SCIENTIFIC PAPER

Multi-modality MRI for Alzheimer's disease detection using deep learning

Latifa Houria[1] · Noureddine Belkhamsa[1] · Assia Cherfa[1] · Yazid Cherfa[1]

Abstract

Diffusion tensor imaging (DTI) is a new technology in magnetic resonance imaging, which allows us to observe the insightful structure of the human body in vivo and non-invasively. It identifies the microstructure of white matter (WM) connectivity by estimating the movement of water molecules at each voxel. This makes possible the identification of the damage to WM integrity caused by Alzheimer's disease (AD) at its early stage, called mild cognitive impairment (MCI). Furthermore, the brain's gray matter (GM) atrophy characterizes the main structural changes in AD, which can be sensitively detected by structural MRI (sMRI) modality. In this research, we aimed to classify the Alzheimer's diseases stages by developing a novel multi-modality MRI (DTI and sMRI) fusion strategy to detect WM alterations and GM atrophy in AD patients. The latter is based on a 2-dimensional deep convolutional neural network (CNN) features extractor and a support vector machine (SVM) classifier. The fusion framework consists of merging features extracted from DTI scalar metrics [(fractional anisotropy (FA) and mean diffusivity (MD)], and GM using 2D-CNN and feeding them to SVM to classify AD versus cognitively normal (CN), AD versus MCI, and MCI versus CN. Our novel multimodal AD method demonstrates a superior performance with an accuracy of 99.79%, 99.6%, and 97.00% for AD/CN, AD/MCI, and MCI/CN respectively.

Keywords Diffusion tensor imaging · Structural MRI · Alzheimer's disease · Convolutional neural network · Support vector machine

Introduction

Alzheimer's disease (AD) is an irreversible progressive neurodegenerative disorder that affects people over the age of 65 and outlines around 60% of dementia worldwide. It is caused by damage to nerve cells in certain brain regions, affecting a person's memory and cognitive abilities, which disrupt their daily life. The Alzheimer's Association declares that AD is the sixth leading cause of death in the USA; around 50 million people were diagnosed with this disease in 2018 and in 2050, this number will be tripled [1]. At present, no effective treatment or prevention is found. Moreover, disease management is prohibitively costly. Early screening of this disease is of primordial importance for researchers to slow down its progression and optimize the treatment. In this context, advances in neuroimaging, primarily magnetic resonance imaging (MRI), have shown the potential to improve the early diagnosis of AD.

AD is characterized by a progressive loss of Gray matter (GM) that occurs pre-symptomatically in certain neuroanatomical structures [2]. Structural MRI (sMRI) is the most used neuroimaging modality to detect brain atrophy. It has already highlighted many biomarkers of Alzheimer's disease; in particular, the atrophy of structures such as the hippocampus, the amygdala, and the thalamus [3]. In fact, hippocampal atrophy in prodromal patients proved to be the best structural predictor of Alzheimer's disease progression [4]. However, it is associated with a large number of neurodegenerative pathologies, thereby limiting its specificity to Alzheimer's disease [5].

✉ Latifa Houria
    houria_latifa@univ-blida.dz

    Assia Cherfa
    bouzouadcherfa_assia@univ-blida.dz

    Yazid Cherfa
    cherfa_yazid@univ-blida.dz

[1] LASICOM Laboratory, Department of Electronics, Faculty of Technology, University of Blida 1, Blida, Algeria

Within this frame of reference, many studies on the AD-prodromal phase called mild cognitive impairment (MCI) have focused their research on the hippocampus. Nevertheless, some other structures appear interesting such as, the volume of the amygdala which could be a structural predictor as powerful or even more efficient than the volume of the hippocampus to predict MCI [6, 7]. Furthermore, there are changes in white matter that preceded gray matter atrophy but are not detectable by sMRI [8]. The introduction of diffusion tensor imaging allows identification of these changes when the patient still presents an MCI [9]. The MCI is the transitory phase between (CN) decline and AD or another dementia.

DTI has conventionally studied the white matter microstructural integrity based on the estimation of the water molecules' diffusion in all directions (six directions at least) [10]. The degree of anisotropy of water diffusion is represented by the fractional anisotropy (FA), while mean diffusivity (MD) represents its magnitude. Studies have shown the importance of measuring these two DTI indices (FA and MD) to describe the physiological aging in the MCI patient phase [11]. Increased MD and decreased FA were reported in AD patients compared to CN. Higher MD in MCI patients was observed in both hippocampi [12]. Indeed, a considerable increase in MD and decrease in FA indicates a progressive loss of the barriers restricting the motion of water molecules in tissue compartments, associated with neuronal loss in AD [13]. It, therefore, seems important to measure the DTI indices because they can provide additional information on the pathophysiology of the disease.

The introduction of machine learning and deep learning techniques has greatly contributed to the diagnosis and prognosis of AD based on neuroimaging data [14]. Numerous research works have been published for the AD classification using DTI, where the FA and MD were the most frequently metrics used as features. The most popular among these machine learning-based methods utilized as classifiers, are the support vector machine (SVM), and random forest (RF) [15–19]. Most of them used the tract-based spatial statistics (TBSS) algorithm [20] to extract the white matter skeleton from FA and MD. They selected only the pertinence WM Skeleton information to perform binary or multi-classification using Alzheimer's disease national initiative (ADNI) data set. The difference was presented in the classification task, where Maggipinto et al. [18] used random forest and Lella et al. [19] proposed to concatenate the best result from different classifiers [(SVM, RF, and Multi-layer perceptron (MLP)] from all features groups [(FA, MD, radial diffusivity (RD), longitudinal diffusivity (LD)]. The use of DTI-based machine learning shows impressive performance. However, it is necessary to extract features and subsequently select the relevant ones to perform classification tasks, which is difficult and time-consuming.

Deep learning is a state-of-the-art machine learning method [21]. Classification techniques using deep convolutional neural networks (CNN) revealed higher AD detection performance [22]. Most of the literature approaches have used CNN-based sMRI to classify the different stages of Alzheimer's disease. CNN can handle low to high automatic feature extraction from complex structures. Some authors used a 2D deep CNN architecture trained on sMRI images of the full-brain from the ADNI database to perform a multi-classification task, which resulted in an accuracy of 99.9%. Wang et al. [23] suggested an eight-layer 2DCNN and achieved an accuracy of 97.65% to classify AD versus CN using the ADNI sMRI dataset.

