Magnetic Resonance Imaging 62 (2019) 167–173

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original Contribution

Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images

Tawseef Ayoub Shaikh*, Rashid Ali

Department of Computer Engineering, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India

ARTICLE INFO

Keywords:
Computer-assisted diagnosis (CAD)
Alzheimer's disease
DWT (Discrete Wavelet Transform)
Support Vector Machine (SVM)
Machine intelligence
Classifier

ABSTRACT

An inventive scheme for automated tissue segmentation and classification is offered in this paper using Fast Discrete Wavelet Transform (DWT)/Band Expansion Process (BEP), Kernel-based least squares Support Vector Machine (KLS-SVM) and F-score, backed by Principal Component Analysis (PCA). Using input as T1, T2 and Proton Density (PD) scans of patients, CSF (Cerebrospinal Fluid), WM (White matter) and GM (Gray matter) are afforded as output, which act as hallmark for brain atrophy and thus sustaining in diagnosis of Alzheimer's disease (AD) from Mild Cognitive Impairment (MCI) and Healthy Controls (HC). The blending of BEP features from DWT and texture features from Gray Level Co-occurrence Matrix (GLC) promises to be a savior in atrophy revelation of the segmented tissues. Data used for evaluation of this study is taken from the ADNI database that encloses T1-weighted s-MRI (Structural Magnetic Imaging Resonance) scans of 158 patients with AD and 145 HC. Preprocessing steps unearthed five parameters for classification (i.e. cortical thickness, curvature, gray matter volume, surface area, and sulcal depth), in the preliminary step. For challenging the classifier performance, ROC (Receiver operating characteristics) curves are painted and the SVM classifiers of two-dimensional spaces took the top two important features as classification features for separating HC and AD to the maximum extent. The final results revealed that Fast DWT + F-Score + PCA + KLS-SVM + Poly Kernel is giving 100% tissue classification accuracy for test samples under consideration with only 7 input features.

1. Introduction

Alois Alzheimer, a German neuropathologist, first labeled Alzheimer's disease (AD) as Dementia Praecox in 1907. Now, the present day research witnesses a growing trend in neurologic diseases like Epilepsy, Parkinson's disease (PD), Multiple Sclerosis (MS), Stroke, AD, Myasthenia gravis, etc. More than 600 neurologic diseases have been documented in the medical literature [1]. Dementia is one common disease in the list hurting cognitive abilities like memory and other grave issues enough to thwart regular life. Parkinson's disease dementia, Mixed dementia, Vascular dementia, Dementia with Lewy bodies, Frontotemporal dementia (FTD), Korsakoff syndrome, Huntington's disease dementia, Creutzfeldt-Jakob disease, Normal pressure hydrocephalus, Down syndrome dementia, Posterior cortical atrophy (PCA), etc. are its compound forms [1]. The literature confronts that 44 million people suffer from dementia globally, anticipated to double by 2030 and more than triple by 2050, i.e. 115 million people. 60% demented are affected by Alzheimer's disease and further 1 in 85 people will be its prey by 2050 [1]. 66% of the demented are partaking AD and only 10% of them are diagnosed at the right time and residual 90% will not be diagnosed at all, according to the survey through research group of Alzheimer's Association [2]. America alone is home to 5 million Alzheimer patients, comprising 11% of those with age 65 and older and one-third of those 85 and older. The disease also impacts more than 15 million family members, friends, and caregivers.

Healthcare sector is at the peak of revolution with new innovations particularly in the design of non-invasive machines/techniques which can draw clear HD (High Definition) painting of interiors of the whole body on the wall. These are predominantly used in brain tumors, strokes, blood clots, or other abnormalities that might account for Multiple Sclerosis (MS) or Alzheimer's. Enhanced with high-resolution imaging and high brain tissue contrast capabilities, Structural Magnetic Resonance Imaging (s-MRI) is one of the vital neuroimaging modalities. Truly, MRI scans are architected by pulse sequences and itemized by three varied MR tissue parameters: spin-lattice (T1), spin–spin (T2) relaxation times, and Proton Density (PD) [3]. In a magnetic field, ingredients of brain such as Gray Matter (GM), Cerebrospinal Fluid (CSF), White Matter (WM), Glial Matter, Fat, Muscle/Skin, etc., display idiosyncratic characteristics making extraction of both spatial and tissue characteristics possible from these MRI scans [4]. CSF, GM, and WM,

* Corresponding author.
E-mail address: tawseef37@gmail.com (T.A. Shaikh).

https://doi.org/10.1016/j.mri.2019.06.019
Received 13 February 2019; Received in revised form 2 June 2019; Accepted 23 June 2019

distributed in T1, T2, and PD are key tissue types of brain. Making it non-reproducible and essentially tough task for manual classification of these colossal data amounts of tissues, ascends the need for development of fully automatic and accurate brain tissue classification from MRI images, in case of various disease symptoms like Tumors, MS, AD and other White Matter Lesions (WML) which is at the heart of present-day research [5,6].

1.1. Need for automation

Dealing with such a heavy burden on the society both economically and health-wise, early and precise diagnosis of Alzheimer's disease is the need of the hour. The current refined non-invasive tests used in the studies of Alzheimer's disease (AD) and mild cognitive impairment (MCI) have sightseen varied neuroimaging modalities with promising results. These include structural MRI (s-MRI), functional MRI (fMRI) [7], Flourodeoxy glucose Positron Emission Tomography (FDG-PET) and amyloid PETs such as Pittsburgh compound (PiB-PET) [8]. Existing neuropsychological testing techniques for AD diagnosis like Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), BDIMC, COG SCORE, BOMC, MOCA, AD8, GP COG, and 10/66 research group practice a sequence of cognitive tests comprehending a set of questions and images [8]. Being time-consuming for visualizing/interpreting multimodal data of patients, the precise outcome is directly proportionate to the expertise of clinician/doctor here. Manual segmentation is not lesser than a nightmare keeping in view the perfections accomplished by imaging tools (e.g., MR scanners resolve images at millimetric resolution) over the past years. To excerpt the contours of the target structures, a trained operator has to endure around eighty 512×512 images, slice by slice. Similarly exploring s-MRI with VBM (Voxel-Based Morphometry), 106 brain dimensions are created [9]. Various studies have assessed manual segmentation as not only tedious but also liable to errors [10,11].

