaz M, Chen L, Er M, Ak N. A general skull stripping of multiparametric brain MRIs using 3d convolutional neural network. Sci Rep. 2022;12(1):10826.
13. Fennema-Notestine C, Ozyurt IB, Clark CP, Morris S, Bischoff-Grethe A, Bondi MW, Jernigan TL, Fischl B, Segonne F, Shattuck DW. Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: effects of diagnosis, bias correction, and slice location. Hum Brain Mapp. 2006;27(2):99–113.
14. Srikanth R, Bikshalu K. Multilevel thresholding image segmentation based on energy curve with harmony search algorithm. Ain Shams Eng J. 2021;12(1):1–20.
15. Cao Q, Qingge L, Yang P. Performance analysis of otsu-based thresholding algorithms: a comparative study. J Sens. 2021;2021:1–14.
16. Zortea M, Flores E, Scharcanski J. A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. Pattern Recognit. 2017;64:92–104.
17. He L, Ren X, Gao Q, Zhao X, Yao B, Chao Y. The connected-component labeling problem: a review of state-of-the-art algorithms. Pattern Recognit. 2017;70:25–43.
18. Jiang W. Applications of deep learning in stock market prediction: recent progress. Expert Syst Appl. 2021;184:115537.
19. Wang J, Wang J. Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. Neurocomputing. 2015;156:68–78.
20. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. Neurocomputing. 2018;300:70–9. https://doi.org/10.1016/j.neucom.2017.11.077.
21. Lv G, Guo S, Chen D, Feng H, Zhang K, Liu Y, Feng W. Laser ultrasonics and machine learning for automatic defect detection in metallic components. NDT E Int. 2023;133:102752.
22. Lamba D, Hsu WH, Alsadhan M. Predictive analytics and machine learning for medical informatics: a survey of tasks and techniques. In: Kumar, P., Kumar, Y., Tawhid, M.A. (eds.) Machine Learning, Big Data, and IoT for Medical Informatics, pp. 1–35. Academic Press, London, UK (2021). https://doi.org/10.1016/B978-0-12-821777-1.00023-9 . https://www.sciencedirect.com/science/article/pii/B9780128217771000239
23. Pink CM. Forensic ancestry assessment using cranial nonmetric traits traditionally applied to biological distance studies. In: Pilloud, M.A., Hefner, J.T. (eds.) Biological Distance Analysis, 2016;213–230. Academic Press, San Diego. https://doi.org/10.1016/B978-0-12-801966-5.00011-1 . https://www.sciencedirect.com/science/article/pii/B9780128019665000111
24. Tong JC, Ranganathan S. Computational t cell vaccine design. In: Tong, J.C., Ranganathan, S. (eds.) Computer-Aided Vaccine Design, pp. 59–86. Woodhead Publishing, Cambridge, UK (2013). https://doi.org/10.1533/9781908818416.59 . https://www.sciencedirect.com/science/article/pii/B9781907568411500052
25. Lipsky RH, Lin M. Genetic predictors of outcome following traumatic brain injury. In: Grafman, J., Salazar, A.M. (eds.) Traumatic Brain Injury, Part I. Handbook of Clinical Neurology, 2015;127: 23–41. Elsevier, Amsterdam. https://doi.org/10.1016/B978-0-444-52892-6.00003-9 . https://www.sciencedirect.com/science/article/pii/B9780444528926000039
26. Abirami S, Chitra P. Energy-efficient edge based real-time healthcare support system. In: Raj, P., Evangeline, P. (eds.) The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases. Advances in Computers, 2020;117:339–368. Elsevier, Amsterdam. https://doi.org/10.1016/bs.adcom.2019.09.007 . https://www.sciencedirect.com/science/article/pii/S0065245819300506
27. Meyer-Baese A, Schmid V. Foundations of neural networks. In: Meyer-Baese, A., Schmid, V. (eds.) Pattern Recognition and Signal Analysis in Medical Imaging (Second Edition), Second edition edn., 2014;197–243. Academic Press, Oxford. https://doi.org/10.1016/B978-0-12-409545-8.00007-8 . https://www.sciencedirect.com/science/article/pii/B9780124095458000078
28. Van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. Diagnostic methods i: sensitivity, specificity, and other measures of accuracy. Kidney Int. 2009;75(12):1257–63.