Others have reported excellent results using transfer learning methods [24–26]. The transfer learning (TL) technique has been widely used as a powerful tool even with a smaller dataset containing a few hundred images. It involves reusing the weights of a pretrained model for other applications. Authors in [27] presented an approach for Alzheimer's Disease multi-classification. They fine-tuned the AlexNet using the preprocessed segmented and non-segmented images taken from the OASIS repository. Sava et al. [28] used sMRI images to train 29 different pre-trained models. The higher accuracy is provided using the EfficientNetB0 model to categorize AD versus CN versus MCI. In [29], the authors reached an accuracy of 98.73 to classify AD versus CN. They segmented the GM tissue using the tissue segmentation method, which was subsequently classified using the VGG family.

However, others have suggested extracting deep discriminative features based on transfer learning methods and classifying them with SVM [30, 31]. Jiang et al. [31] applied VGG16 to extract features from T1w images and classfiying them via SVM to identify EMCI from CN. Naz et al. [25] used deep features extracted from fully connected layers of pre-trained CNN. The study is realized using eleven pre-trained CNNs and the classification task is done using SVM where the VGG family presented the best performance. In [32], sMRI features were retrieved from Resnet101, InceptionV3, and Darknet53 models, then concatenated and optimized using the mRMR approach. KNN and SVM were utilized to classify these features. An accuracy of 99.1% was obtained.

In recent years, DTI indices, principally MD, combined with sMRI information have been adopted by many researchers. They proposed different techniques to combine DTI and sMRI. Massalimova et al. [33] have tried multi-modal Resnet-18 network (sMRI and DTI) in classifying CN, MCI, and AD from OASIS-3 datasets. They managed to suggest that the classification performed by the softmax layer could be preferable than another classifier in contrast to Kang et al. [30]. Kang et al. [30] suggested a fusion technique consisting of merging slices with the same index of the

T1w, FA, and MD images into an RGB slice. After that, the pre-trained VGG16 network is used to extract the features and SVM classifier to discriminate MCI patients, from CN using the ADNI dataset.

Aderghal et al. [34] proposed LeNet-like CNN based on sMRI and DTI-MD images. They selected the median slice Hippocampal and its two neighbors in each projection (axial, sagittal, and coronal). The proposed CNN is trained on the MNIST database. They first retrained the model on sMRI then on DTI-MD. They achieved a classification accuracy of 86.83% for AD versus CN, 69.85% for MCI versus CN, and 71.75% for AD versus MCI. Marzban et al. [35] proposed a simple 2DCNN based on a single convolution layer. They trained the model on diffusion scalars metrics (FA, MO, and MD) and GM. The cascaded MD and GM volumes achieved an overall accuracy of 88.9% and 79.6% respectively for AD versus CN and MCI versus CN. Ahmed et al. [36] extracted visual features from the hippocampus ROI in both sMRI and MD images. The extracted features and the amount of CSF calculated on the sMRI are combined and classified using multi-kernel learning (MKL).

Perez-Gonzalez et al. [37] computed the relevant features from sMRI and DTI and combined them with neuropsychological scores before feeding them to RF classifier. They achieved an accuracy of 88.8% using hippocampus, thalamus, and amygdala features for the right and left hemispheres to distinguish between MCI and CN.

Fang et al. [38] proposed re-transfer learning approach based AlexNet model to classify AD versus EMCI, AD versus LMCI, AD versus CN and EMCI versus CN. The first transfer learning process train sMRI with the pretrained AlexNet model and the model weight obtained is then used to train DTI images (FA and MD). Agostinho et al. [39] compared the performance of combining sMRI+DTI and sMRI+positron emission tomography (PET). They used various anatomical labeling atlases to divide the brain into ROIs and then extracted different metrics. After that, an embedded-based technique was used to minimize the features and SVM to classify them into CN and AD. The finding results achieved an accuracy of 97.00%, and 98.00% using sMRI+DTI and sMRI+ PET respectively.

Assessment of pathophysiological changes by neuroimaging would be essential to predict AD. Single modality cannot provide enough information, therefore, multi-modality must be combined to detect AD. sMRI and DTI have received more attention in recent years to study the progression of Alzheimer's disease. These two modalities are complementary; the sMRI detect the shrinkage of gray matter and changes in the brain volume. Moreover, the DTI is a useful prediction marker to detect the WM deterioration. In this context, we aim to detect patterns of micro and macrostructural changes in the different AD stages using the multi-modality MRI (sMRI and DTI) fusion process. We propose a new methodology that consists of a new CNN to extract the salient visual features from the DTI measurements and the GM images separately. After that, these features are concatenated and transmitted to SVM to identify AD from MCI, AD from CN, and MCI from CN.

Methodology

Our proposed strategy consists of pre-processing, a 2D slice selection, a features extraction, and a classification. We work on DTI measurements (FA, MD) and GM brain segmented from T1-weighted sMRI to classify (CN vs. AD), (AD vs. MCI), and (CN vs. MCI). New 2DCNN architecture was trained by slice-level dataset (only the 32 relevant slices selected from FA, MD, GM images) to extract the salient features from DTI maps and GM. The optimal FA-CNN, MD-CNN, and GM-CNN models are saved depending on lower loss value during the training process, then adapted to extract features from the last fully connected layer. After that, the features of each slice in the subject-level dataset (FA, MD, GM) are extracted by their optimal model (FA-CNN, MD-CNN, and GM-CNN). These features are merging and feeding to the SVM classifier to improve the performance as is illustrated in Fig. 1. The detailed description is found in the following subsections.

Database

Dataset used in this work has been obtained from the Alzheimer's disease neuroimaging initiative (ADNI) (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The objectives of the ADNI study are the identification of biomarkers for clinical use and early detection of AD [40]. The selected balanced dataset includes both Diffusion-weighted images (DWI) and sMRI brain scans from 150 individuals of both genders (50 AD, 50 CN, and 50 MCI), with ages varying from 55 to 90 acquired by GE medical system scanners. The 50 MCI subjects are selected with 25 early MCI and 25 Late MCI. The selected subjects coming from ADNI-GO and ADNI-2 phases.