All the features generated are not always worthy of the task at hand. Many multivariate approaches show accuracy drop if redundant information exists in features for disease prediction. To pronounce the dataset at least as well as the original set of features, Feature Selection (FS), an energetic research area in machine learning (ML) communities, aims to find a small number of relevant features according to given criterion, maintaining their original values. It is not always the case that more information would lead to higher accuracy. For differentiating both groups, a single feature is not on the road to afford information in many situations. On the other hand, multivariate analyzes contemplate sets of features from small individual contributions, whose integration lead to the identification of different patterns from each group. Thus, not all brain structures are crucial to set apart abnormal patients from healthy controls. These all driving factors created a thrust to develop an automated solution for analysis and interpretation of such multi-featured and multi-modal data. Machine learning has come as a Messiah for relieving this burden, by producing offshoots commonly known to us as Computer-aided/assisted diagnosis (CAD) systems. These systems can handle multi-featured, complex and heterogeneous data at a very fast pace and beat the human brain when it comes to interpreting images. Lot of pioneering work in this direction has already geared up.

Morphological changes in the brain can be measured using manual, semi-automated and fully automated volumetric techniques to study the whole brain and medial temporal volumes. In recent years, major advances in neuroimaging have provided opportunities to study neurological-related diseases, resulting in improvements in early and accurate detection of AD [12–15]. The functional network techniques are used by various researchers in the literature for the automated diagnosis of Alzheimer's disease. Xiaohong Cui et al. [16] proposed a minimum spanning tree (MST) classification framework to identify Alzheimer's disease (AD), MCI, and normal controls (NCs). The proposed method mainly uses the MST method, graph-based Substructure Pattern mining (g-Span), and graph kernel Principal Component Analysis (graph kernel PCA). Specifically, MST is used to construct the brain functional connectivity network g-Span, to extract features and subnetwork selection and graph kernel PCA, to select features. Finally, the support vector machine is used to perform classification. The method was evaluated on MST brain functional networks of 21 AD, 25 MCI, and 22 NC subjects. The experimental results show that the proposed method achieves classification accuracy of 98.3, 91.3, and 77.3%, for MCI vs. NC, AD vs. NC, and AD vs. MCI, respectively. Similar latest graph connectivity work for Alzheimer's disease diagnosis has been done [17–23].

The remainder of the paper is organized as follows: First, in Section 2, data sets, both benchmark and local, used in the study are ignited. Initial preprocessing techniques with proper feature vector extraction from input raw MRI brain images are also part of this domain. The novelty of the work is at the heart of the proposed methodology in Section 3 with a torchlight also thrown on the evaluating parameters used in this paper. The final results in the form of tables, graphs, ROC/AUC curves are drawn in Section 4. The clinical importance of the obtained results is thoroughly discussed here with a comparison of our techniques with existing ones in terms of its improvements. Finally, the output results are concluded in Section 5 and some promising future directions are also highlighted.

2. Methods

Both datasets, benchmark and local, used in the study are the topics of description in this section. Initial preprocessing techniques with proper feature vector extraction from input raw MRI brain images with a flavor of the used machine learning classifier are also part of this domain.

2.1. Data

The data in this research is taken from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) [24]. To quantify the headway of mild cognitive impairment (MCI) and early Alzheimer's disease (AD), using MRI, PET, other biological markers, whether and how clinical and neuropsychological assessments could be combined, is the focal goal of ADNI. The data entailed high-resolution T1-weighted s-MRI of 303 participants. T1-weighted structural image parameters of all participants were roughly labeled as TR = 2400 ms, TE = 4 ms, slice thickness = 1.0 mm, voxel size 1.0 mm × 1.0 mm × 1.0 mm. Cognitive and behavioral assessments of all participants are done before scanning. While relating age and gender, no significant difference ($p > 0.05$) between the HC group and AD group is found. Certain differences between groups for demographics including mini-mental state examination (MMSE), clinical dementia rating (CDR) existed, such as AD had a lower score of MMSE but a higher score of CDR than HC. SPSS 22.0 initially is used for statistical analysis of basic information.

The data is represented as a mean ± standard deviation. *p*-Values are displayed on right columns for two sample *t*-tests for each sample characteristic except for gender, which displays *p*-value from a χ^2 test. The second local dataset for this work was attained from 2 known medical imaging centers of Srinagar, India namely Medicare Diagnostic Center and DM diagnostic clinic over the last 2 years [25]. Images from three sections, i.e., axial, coronal and sagittal compose every study of the dataset. The number of images varies across three sections depending upon the parameter selection of axial resolution and slice thickness. The results are obtained here in two experiments, viz., segmentation of brain tissues and the estimation of brain volume atrophy (Table 1).

2.2. Neuroimaging processing

Initial processing of MRI images mined two types of information, i.e.

Table 1
Mean and standard deviation of sample demographics of ADNI dataset.