29. Glozman T. Hidden cues : Deep learning for alzheimer ' s disease classification cs 331 b project final report. (2016). https://api.semanticscholar.org/CorpusID:44060079
30. Setti SE, Hunsberger HC, Reed MN. Alterations in hippocampal activity and Alzheimer's disease. Transl Issues Psychol Sci. 2017;3(4):348.
31. Sadhukhan D, Veluppal A, Ramaniharan AK, Swaminathan R. Lateral ventricle texture analysis in Alzheimer brain mr images using kernel density estimation. Biomed Sci Instrum 2021;57(2)
32. Allgaier M, Allgaier C. An update on drug treatment options of Alzheimer's disease. Front Biosci-Landmark. 2014;19(8):1345–54.

33. Kim J, Jeong M, Stiles WR, Choi HS. Neuroimaging modalities in Alzheimer's disease: diagnosis and clinical features. Int J Mol Sci. 2022;23(11):6079.
34. Ranschaert ER, Morozov S, Algra PR. Artificial Intelligence in Medical Imaging: Opportunities. Applications and Risks. Cham: Springer; 2019.
35. Leandrou S, Petroudi S, Kyriacou PA, Reyes-Aldasoro CC, Pattichis CS. Quantitative MRI brain studies in mild cognitive impairment and Alzheimer's disease: a methodological review. IEEE Rev Biomed Eng. 2018;11:97–111.
36. Acharya UR, Fernandes SL, WeiKoh JE, Ciaccio EJ, Fabell MKM, Tanik UJ, Rajinikanth V, Yeong CH. Automated detection of Alzheimer's disease using brain MRI images-a study with various feature extraction techniques. J Med Syst. 2019;43:1–14.
37. El Naqa I, Murphy MJ. What Is Machine Learning? Cham: Springer; 2015.
38. Samuel AL. Some studies in machine learning using the game of checkers. IBM J Res Dev. 1959;3(3):210–29. https://doi.org/10.1147/rd.33.0210.
39. Sakai K, Yamada K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018. Jpn J Radiol. 2019;37(1):34–72. https://doi.org/10.1007/s11604-018-0794-4.
40. Jo T, Nho K, Saykin AJ. Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci. 2019;11:220.
41. Tjøstheim D, Otneim H, Støve B. Dependence. In: Tjøstheim, D., Otneim, H., Støve, B. (eds.) Statistical Modeling Using Local Gaussian Approximation, pp. 49–86. Academic Press, London, UK (2022). https://doi.org/10.1016/B978-0-12-815861-6.00010-9 . https://www.sciencedirect.com/science/article/pii/B9780128158616000109
42. Singh A, Porwal U, Bhardwaj A, Jin W. Multiscale representation learning for biomedical analysis. In: Govindaraju, V., Rao, A.S.R.S., Rao, C.R. (eds.) Deep Learning. Handbook of Statistics, 2023;48:9–27. Elsevier, Amsterdam, Netherlands. https://doi.org/10.1016/bs.host.2022.12.004 . https://www.sciencedirect.com/science/article/pii/S0169716122000591
43. David N. Selection of variables and factor derivation. Commercial data mining: Processing, analysis and modeling for predictive analytics projects, 2014;79–104.
44. Sinaga MA. On study of mutual information and its estimation methods. arXiv preprint arXiv:2106.14646 (2021)
45. ; Cheng P, Hao W, Dai S, Liu J, Gan Z, Carin L. Club: A contrastive log-ratio upper bound of mutual information. In: International Conference on Machine Learning, 2020;1779–1788. PMLR
46. Hoque N, Bhattacharyya DK, Kalita JK. Mifs-nd: a mutual information-based feature selection method. Expert Syst Appl. 2014;41(14):6371–85.
47. Ready RE, Ott BR, Grace J. Validity of informant reports about ad and mci patients' memory. Alzheimer Dis Assoc Disord. 2004;18(1):11–6.