The demographic information of the participants employed in this study is listed in Table 1, where the mini-mental state examination (MMSE) is a mental test that measures cognitive function. The lower MMSE score, as an auxiliary diagnostic index, implies poor cognitive ability.

The raw 3D T1-weighted sMRI were acquired with $256 \times 256 \times 196$ voxels per volume, a voxel size of $1.0 \times 1.0 \times 1.2$ mm3, inversion time = 400 ms, and flip angle = 11. The raw DWI data was acquired with 128×128 matrix, a voxel size of $2.7 \times 2.7 \times 2.7$ mm^3, and 41 gradient directions (b = 1000 s / mm^2). In addition to these images, 5 T2-weighted images
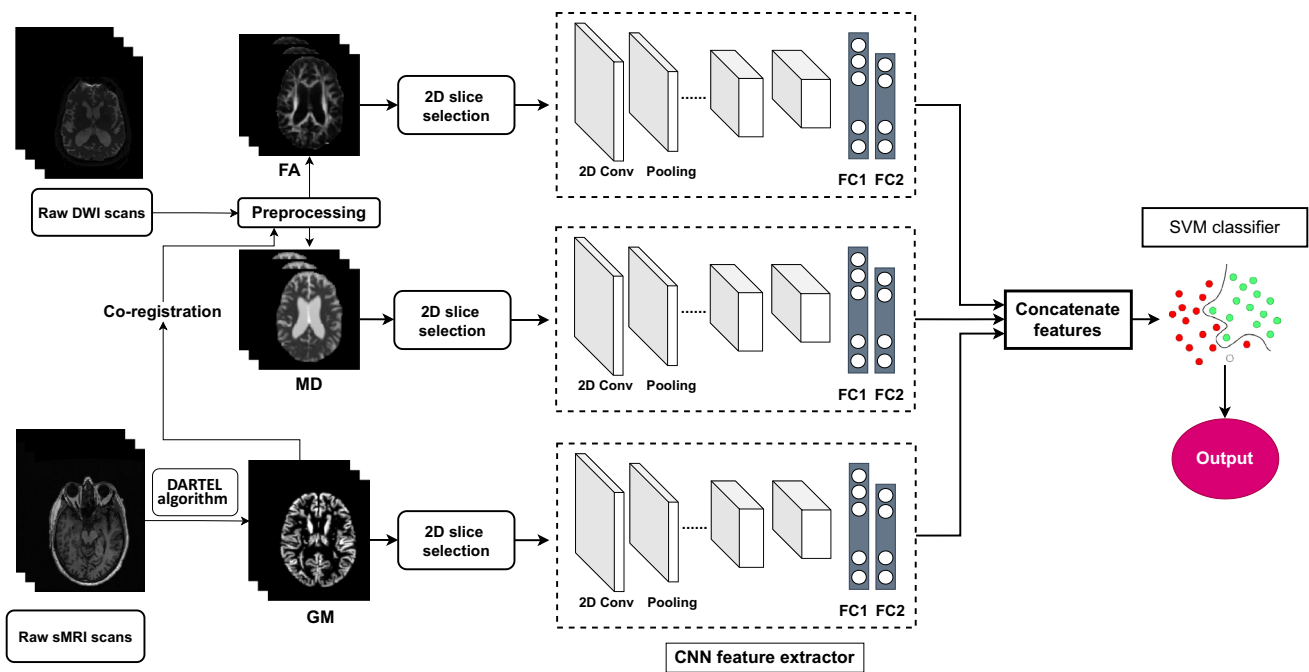
Fig. 1 Flowchart of the proposed fusion multi-modalities system using the 2DCNN-SVM approach for AD identification

Table 1 Subject-related statistics data

Classes	CN	MCI	AD
Number	50	50	50
Gender (F/M)	27/23	20/30	28/22
Age	72.5± 6.1	74.40± 7.47	75.60 ± 8.63
MMSE	28.93 ± 1.18	27.4 ± 1.99	23.0 ± 2.42

without diffusion (b = 0) are used as reference scans. More informtion about the acquisition parameters can be found in the ADNI2 protocol.

Pre-processing

All downloaded data are converted from DICOM format to Neuroimaging InFormatics Technology Initiative (NIFTI) format, using MRICro software (http://www.mricro.com/).

The pre-processing steps of the raw sMRI volumes to segment the GM are performed by the CAT12 toolbox (http://www.neuro.uni-jena.de/cat/). The CAT12 toolbox is an extension of SPM12 software [41]. In short, all T1-weighted 3D sMRI are normalized by the diffeomorphic anatomic registration through exponentiated lie algebra (DARTEL algorithm) using an affine transformation followed by a nonlinear registration, corrected for bias field inhomogeneities, and then segmented into GM, WM components.

DWI volumes are preprocessed using functional magnetic resonance imaging of the brain (FMRIB) software library (FSL) [42]. First DWI scans are corrected for eddy current distortions and susceptibility artefacts by the FSL-eddy correct. FSL's brain extraction tool was used to remove the brain skull. The diffusion tensor calculations are performed by the FSL dtifit at each voxel of fixed DWI scans. The eigenvalues of the diffusion tensor ($\lambda 1$, $\lambda 2$, $\lambda 3$) were utilized to obtain maps of scalar anisotropy and diffusivity. Several diffusion metrics can be calculated. The widely used diffusion metrics are fractional anisotropy (FA) and mean diffusivity (MD). FA is calculated using Eq. (1). MD represents the magnitude of diffusion which is calculated by averaging the three eigenvalues as it is mentioned in Eq. (2). Finally, FA and MD are co-registered with the corresponding sMRI scans and each scan contains 121×145×121 voxels using SPM12.

$$FA = \sqrt{\frac{1}{2}} \frac{\sqrt{(\lambda_1^2 - \lambda_2^2) + (\lambda_2^2 - \lambda_3^2)(\lambda_3^2 - \lambda_1^2)}}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \tag{1}$$

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2}$$

2D slice selection

Each FA, MD, and GM volume is decomposed into 2D slices along the axial view to highlight the most distinctive features and ensure improved classification efficiency. We select 32 slices from each subject based on higher entropy

information (slices with indices 34 65). The selected slices are associated with most of the deteriorated AD brain regions mentioned in literature such as the hippocampus, the entorhinal cortex, and thalamus. As a result, a total of 1600 (32×50) of each class (CN, MCI, and AD) are selected.