Characteristic	AD	HC	*p*-Value
Sample size	158	145	
Gender (male/female)	86/72	71/74	0.342
Age (years)	75.21 ± 7.47	75.76 ± 4.42	0.432
CDR	4.68 ± 1.74	0.04 ± 0.14	< 0.000001
MMSE	23.30 ± 2.05	29.14 ± 0.94	< 0.000001

volumetric features and cortical thickness features. This initial preprocessing was carried with the Freesurfer image analysis suite, which is documented and open source brain suite software (http://surfer.nmr.mgh.harvard.edu). Averaging [26] of multiple volumetric T1 weighted images (when more than one is available), removal of non-brain tissue using a hybrid watershed/surface deformation procedure [27], motion correction, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures (including hippocampus, amygdala, caudate, putamen, ventricles) [28], Image Registration, Longitudinal Processing, intensity normalization, Skull stripping, Subcortical Segmentation, Cortical Thickness Estimation, Cortical Surface Reconstruction, Cortical Segmentation, fMRI Analysis, Tractography, FreeView Visualization GUI, tessellation of the gray matter-white matter boundary, automated topology correction and surface deformation are some of the services offered by this tool. Surface inflation, registration to a spherical atlas which utilizes individual cortical folding patterns to match cortical geometry across subjects [29], fragmentation of the cerebral cortex into units based on gyral and sulcal structure and creation of a variety of surface based data including maps of curvature and sulcal depth [29] are the succeeding deformable procedures after initial cortical processing.

Cortical thickness (CT) and gray matter probability (GMP) are dug out in the s-MRI case. Segmentation and deformation trials from all-inclusive three dimensional MR volume produce representations of Cortical thickness, calculated as the closest distance from the gray/white boundary to the gray/CSF (Cerebrospinal Fluid) boundary at each vertex on the tessellated surface. Using spatial intensity gradients across the tissue, maps are fashioned which are not constrained to the voxel resolution of the original data, and thus are capable of detecting submillimeter differences between groups. The cortical thickness features are average values for each region. A measure of roughness for each cortical region, the standard deviation of the cortical thickness was also calculated additionally. Also, anatomical localization of the cerebral areas of reformed white and gray matter was performed. Individual Brain Atlases using Statistical Parametric Mapping (IBASPM) are used for Volumetric analysis of brain MRIs utilized, which is an extension of SPM-5. The calculation of the brain structure was performed as follows: 1) MRIs were segmented into gray matter, white matter and cerebral spinal fluid (CSF), using IBASPM segmentation. 2) MRI scans were spatially transformed into Montreal Neurological Institute (MNI) space, using affine transformation for approximate registration, and nonlinear transformation for fine registration to obtain the transformation parameters [30].

2.3. Feature reduction

There are two different feature reduction phases in the proposed scheme. Here, we briefly describe these phases. Given that some features are uninformative, irrelevant or redundant for classification, reducing the number of features not only speed up computation but also improve classification performance. Therefore, an initial feature selection step was adopted. The feature ranking approach [31,32], has been widely used in feature selection. F-score is a simple, generally quite effective technique which measures the discrimination of two sets of real numbers and used in previous studies [33,34]. In this study, the F-score method is employed for feature ranking.

Given training vectors, xk, $k = 1, \ldots, m$, if the number of positive instances (i.e., HC) and negative instances (i.e., AD) are n_+ and n_-, respectively, then the F-score of the ith feature is defined as:

$$F(i) = \frac{(\bar{x}_i^{(+)} - \bar{x}i)^2 + (\bar{x}_i^{(-)} - \bar{x}i)^2}{\frac{1}{n+-1}\sum_{k=1}^{n+}(x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n--1}\sum_{k=1}^{n-}(x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2} \quad (1)$$

where $\bar{x}i$, $\bar{x}_i^{(+)}\bar{x}_i^{(-)}$ are the average of the ith feature of the whole, positive, and negative datasets, respectively; $x_{k,\ i}{}^{(+)}$ is the ith feature of the kth positive instance, and $x_{k,\ i}{}^{(-)}$ is the ith feature of the kth negative instance. The numerator indicates the discrimination between the positive and negative sets, and the denominator indicates the one within each of the two sets. The larger the F-score is, the more likely this feature is more discriminative. Therefore, we use this score as a feature selection criterion.

2.4. PCA

The second feature reduction is achieved using PCA [35]. Let *Tvar* be the total variance of the original feature set, and *Svar* be the total variance of the reduced feature set having dimension d, $d < < M^2$. The proper value of d is selected: $\frac{Svar}{Tvar} \approx 0.9$. Let *PERred* be the achieved feature reduction percentage.

$$\therefore PERred = \frac{N^2 - d}{N^2} \times 100\% \quad (2)$$

For a 256-by-256 size T2-weighted MR image, its dimension is $256 * 256 = 65{,}535$; after a three-level DWT decomposition, the dimension of the wavelet coefficients is $32 * 32 = 1024$. Therefore, it is still a high computation cost if we directly submit the 1024 dimensional data to classifier. In this paper, we used PCA to reduce the 1024 dimensional data to only 7 principal components. The primary function of PCA is to reduce a larger set of variables into a smaller set of' artificial variables (called principal components) that account for most of the variance in the original variables [36]. Variables are supposed to be continuous and have linear relationships [36], and PCA reduces dimensionality based on the relationships among variables. The major strength of PCA is that it is based on whole brain voxel data. Several machine learning studies have used PCA to reduce dimensionality.

3. Methodology

This section is the backbone of the present work in in the form of Proposed Methodology with evaluating parameters. The novelty of the present work is at the heart of this section.