Feature extraction using 2DCNN

The handcrafted features extraction was the main problem in the traditional machine learning algorithms which is hard and time-consuming. CNN can perform this task automatically without human intervention.

CNN is the most common deep learning model used among neural networks. It is inspired by the human visual system. A typical CNN architecture comprises principally an input layer, convolution layer, pooling layer, fully connected layer, and classification layer. The convolution layer extracts automatically the features from the input FA, MD, or GM images by multiplying element-wise with a filter. The pooling layer aimed to reduce the redundant information by acquiring the average of a region or the maximum. The fully connected layer is used to reduce and transform the feature maps to a column feature map. The classifiers are finally used for AD prediction.

In short, the 2DCNN architecture consists of three convolutional layers with 3×3 size filters. Each convolutional layer is followed by a RELU layer, batch normalization (BN) layers, and a max-pooling layer, then two fully connected layers, softmax layer, and output layer. The RELU layer sets the negative values to zero and BN accelerates the training process. More details are tabulated in Table 2.

Classification using support vector machine (SVM)

SVM is a widely applied supervised learning method that treats small high-dimensional data by finding a maximal margin hyperplane to separate classes and solve a binary classification problem [43]. SVM is considered better to use than the Softmax layer as is mentioned in previously published studies [44, 45].

The trained FA-CNN, MD-CNN, and GM-CNN are adopted to extract the features. These features are then transmitted to the SVM classifier instead of the Softmax layer for AD classification. These features extracted from FA, MD, and GM images is a matrix whose size is the number of slices multiplied by the number of features selected from each slice. For 32 slices of each subject, the feature representation has the dimension of 32×2. For all subjects (100), the output of each model is a matrix of 100 32×2. They are then concatenated into a total feature matrix with the dimension of 3200×2. SVM classifier is trained and tested using these deep extracted features as is shown in Fig. 2.

Multi-modality MRI fusion process

The automatic AD screening fusion algorithm developed using multi-modalities MRI is illustrated in Fig. 1. The three optimal CNN (FA-CNN, MD-CNN and, GM-CNN) are used to extract features. We tried several fusion procedures experiences (FA and MD), (FA and GM), (MD and GM), and (FA and MD and GM) to choose the best model score. The fusion process consists of merging the features extracted from FA, MD, and GM into a global feature vector. Accordingly, the size of the fused FA + MD + GM feature matrix is 3200×6.

Table 2 Layers proprieties for the proposed 2DCNN architecture

Layers (#)	Layer name	Layer properties
1	Input layer	145×121×1 images
2	Convolutional	32 3× 3×1 convolutions with stride [1 1] and padding same
3	Batch normalization	32
4	Rectified linear unit	Rectified linear unit
5	Max pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
6	Convolutional	16 3×3×32 convolutions with stride [1 1] and padding same
7	Batch normalization	16
8	Rectified linear unit	Rectified linear unit
9	Max pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
10	Convolutional	8 3×3×16 convolutions with stride [1 1] and padding same
11	Batch normalization	8
12	Rectified linear unit	Rectified linear unit
13	Fully connected	256 hidden neurons in fully connected layer
14	Fully connected	2 hidden neurons in fully connected layer
15	Softmax	Softmax
16	Classification output	2 output classes

Fig. 2 The pipeline of proposed GM-CNN with SVM method to distinguish between AD and CN
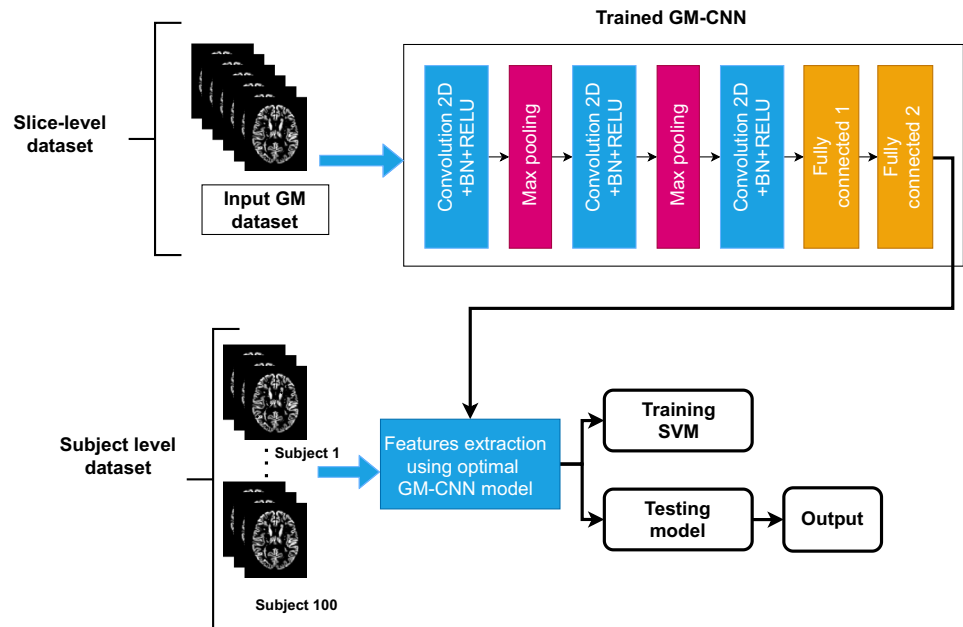
Results

Experiments

In this work, several experiments are carried out to validate the effectiveness of our proposed method to classify (AD vs. CN), (CN vs. MCI), and (AD vs. MCI). In the first experience, we performed a direct unimodal classification of features extracted from FA, MD, and GM. This gives us information about the best modality and map. In the second experiment, we study whether multi-modality increases performance and allows better discrimination between the different classes or not. This is achieved by studying the impact of merging features of the two modalities.

The 2DCNN-SVM proposed has been implemented using MATLAB ver. R2019a and running on a 3.1 GHz Intel-i7 processor, 16 GB of RAM. The CNN model was trained using an optimized stochastic gradient descent momentum (SGDM) using the back-propagation algorithm and cross-entropy as a loss function. The batch size is 64, the learning rate is 0.0001 for 25 epochs. There is a total of 3200 images of each map (FA, MD, and GM), 1600 images for each class. The dataset is divided into 75% for training, 15% for validation and the remaining 15% for testing the SVM. The same CNN architecture is used to train FA slices, MD slices and GM slices.

For the SVM classifier, the extracted data is categorized into training, validation, and test data. We used the extracted features from 2720 images for the training and 480 images for the test.

The best SVM using radial basis function (RBF) (Gaussian kernel) classification score was obtained by 10-fold cross-validation. The optimal hyperparameters (cost and gamma) were determined using the grid search technique. It finds the best model result from different combinations of parameters; where cost controls the error and gamma gives the curvature weight of the decision boundary.

Evaluation

The performance of our method was validated using accuracy and the area under the receiver operating characteristic curve (AUC). The validation results are illustrated in Table 3 and the ROC curves of 10-fold cross-validation are shown in Figs. 3, 4, 5.

The fused FA, MD, and GM improved better the result and outperformed the single modality and the sMRI+MD fused procedures adopted in many previous studies [30, 34, 35].

We tested our method using 240 AD images, 240 CN images, and 240 MCI images. The used evaluation metrics are the accuracy, sensitivity, and specificity determined by the confusion matrices. An example of the confusion matrix of the fused characteristics FA, MD, and GM is shown in Figs. 6, 7, 8. All test results are summarized in Table 4.

Table 4 shows that the FA, MD, and GM are important to discriminate the different AD stages. For the use of FA, MD, GM independently, we report that MD obtained the best result in the case of AD versus CN with an accuracy of 98.96%. However, the GM yields better results in classifying AD versus MCI and CN versus MCI with an accuracy of 96.88% and 93.50% respectively.

We investigated the best combination of features (FA and MD, FA and GM, and MD and GM). Fused FA and MD

Table 3 The performance of the validation dataset

Classes	Performances	FA	MD	FA+ MD	GM	GM+ FA	GM+ MD	GM+ FA + MD
AD/CN	Accuracy	95.48	97.94	97.79	96.91	99.12	98.49	**99.52**
	AUC	98.79	99.52	98.91	99.51	99.96	98.97	**99.99**
AD/MCI	Accuracy	92.98	96.21	96.95	96.87	98.49	98.68	**99.08**
	AUC	96.42	99.02	99.54	99.31	99.98	99.93	**99.96**
CN/MCI	Accuracy	93.31	92.76	94.85	93.03	97.17	95.99	**97.79**
	AUC	96.43	95.18	98.83	97.72	99.06	98.65	**99.60**

Bold values indicate the best results
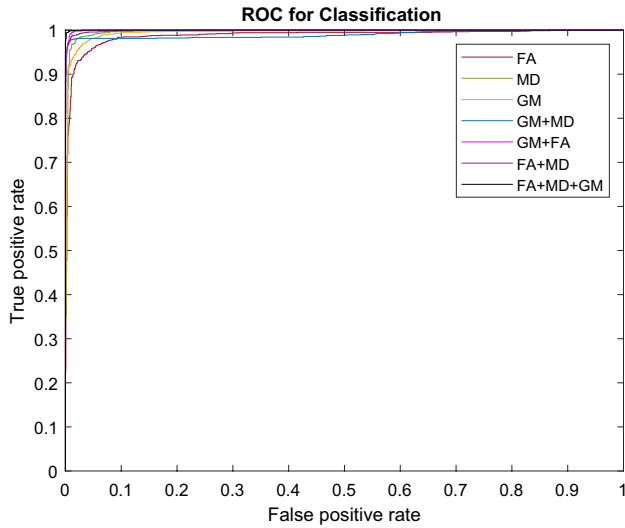
Fig. 3 ROC curves of AD versus CN
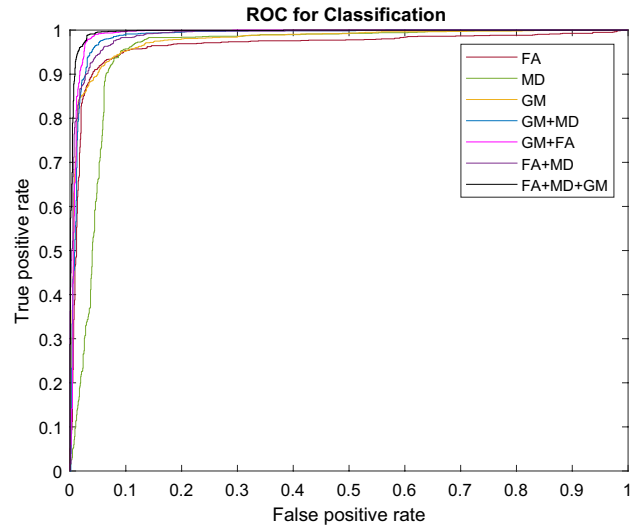
Fig. 5 ROC curves of CN versus MCI
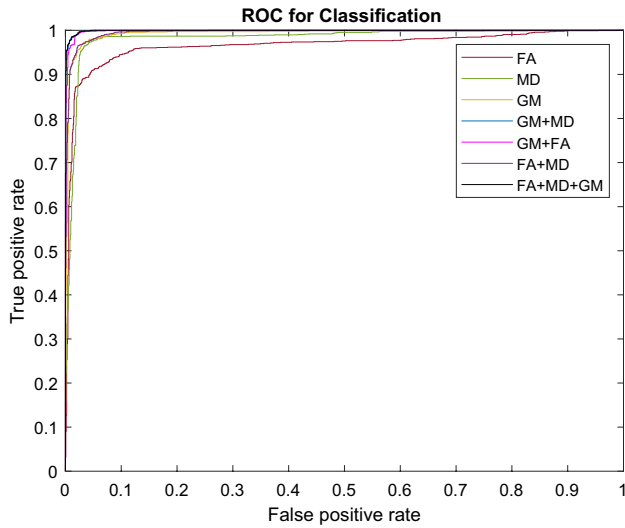
Fig. 4 ROC curves of AD versus MCI