3.1. Proposed methodology

An innovative automated tissue segmentation and classification scheme established on F-score backed with Principal Component Analysis (PCA), fast Discrete Wavelet Transform (DWT)/Band Expansion Process (BEP), and least squares based Kernel Support Vector Machine (LSK-SVM) classifier is at the heart of this paper. Using input as T1, T2 and Proton Density (PD) scans of patients, CSF (Cerebrospinal Fluid), WM (White matter) and GM (Gray matter) are afforded as output, which act as the hallmark for brain atrophy, thus sustaining in diagnosis of Alzheimer's disease (AD) from Mild Cognitive Impairment (MCI) and Healthy Controls (HC). Blending BEP features from DWT with texture features from Gray Level Co-occurrence Matrix (GLCM) promises to be a Messiah in revealing the atrophy of segmented tissues. T1-weighted s-MRI scans of 158 patients with the AD and 145 HC are engaged from the ADNI database. Preprocessing steps unearthed five parameters (i.e. cortical thickness, curvature, gray matter volume, surface area, and sulcal depth) in the preliminary step. The Fast DWT + F-Score + PCA + KLS-SVM + Poly Kernel is giving 100% tissue classification accuracy for test samples under consideration. MRI

raw images are made input initially followed by DWT and FreeSurfer operations in order to extract the valuable features using Gray Level Co-occurrence Matrix (GLCM) and texture and volumetric features. Corresponding acquired features from the above two steps are integrated and in the succeeding steps, feature reduction using F-score backed by PCA is applied in order to extract the most relevant and meaningful features for optimal classification in the following steps. The feature vector is constructed from the optimal relevant features obtained in the previous step and used as a tool for training KLS-SVM in order to make a differentiation between AC and HC patients. The evaluation of the trained classifier is performed using new test instances of MRI scans and noting down the corresponding classification accuracies.

3.2. Evaluating parameters

The recall is the probability that a diagnostic test is positive, given that the person has the disease.

$$\text{Recall or True Positive Rate } (TPR) = \text{Sensitivity} = \frac{\sum \text{Truepositive}}{\sum \text{Conditionpositive}} = \frac{TP}{TP + FN} \tag{3}$$

Specificity is the probability that a diagnostic test is negative, given that the person does not have the disease.

$$\text{True Negative Rate } (TNR)\ [36] = \text{Specificity} = \frac{\sum \text{Truenegative}}{\sum \text{Conditionnegative}} = \frac{TN}{TN + FP} \tag{4}$$

Accuracy is the probability that a diagnostic test is correctly performed.

$$\text{Accuracy } [36] = \frac{TP + TN}{TP + TN + \text{FP} + \text{FN}} \tag{5}$$

Also, ML models used in this study are assessed with the Receiver Operating Characteristics (ROC) [37,38]. The area under the curve (AUC) has a public use for performance evaluation of clinical diagnostic and predictive models [39] while creating a tradeoff between the sensitivity (true positive rate) and the specificity (true negative rate). Percentage of positive instances truthfully categorized as positive is known as sensitivity and specificity calculates the percentage of negative instances which were truthfully labeled as negative. ROC curve designs the sensitivity against 1-specificity (false positive rate) while evaluating a classifier [39]. That curve illustrates the development of the classification threshold from the very positive threshold where every instance is labeled as positive to the very negative threshold where every instance is labeled as negative. The confidence score of the classifier is beneath the set threshold in the setting when the sensitivity of a classifier is 0.0 and the specificity being 1.0. Contrariwise, the confidence score of the classifier is overhead the set threshold provided specificity is 0.0 and sensitivity is 1.0. Diagonal line relating origin (0, 0) to the final point (1, 1) possessing AUC of 0.5, are characteristics of a random classifier. Similarly, a perfect classifier possesses AUC of 1.0 with a starting of ROC curve from (0, 0) to (1, 0) [39], and has the signature in ranking all positive instances above all negative instances. AUC fundamentally is like Wilcoxon-Mann-Whitney statistic [40], which has an appearance that a classifier ranking of randomly selected positive cases is higher than randomly selected negative cases. While diverse tradeoff arises from diverse clinical diagnostic scenarios, an AUC that is greater than 0.75 is usually recommended for clinical purposes [41]. The AUC is defined as follows:

$$c(p,n) = \begin{cases} 1 \text{ if } c(p) > c(n) \\ 0 \text{ otherwise} \end{cases}$$

$$AUC(c,P,N) = \frac{1}{|P||N|} \sum_{p \varepsilon P} \sum_{n \varepsilon N} c(p,n) \tag{6}$$

where $c(e)$ being the classifier's confidence score, P is a set of positive examples, and N is a set of negative (HEG) examples. Applicability of ML in clinical practice is not governed by the significance of accuracy indices being above chance level. Training and testing of our pattern recognition algorithms were on groups enjoying the relatively modest size and therefore, achieved accuracies from those samples are not unswervingly archetypal for predictions in the clinical world [42,43]. Performance indices with reasonable accuracy gained in the present study are reliable with the notion that there are important limitations in the application of ML techniques for predicting diagnosis in clinical neuroimaging research.

4. Discussion and experimental results

The final results in the form of tables, graphs, ROC/AUC curves are drawn here in the present section. Also, the clinical importance of the obtained results is thoroughly discussed in addition to a comparison of existing techniques with that of ours in terms of their improvements.

4.1. Experimental setup

Using PC with Intel Core i5-4590 processor, 3.30 GHz, 8 GB of RAM and 64 Bit Operating System and in MATLAB environment, the offered technique and experiments are probed on two different MR image datasets. Partaking in-plane resolution of $256 * 256$, both datasets consist of T1-weighted MR brain images, which were downloaded from the ADNI website (adni.loni.usc.edu). The first dataset being benchmark dataset, extensively castoff in brain MR image classification problem, and consist of abnormal images of AD diseases besides normal images of HC. The benchmark dataset entails of high-resolution T1-weighted s-MRI of 303 participants with 158 AD and 145 HC. Several prevailing schemes are instigated for equating the performance of our proposed method with contemporary techniques. Several different experiments are conceded for assessing the performance of the proposed system in light of feature reduction, efficiency, classification accuracy, comparisons with other state of the art schemes and computation complexity analysis.

4.2. Results and discussions

The 10-fold cross-validation strategy is performed for all algorithms, and this process is repeated 5 times independently so as to avoid the sampling bias introduced by randomly partitioning dataset in the cross-validation. The backbone for the proposed system is the power of Fast DWT + F-Score + PCA + KLS-SVM + Poly kernel for tissue segmentation and classification. Conserving 90% of the total variance of the disintegrated features, PCA yields feature vector of size 7 only, which being the first 7 principal components and outputs feature set of 0.88% and 0.014% of the initial feature set. Therefore, due to the Fast DWT + F-Score + PCA + KLS-SVM + Poly kernel combination, we have achieved a 99% feature reduction (See Table 2).

Triumphing of 47.39% feature reduction from the state of the art brain MR image classification techniques is not the only contribution of this work but also higher performance in terms of accuracy lies at its core. To discover the suitable number of principal components which contributed for best results, the performance of the proposed system experimented with different numbers of principal components. It is clear from the results given in Fig. 1, that our proposed system competently achieved paramount results in terms of sensitivity (1.00, 1.00), specificity (1.00, 1.00) and classification accuracy (100%, 100%), using only 7 principal components for image representation. The classification accuracy of the proposed system can also be evaluated through receiver operating characteristic (ROC) curves, shown in Fig. 1 below.

With an average area under the curve (AUC) of 100% and 0% standard deviation, the specified proposal correctly classified the MRI images of both datasets. To challenge the strength and offers of our proposed model, we paralleled its performance with 11 state of the art brain MR image classification schemes. The comparison results plus feature vector dimension for each of the schemes are shown in Table 3

Table 2
Automated AD classification using KLS-SVM classifier.

Test images from different models	Average classification with KLS-SVM		
	Linear	SVM	Polynomial
Texture/GLCM	77.13	80.27	100.00
DWT/BEP	85.31	89.17	100.00
Fast DWT + F-Score + PCA + KLS-SVM + Poly	100.00	100.00	100.00

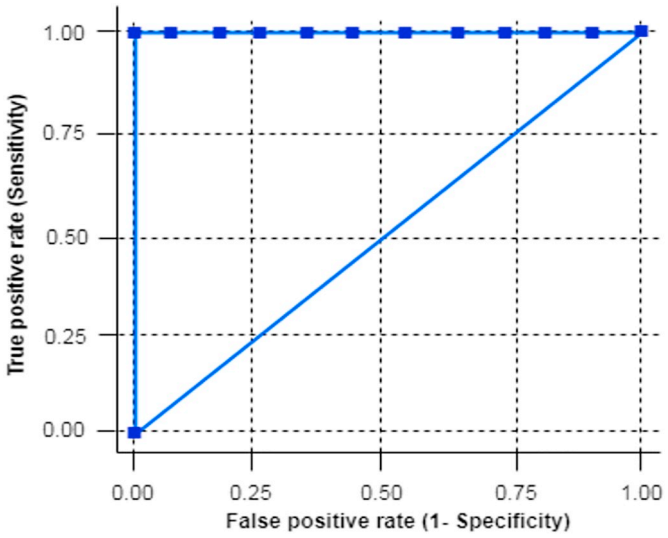

Fig. 1. Performance evaluation using ROC curves.

Table 3
Performance comparison using two different dataset groups.

Approach	Feature dimension	Accuracy (%)	
		Image Group 1	Image Group 2
DWT + SVM + POLY [44]	4761	98.00	97.15
DWT + SVM + RBF [44]	4761	98.00	97.33
DWT + PCA + FNN + ACPSO [45]	19	100.00	98.75
DWT + PCA + BPNN + SCG [46]	19	100.00	98.29
DWT + PCA + FNN + SCABC [47]	19	100.00	98.93
DWT + PCA + KSVM + IPOLY [48]	19	100.00	98.12
DWT + PCA + k-NN [49]	7	98.00	97.54
RT + PCA + LS-SVM + RBF [50]	9	99.39	96.47
Fast DWT + PCA + LS-SVM + RBF [51]	8	99.56	96.78
PLS + PCA [52]	10	84.59	85.7
Curve let + PCA + KNN [53]	10	89.47	87.10
LUPI-DL [54]	93	91.45	88.70
MM-SDPN-SVM [55]	93	97.13	96.93
Fast DWT + F-Score + PCA + LS-SVM + POLY (Proposed)	7	100.00	100.00

below. Worst accuracy performance results are described in [44] with the highest feature dimension (4761 features/image) also, thus results in high computational complexity. The techniques described in [46–48] show amended results in brain MR image classification, with lower feature vector dimension (19), but these schemes use various complex weight optimization techniques, which themselves require high computational complexity. Feature vector dimension (7), used in [49], is equal to our proposed method (7), but it is obvious from the results of Table 3, that the method described in [49] is less effectual and general than the proposed scheme, in terms of classification accuracy whereas, our proposed system only requires a feature vector of dimension 7, with the highest retrieval accuracies. Table 3 below shows a comparison of other techniques with our proposal [50–53]

The results of the proposed method when compared with the much-hyped deep learning methods for AD diagnosis from neuroimaging data, again showed marginal results [54,55]. Yan Li et al. [54] proposed a deep learning LUPI-based CAD framework for AD/MCI diagnosis using an approach of Transfer Learning (Learning using privileged information (LUPI)). The LUPI classifiers are used at the classifier level and deep learning at the feature level. Using a single model PET data and MRI as PI (privileged information), the authors got marginal performances from their proposed work. The deep learning algorithms MDL-CW, MRBM, and DCCA algorithms outperform the classical SVM classifier. DCCA achieves the overall best performance with the best classification accuracy, sensitivity, YI, NPV, and F1 score of 91.45 ± 5.42%, 90.60 ± 10.71%, 82.73 ± 11.02%, 92.34 ± 8.52%, and 91.19 ± 5.68%, respectively, and the second-best specificity and PPV of 92.13 ± 11.30% and 93.44 ± 9.14%, respectively [54].