Fig. 6 Matrice of confusion of AD versus CN

Confusion Matrix

Fig. 7 Matrice of confusion of AD versus MCI

Confusion Matrix

Fig. 8 Matrice of confusion of CN versus MCI

outperformed the other combined features with an accuracy of 99.98 % and 98.33% to classify AD versus CN and AD versus MCI. On the other hand, fused GM and MD achieved

higher results to classify CN versus MCI with an accuracy of 97.00%, a sensitivity of 97.20%, and a specificity of 96.80%. Compared to the use of the single modality, the merged FA, MD, and GM led to an increase of approximately 3.54%, 5.83%, 4.1% for the accuracy, 3.55%, 5,85%, 4.75% for the sensitivity, and 2.8%, 5.85%, 5.2% for the specificity in the three cases of classification (AD versus CN), (AD versus MCI), and (CN versus MCI) respectively. So, the multimodality gives the best performance as clearly shown in Fig. 9.

Discussion

To validate the performance and efficiency of our novel workflow, we compared it to the previous approaches presented in the literature and dealing with the same databases (ADNI) and the same modalities (sMRI and DTI). Our results gained higher accuracy in the AD detection compared to other studies as is shown in Table 5.

In general, our results concerning AD early detection imply the existence of distinct pathophysiological processes. In fact, the hippocampus is known to be one of the earliest and most severely damaged structures affected by AD. However, there are other structures involved in AD detection such as the amygdala, thalamus, and putamen. The relevant slices selection seems a powerful and easy method than segmenting the hippocampus or other brain's regions which requires a human expert. Our network learns the complex patterns of brain atrophy from relevant sections that contain almost all of the AD-affected regions mentioned in the literature, for each patient. This eliminates the process of segmentation of the hippocampus and other regions of the brain. Moreover, a subsequent selection of the most discriminating characteristics is avoided in our approach.

Our results confirm the effectiveness of the DTI measurement FA and MD in the classification of AD versus CN, AD versus MCI, and CN versus MCI which is consistent with the previous works [18, 19]. In addition, The GM

Table 4 Performance evaluation of the proposed 2DCNN-SVM technique on the test dataset

Classes	Performances	FA	MD	FA+MD	GM	GM+FA	GM+MD	GM+FA + MD
AD/CN	Accuracy	97.50	98.96	**99.89**	96.25	98.50	99.38	99.79
	Sensitivity	97.50	99.00	98.75	96.25	98.55	99.40	**99.80**
	Specificity	97.50	99.00	98.75	96.25	98.50	99.40	**99.80**
AD/MCI	Accuracy	93.75	95.42	98.33	96.88	98.96	98.13	**99.58**
	Sensitivity	94.00	95.45	98.35	96.90	99.00	98.20	**99.60**
	Specificity	93.75	95.45	98.35	96.85	98.95	98.15	**99.60**
CN/MCI	Accuracy	92.30	92.10	94.40	93.50	95.60	**97.00**	**97.00**
	Sensitivity	92.25	92.20	94.50	93.60	95.90	**97.20**	96.95
	Specificity	92.20	91.80	94.20	93.25	95.20	96.80	**97.00**

Bold values indicate the best results

Fig. 9 Comparison of performance of proposed technique for binary classification of AD versus CN, AD versus MCI, and CN versus MCI
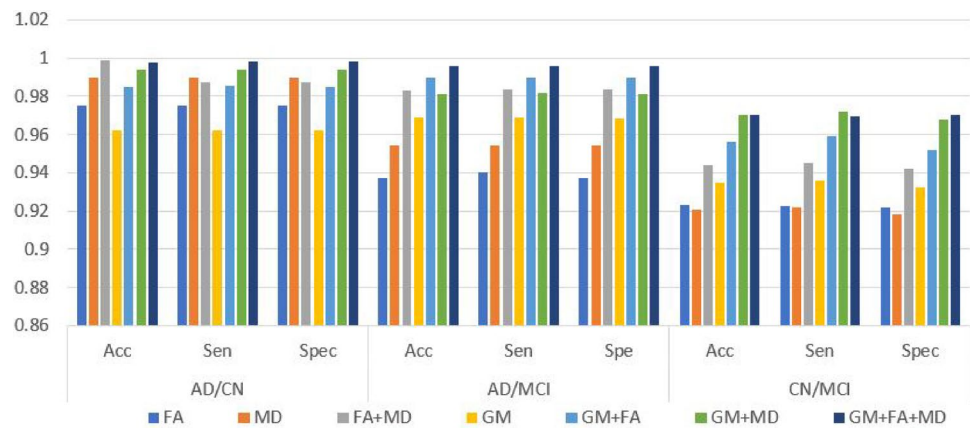
Table 5 Comparison of results with state-of-the-art techniques applied to AD detection

Study	Modality	Subject size			Method (Maps and classifier)	Classification accuracy (%)		
		AD	CN	MCI		AD /CN	CN / MCI	MCI/ AD
Lella et al. (2017) [17]	DTI	40	40	–	(FA, MD), RF,SVM and naive Bayes	78.00	–	–
Maggipinto et al. (2017) [18]	DTI	89	90	90	(FA MD), RF	87.00	81.00	–
Ahmed et al. (2017) [36]	DTI+sMRI	45	52	58	Hippocampus ROI (sMRI+MD)+CSF volume, SVM	90.20	79.42	76.63
Aderghal et al. (2020) [34]	DTI+sMRI	64	399	273	Hippocampus ROI (sMRI+MD), LeNet-CNN	92.30	78.48	79.16
Jiang et al. (2020) [31]	sMRI	–	50	70	VGG16 with SVM	–	89.40	–
Marzban et al. (2020) [35]	DTI+sMRI	115	185	106	(MD, FA, MO, GM), 2DCNN	93.50	–	79.6
Kang et al. (2020) [30]	DTI+sMRI	–	50	70	(FA+MD+sMRI), VGG16+SVM	–	94.20	–
Feng et al. (2020) [45]	sMRI	130	135	133	3D-CNN+SVM	99.10	98.90	89.40
Lella et al. (2021) [19]	DTI	40	40	–	(FA+MD+RD+LD),SVM, RF, PML	88.50	–	–
Naz et al. (2021) [25]	sMRI	95	146	138	VGG16, VGG19	98.89	97.06	**99.27**
Our approach	DTI+sMRI	50	50 50	50 50	(FA+MD,+GM), 2DCNN-SVM	**99.79**	**99.58**	97.00