Similarly, J. Shi et al. [55] proposed a multimodal stacked deep polynomial network (MM-SDPN), which consists of two-stage SDPNs, to fuse and learn feature representation from multimodal neuroimaging data for AD diagnosis. Experimental results indicate that MM-SDPN is superior over the state-of-the-art multimodal feature learning-based algorithms for AD diagnosis, like Multiple kernel learning (MKL), Multi-tasking learning (MTL), Deep belief networks (DBN), Multimodal restricted Boltzmann machines (M-RBM), Stacked auto encoders (SAE), Nonlinear graph fusion (NGF). In AD vs. NC case, the proposed MM-SDPN algorithm achieves the best performance with a mean classification accuracy of 97.13 ± 4.44%, sensitivity of 95.93 ± 7.84% and specificity of 98.53 ± 5.05%, in comparison with single modal deep learning algorithms like DPN-3-MRI, DPN-6-MRI, SDPN-PET, etc. Both of the works extracted a feature vector of 93 features whereas, our method relies only on 7 extracted features and outperforms the said methods [54,55]. Also, the deep learning complex network increases the computational complexity which makes our proposed method relatively faster (Table 3).

The atrophy change in various brain sub-structures revealed from our method is in close agreement with the findings in previous studies. The results disclosed smaller volumes of ptIFC (related with inhibitory control processes) in AD patients than in HC controls in both hemispheres and the left triangular shows diminution conferring to the group level (HC > MCI > AD). Postcentral gyrus with abridged white matter (WM) and volume below the right postcentral gyrus are other zones of enormous focus in AD patients [56,57]. Bilateral insula, right superior temporal gyrus, bilateral anterior cingulate cortex, and left superior medial frontal cortex are usually abridged gray matter provinces obtained in AD results [58,59]. Cortical thickness measured by voxel-based morphometry (VBM) embodies the shortest distance between the boundary of gray/white matter and the pial surface, whereas gray matter volume denotes the amalgamation of two genetically independent anatomical properties [60,61]. The cortical thickness of the bilateral entorhinal, bilateral paracentral, left medial orbitofrontal, left superior parietal, right cuneus, right postcentral was thinner in AD compared with HC. In the surface area, AD has smaller areas than HC in the left inferior temporal, left precuneus, right entorhinal, right banksters, right inferior parietal and have greater areas than HC in the left precentral, left superior parietal, right insula. Compared with HC, the gray matter volume of the bilateral entorhinal, bilateral precentral, left

Table 4
Radiologist's feedback for the second local dataset.

Methods	Acceptable			Improvement needed			Unacceptable		
	CSF	GM	WM	CSF	GM	WM	CSF	GM	WM
Alzheimer image set1									
Texture features				✓		✓			✓
DWT				✓	✓				
Texture + DWT	✓	✓	✓						
Alzheimer image set2									
Texture features				✓	✓				✓
DWT	✓				✓	✓			
Texture + DWT	✓	✓	✓						
Alzheimer image set3									
Texture features				✓	✓	✓			✓
DWT	✓				✓	✓			
Texture + DWT	✓	✓	✓						

superior frontal, left caudal middle frontal, right posterior cingulate and right rostral middle frontal was smaller in the AD.

In the curvature, HC has smaller curvatures than an AD in the bilateral parahippocampal, bilateral superior temporal, left insula, left entorhinal, right supramarginal and right temporal pole. Moreover, AD has greater depths than HC in the bilateral posterior cingulate, left insula, left rostral anterior cingulate, left posterior cingulate, right superior temporal, right parahippocampal and have smaller curvatures than HC in the left paracentral, right banksters in the sulcal depth [62,63]. Severe shrinking in cortical thickness, surface area and gray matter volume of the entorhinal cortex and atrophy predominantly appears in the temporal lobe, frontal lobe, occipital lobe, parietal lobe and cingulate gyrus in the AD, which almost reach for agreement with pre-existing literature involved in entorhinal cortex's atrophy [64,65]. Interim, the surface areas of left precentral, left superior parietal and right insula show a small intensification in the AD. Evolving indication shows altered neurogenesis in the adult hippocampus epitomizes an early critical event in the course of AD [66,67].

The results obtained in case of local dataset indicate the efficacy of the segmentation method with an accuracy of 92.18 ± 5.32, using our proposed method in comparison to an accuracy of 81.76 ± 9.20 for AD/HC. Based on the overlap index, the proposed segmentation method shows a value of 93.19 ± 3.62 compared to an overlap value of 78.66 ± 7.3 for AD/HC. The segmented ROI is then used for the precise estimation of the AD volume from brain MRI images. Also, the radiologist's feedback has encouraged the worthiness of our proposed method. (See Table 4).

5. Conclusion

Diagnosing human brain disorders from MR images manually is not only the toughest task because of its highly complex dimensionality but is also very prone to human errors, making it almost a hard problem. This creates a thirsting obligation for emerging diagnostic tools, possessing the capability of automatically and accurately classifying brain MR images as normal or abnormal. We put forward our effort toward the same need by proposing a novel method of combining the strengths of Fast DWT + F-Score + PCA + Poly Kernel, and a computationally less expensive KLS-SVM to physique a fully automatic and accurate brain MR image classification system. With this combination, we not only achieve higher feature reduction but also acquire superior performance than the state of the art schemes. Extensive experiments and comparisons show the effectiveness of the proposed system. Using supervised and unsupervised classification schemes for brain MR image classification, the future is promising for investigating the effectiveness of other transforms along the same dimension.

Author contributions

Mr. Tawseef Ayoub Shaikh conceived, designed and performed the experiments. Mr. Rashid Ali analyzed the data and he offered his valuable suggestions and feedbacks. Both contributed to reagents/materials/analysis tools and paper writing jointly.

Conflicts of interest

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgment

This work is partly supported by Research Fellowship of the “Visvesvaraya Ph.D. Scheme for Electronics & IT”, Ministry of Electronics and Information technology (MeitY), Government of India (GoI), Vide Grant no. PHD-MLA/4(39)/2015-16. Authors are very much thanked full to doctors/radiologists and staff team of both medical imaging centers i.e. Medicare Diagnostic Center and DM diagnostic clinic, Srinagar India for verifying all the tissue segmentation results carefully and rating the results on the scale of ‘Acceptable’, ‘Needs Improvement’ and ‘Not Acceptable.

References

[1] Fox NC, Crum VR, Scahill RI, Stevens JM, Janssen JC, Rossar MN. Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. Lancet 2001;358:201–5.
[2] Thies W, Bleiler L. 2013 Alzheimer's facts and figures, Alzheimer's $ dementia. J Alzheimer's Assoc 2013;9(2):208–45.
[3] Ouyang YC, Chen HM, Chai JW, Chen C, Poon SK, Yang CW, et al. Band expansion based over complete independent component analysis for multispectral processing of magnetic resonance images. IEEE Trans Biomed Eng 2008;55(6):1666–77.
[4] Ouyang YC, Chen HM, Chen C, Poon SK, Yang CW, Lee SK. Independent component analysis for magnetic resonance image analysis. EURASIP J. on Adv. in Sig. Proc., Hindawi 2008;2008(780656):1–14.
[5] Cocosco CA, Zijdenbos AP, Evans AC. A fully automatic and robust brain MRI tissue classification method. Med. Imag. Anal. 2003;7(4):513–27. Elsevier.
[6] Wang J, Chang CI. Independent component analysis-based dimensionality reduction with applications in hyper spectral image analysis. IEEE Trans Geo Remote Sens 2006;44(6):1586–600.
[7] Lenzi D, Serra L, Perri R, Pantano P, Lenzi GL, Paulesu E, et al. Single domain amnestic MCI: a multiple cognitive domains fMRI investigation. Neurobiol. Aging 2011;32(9):1542–57. Elsevier.
[8] Harrison J, et al. Cognitive approaches to early Alzheimer's disease diagnosis. Med Clin North Am 2013;97(3):425–38. Elsevier.
[9] Keogh E, Mueen A. Curse of dimensionality. Encyclopaedia of machine learning and data mining. Springer; 2017. p. 257–68.
[10] Collier DC, Burnett SSC, Amin M. Assessment of consistency in contouring of normal-tissue anatomic structures. J Appl Clin Med Phy 2003;4(1):17–24.
[11] Vansteenkiste E, et al. Quantitative analysis of ultrasound images of the preterm brain Ph.D. dissertation Ghent University; 2007. [Last visited 5-Nov-2018].
[12] Hanyu H, Sato T, Hirao K, Kanetaka H, Iwamoto T, Koizumi K. The progression of cognitive deterioration and regional cerebral blood flow patterns in Alzheimer's