Bold values indicate the best results

atrophy in sMRI is of great interest to researchers for the AD early detection. The sMRI based transfer learning has proven impressive results [24, 26]. Generally, the VGG16 and VGG19 models have gained higher accuracy than other pre-trained models [25]. Recently, some of the authors [30, 31] succeeded in using a pre-trained (VGG16) model for automatic extraction of features and SVM for the classification; they achieved a higher accuracy. However, the transfer learning technique relies generally on natural images whose models are trained using the Imagenet database [46]. Conversely, our simple networks learn and extract from scratch the most pertinent features.

In the past few years, the multi modalities (DTI-MD and sMRI) were reported by many researchers. They proposed different combination techniques to ensure the best classification. Aderghal et al. [34] suggested the transfer learning technique to perform the fusion process and Marzban et al. [35] adopted a cascaded CNN. However, they achieved lower accuracy than what we got which is over 97%. This is probably due to the small sample size we used compared to them, or the fact that we didn't work on specific ROI, or the impact of adding FA.

In summary, both, diffusion scales metrics and the GM are powerful elements and important for AD stage discrimination. The multi-modality fusion process (FA+MD+GM) seems to be the best technique to improve the AD classification performance.

Conclusions

This paper proposes a computer-assisted diagnosis for Alzheimer's disease classification using multimodality MRI imaging. A new CNN is used to extract features from DTI-FA, DTI-MD, and GM images. Deep features are concatenated and fed to the SVM classifier. The proposed

CNN-SVM approach demonstrates the effectiveness of multimodal MRI use by achieving a classification accuracy of 99.79%, 99.85%, and 97.00% for AD/CN, AD/MCI, and CN/MCI, respectively. The recent findings show that the proposed approach offers value to the healthcare profession by allowing for the precise classification of Alzheimer's disease.

In the future, a new fusion technique will be present. We will incorporate other neuroimaging techniques or DTI measurements like LD and RD. The three planes (axial, sagittal, and coronal) can also be merged instead of limiting the study to the axial plane.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Reserach involving human and animal participants This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Patterson C(2018) The state of the art of dementia research: New frontiers. World Alzheimer Report **2018**
2. Fox N, Warrington E, Freeborough P, Hartikainen P, Kennedy A, Stevens J, Rossor MN (1996) Presymptomatic hippocampal atrophy in Alzheimer's disease: a longitudinal MRI study. Brain 119(6):2001–2007
3. Pini L, Pievani M, Bocchetta M, Altomare D, Bosco P, Cavedo E, Galluzzi S, Marizzoni M, Frisoni GB (2016) Brain atrophy in Alzheimers disease and aging. Ageing Res Rev 30:25–48
4. Apostolova LG, Dinov ID, Dutton RA, Hayashi KM, Toga AW, Cummings JL, Thompson PM (2006) 3d comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease. Brain 129(11):2867–2873
5. Van de Pol L, Gertz H-J, Scheltens P, Wolf H (2011) Hippocampal atrophy in subcortical vascular dementia. Neurodegener Dis 8(6):465–469
6. Bell-McGinty S, Lopez OL, Meltzer CC, Scanlon JM, Whyte EM, Dekosky ST, Becker JT (2005) Differential cortical atrophy in subgroups of mild cognitive impairment. Arch Neurol 62(9):1393–1397
7. Klein-Koerkamp Y, A Heckemann R, T Ramdeen K, Moreaud O, Keignart S, Krainik A, Hammers A, Baciu M, Hot P (2014) Amygdalar atrophy in early Alzheimers disease. Curr Alzheimer Res 11(3):239–252
8. Ringman JM, O'Neill J, Geschwind D, Medina L, Apostolova LG, Rodriguez Y, Schaffer B, Varpetian A, Tseng B, Ortiz F et al (2007) Diffusion tensor imaging in preclinical and presymptomatic carriers of familial Alzheimer's disease mutations. Brain 130(7):1767–1776
9. Medina D, deToledo-Morrell L, Urresta F, Gabrieli JD, Moseley M, Fleischman D, Bennett DA, Leurgans S, Turner DA, Stebbins GT (2006) White matter changes in mild cognitive impairment and ad: a diffusion tensor imaging study. Neurobiol Aging 27(5):663–672
10. Basser PJ, Mattiello J, LeBihan D (1994) MR diffusion tensor spectroscopy and imaging. Biophys J 66(1):259–267
11. Cherubini A, Péran P, Caltagirone C, Sabatini U, Spalletta G (2009) Aging of subcortical nuclei: microstructural, mineralization and atrophy modifications measured in vivo using MRI. Neuroimage 48(1):29–36
12. Fellgiebel A, Wille P, Müller MJ, Winterer G, Scheurich A, Vucurevic G, Schmidt LG, Stoeter P (2004) Ultrastructural hippocampal and white matter alterations in mild cognitive impairment: a diffusion tensor imaging study. Dement Geriatr Cogn Disord 18(1):101–108
13. Kantarci K, Avula R, Senjem M, Samikoglu A, Zhang B, Weigand S, Przybelski S, Edmonson H, Vemuri P, Knopman DS et al (2010) Dementia with Lewy bodies and Alzheimer disease: neurodegenerative patterns characterized by DTI. Neurology 74(22):1814–1821
14. Billeci L, Badolato A, Bachi L, Tonacci A (2020) Machine learning for the classification of Alzheimers disease and its prodromal stage using brain diffusion tensor imaging data: a systematic review. Processes 8(9):1071
15. Dyrba M, Ewers M, Wegrzyn M, Kilimann I, Plant C, Oswald A, Meindl T, Pievani M, Bokde AL, Fellgiebel A et al (2013) Robust automated detection of microstructural white matter degeneration in Alzheimers disease using machine learning classification of multicenter DTI data. PLoS ONE 8(5):64925
16. O'Dwyer L, Lamberton F, Bokde AL, Ewers M, Faluyi YO, Tanner C, Mazoyer B, O'Neill D, Bartley M, Collins DR et al (2012) Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. PLoS ONE 7(2):32441
17. Lella E, Amoroso N, Bellotti R, Diacono D, La Rocca M, Maggipinto T, Monaco A, Tangaro S (2017) Machine learning for the assessment of Alzheimer's disease through DTI. Applications of Digital Image Processing XL, vol 10396. SPIE, Bellingham, pp 239–246
18. Maggipinto T, Bellotti R, Amoroso N, Diacono D, Donvito G, Lella E, Monaco A, Scelsi MA, Tangaro S, Initiative ADN et al (2017) DTI measurements for Alzheimers classification. Phys Med Biol 62(6):2361
19. Lella E, Pazienza A, Lofù D, Anglani R, Vitulano F (2021) An ensemble learning approach based on diffusion tensor imaging measures for Alzheimers disease classification. Electronics 10(3):249
20. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ,

Matthews PM et al (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31(4):1487–1505

21. Lee G, Fujita H (2020) Deep learning in medical image analysis: challenges and applications, vol 1213. Springer, Cham

22. Hazarika RA, Abraham A, Sur SN, Maji AK, Kandar D (2021) Different techniques for Alzheimers disease classification using brain images: a study. Int J Multimed Inf Retr 2021:1–20

23. Wang S-H, Phillips P, Sui Y, Liu B, Yang M, Cheng H (2018) Classification of Alzheimers disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. J Med Syst 42(5):1–11

24. Ashraf A, Naz S, Shirazi SH, Razzak I, Parsad M (2021) Deep transfer learning for Alzheimer neurological disorder detection. Multimed Tools Appl 80(20):30117–30142

25. Naz S, Ashraf A, Zaib A (2022) Transfer learning using freeze features for Alzheimer neurological disorder detection using Adni dataset. Multimed Syst 28(1):85–94

26. Maqsood M, Nazir F, Khan U, Aadil F, Jamal H, Mehmood I, Song O-y (2019) Transfer learning assisted classification and detection of Alzheimers disease stages using 3d MRI scans. Sensors 19(11):2645

27. Mehmood A, Maqsood M, Bashir M, Shuyuan Y (2020) A deep Siamese convolution neural network for multi-class classification of Alzheimer disease. Brain Sci 10(2):84

28. Savaş S (2022) Detecting the stages of Alzheimers disease with pre-trained deep learning architectures. Arab J Sci Eng 47(2):2201–2218

29. Mehmood A, Yang S, Feng Z, Wang M, Ahmad AS, Khan R, Maqsood M, Yaqub M (2021) A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. Neuroscience 460:43–52

30. Kang L, Jiang J, Huang J, Zhang T (2020) Identifying early mild cognitive impairment by multi-modality MRI-based deep learning. Front Aging Neurosci 12:206

31. Jiang J, Kang L, Huang J, Zhang T (2020) Deep learning based mild cognitive impairment diagnosis using structure MR images. Neurosci Lett 730:134971

32. Eroglu Y, Yildirim M, Cinar A (2022) MRMR-based hybrid convolutional neural network model for classification of Alzheimer's disease on brain magnetic resonance images. Int J Imaging Syst Technol 32(2):517–527

33. Massalimova A, Varol H.A (2021) Input agnostic deep learning for alzheimers disease classification using multimodal MRI images. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 2875–2878. IEEE

34. Aderghal K, Afdel K, Benois-Pineau J, Catheline G, Initiative ADN et al (2020) Improving Alzheimer's stage categorization with convolutional neural network using transfer learning and different magnetic resonance imaging modalities. Heliyon 6(12):05652

35. Marzban EN, Eldeib AM, Yassine IA, Kadah YM, Initiative ADN (2020) Alzheimers disease diagnosis from diffusion tensor images using convolutional neural networks. PLoS ONE 15(3):0230409

36. Ahmed OB, Benois-Pineau J, Allard M, Catheline G, Amar CB, Initiative ADN et al (2017) Recognition of Alzheimer's disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning. Neurocomputing 220:98–110

37. Perez-Gonzalez J, Jiménez-Ángeles L, Saavedra KR, Morales EB, Medina-Bañuelos V, Initiative ADN et al (2021) Mild cognitive impairment classification using combined structural and diffusion imaging biomarkers. Phys Med Biol 66(15):155010

38. Fang M, Jin Z, Qin F, Peng Y, Jiang C, Pan Z (2022) Re-transfer learning and multi-modal learning assisted early diagnosis of Alzheimers disease. Multimed Tools Appl 2022:1–17

39. Agostinho D, Caramelo F, Moreira AP, Santana I, Abrunhosa A, Castelo-Branco M (2021) Combined structural MR and diffusion tensor imaging classify the presence of Alzheimers disease with the same performance as MR combined with amyloid positron emission tomography: a data integration approach. Front Neurosci 15:638175

40. Petersen RC, Aisen P, Beckett LA, Donohue M, Gamst A, Harvey DJ, Jack C, Jagust W, Shaw L, Toga A et al (2010) Alzheimer's disease neuroimaging initiative (Adni): clinical characterization. Neurology 74(3):201–209

41. Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE (2011) Statistical parametric mapping: the analysis of functional brain images. Elsevier, Amsterdam

42. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62(2):782–790

43. Cusano C, Ciocca G, Schettini R (2003) Image annotation using SVM. Internet imaging V, vol 5304. SPIE, Bellingham, pp 330–338

44. Raju M, Gopi VP, Anitha V, Wahid KA (2020) Multi-class diagnosis of Alzheimers disease using cascaded three dimensional-convolutional neural network. Phys Eng Sci Med 43(4):1219–1228

45. Feng W, Halm-Lutterodt NV, Tang H, Mecum A, Mesregah MK, Ma Y, Li H, Zhang F, Wu Z, Yao E et al (2020) Automated MRI-based deep learning model for detection of Alzheimers disease process. Int J Neural Syst 30(06):2050032

46. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25:1