disease: a longitudinal SPECT study. J Neurol Sci 2010;290:96–101. Elsevier.
[13] Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. NeuroImage. 60(1). Elsevier; 2012. p. 221–9.
[14] Chen YJ, Deutsch G, Satya R, Liu HG, Mountz JM. A semi-quantitative method for correlating brain disease groups with normal controls using SPECT: Alzheimer's disease versus vascular dementia. Comput Med Imaging Graph 2013;37(1):40–7. Elsevier.
[15] Liu F, Zhou L, Shen C, Yin J. Multiple kernel learning in the primal for multi-modal Alzheimer's disease classification. IEEE J Biomed Health Infor 2014;18(3):984–90.
[16] Xiaohong C, Jie X, Hao G, Guimei Y, Huijun Z, Fangpeng L, et al. Classification of Alzheimer's disease, mild cognitive impairment, and normal controls with sub network selection and graph kernel principal component analysis based on minimum spanning tree brain functional network. Front Comput Neurosci Methods 2018;12(31):1–11.
[17] Hanyu H, Sato T, Hirao K, Kanetaka H, Iwamoto T, Koizumi K. The progression of cognitive deterioration and regional cerebral blood flow patterns in Alzheimer's disease: a longitudinal SPECT study. J Neuro Sci 2010;290:96–101. Elsevier.
[18] Jie B, Wee Y, Shen D, Zhang D. Hyper-connectivity of functional networks for brain disease diagnosis. Med. image anal. 32. Elsevier; 2016. p. 84–100.
[19] Jie B, Zhang D, Gao W, Wang Q, Wee CY, Shen D. Integration of network topological and connectivity properties for neuroimaging classification. IEEE Trans Biomed Eng 2014;61:576–89.
[20] Ashraf J, Ahmad J, Ali A, Zaheer UH. Analyzing the behavior of neuronal pathways in Alzheimer's disease using petri net modeling approach. Front Neuroinform 2018;12(26):1–24.
[21] Hart B, Cribben I, Fieca M. A longitudinal model for functional connectivity networks using resting-state fMRI. 2018;178(687):687–701. NeuroImage, Elsevier.
[22] Tijn MS, Koini M, Vosa F, Seiler S, Rooij M, Lechner A. Individual classification of Alzheimer's disease with diffusion magnetic resonance imaging. NeuroImage. 152. Elsevier; 2017. p. 476–81.
[23] Vos F, Koini M, Tijn MS, Seiler S, Grond J, Lechner A. A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. NeuroImage. 167. Elsevier; 2018. p. 62–72.
[24] http://adni.loni.usc.edu/ [Last accessed 10-11-2018].
[25] Banday SA, Mir AH. Statistical textural feature and deformable model based brain tumor segmentation and volume estimation. Multimed tools appl. Springer; 2016.
[26] M. Reuter, H. D. Rosas, and B. Fischl B, Highly accurate inverse consistent registration: A robust approach, NeuroImage, Elsevier, 53(7), (2010), 1181–1196.
[27] Segonne F. A hybrid approach to the skull stripping problem in MRI. NeuroImage. 22. Elsevier; 2004. p. 1060–75.
[28] Fischl B. Sequence-independent segmentation of magnetic resonance images. NeuroImage. 23. Elsevier; 2004. p. 69–84.
[29] Desikan RS, et al. An automated labelling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 31. Elsevier; 2006. p. 968–80.
[30] https://omictools.com/image-segmentation-2-category?tab=software [Last accessed 10-11-2018].
[31] Chu C, Hsu LA, Chou KH, Bandettini P, Lin C, Initiative ADN. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage 2012;60(1):59–70. Elsevier.
[32] Liu F, Suk HI, Wee CY, Chen H, Shen D. High-order graph matching based feature selection for Alzheimer's disease identification. Med image computing and com-assisted interv—MICCAI 2013. Berlin: Germany, Springer; 2013. p. 311–8.
[33] Gao J, et al. A novel approach for lie detection based on F-score and extreme learning machine. PloS One 2013;8(6).
[34] Liu F, et al. Multivariate classification of social anxiety disorder using whole-brain functional connectivity. Brain Struct Function 2015;220(1):101–15.
[35] I. T. Jollie, et al., Principal component analysis, Springer, 2002.
[36] Kim YK, Sae KN. Application of machine learning classification for structural brain MRI in mood disorders: a critical review from a clinical perspective. Prog Neuropsychopharmacol Biol Psychiatry 2017;3:71–80. Elsevier.
[37] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf process manag. 45. Elsevier; 2009. p. 427–37.
[38] https://www.cs.cornell.edu/courses/cs578/2003fa/performance_measures.pdf. [Last Visited 15-11-2017].
[39] Segonne F, Pacheco J, Fischl B. Geometrically accurate topology-correction of cortical surfaces using non separating loops. IEEE Trans Med Imaging 2007;26:518–29.
[40] Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis: inflation, flattening, and a surface-based coordinate system. NeuroImage. 9. Elsevier; 1999. p. 195–207.
[41] Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: segmentation and surface reconstruction. NeuroImage. 9. Elsevier; 1999. p. 179–94.
[42] Chapelle O, Haffner P, Vapnik VN. Support vector machines for histogram-based image classification. IEEE Trans Neural Netw 1999;10(5):1055–64.
[43] Hofmann T, Scholkopf B, Smola AJ. Kernel methods in machine learning. Ann Stat 2008;36(3):1171–220. Inst Mathematical Stat, United States.
[44] Carlos A, Eric W, Sebastian MJ, Patrizia M, Bruno V, Magda T, et al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. Psychiatry res.: neuroimaging. 212. Elsevier; 2013. p. 89–98.
[45] Rowayda et al., Regional atrophy analysis of MRI for early detection of Alzheimer's disease, Image Process Pattern Recognit, Wiley, 6 (2013), no. 1, 49–58.
[46] Muwei L, Yuanyuan Q, Fei G, Wenzhen Z, Xiaohai H. Discriminative analysis of multivariate features from structural MRI and diffusion tensor images. Magn Reson Imaging 2014;32:1043–51. Elsevier.
[47] Yingling Z, Shenquan L. Analysis of structural brain MRI and multiparameter classification for Alzheimer's disease. Biomed Eng-Biomed Tech 2017;63(4):1–11.
[48] Qing L, Xia W, Lele X, Kewei C, Li Y, Rui L. Multi-modal discriminative dictionary learning for Alzheimer's disease and mild cognitive impairment. Comput Methods Prog Biomed 2017;150(18):1–8. Elsevier.
[49] Chaplot S, Patnaik LM, Jagannathan NR. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomed Signal Process Control 2006;1(1):86–92. Elsevier.
[50] Zhang Y, Wang S, Wu L. A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO. Prog Electromagn Res 2010;109:325–43.
[51] Zhang Y, Dong Z, Wu L, Wang S. A hybrid method for MRI brain image classification. Expert Syst Appl 2011;38(8):10049–53.
[52] Zhang Y, Wu L, Wang S. Magnetic resonance brain image classification by an improved artificial bee colony algorithm. Prog Electromagn Res 2011;116:65–79.
[53] Zhang Y, Wu L. An MR brain images classifier via principal component analysis and kernel support vector machine. Prog Electromagn Res 2012;130:369–88.
[54] Li Y, Meng F, Shi J, A. D. N. I. Learning using privileged information improves neuroimaging-based CAD of Alzheimer's disease: a comparative study. Med Biol Eng Comput 2019. https://doi.org/10.1007/s11517-019-01974-3. Springer.
[55] Shi J, Zheng X, Li Y, Zhang Q, Ying S. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE J Biomed Health Inform 2018;22(1):173–83.
[56] Dahshan El, Hosny ESA, Salem ABM. Hybrid intelligent techniques for MRI brain images classification. Digit Signal Process 2010;20(2):433–41.
[57] Khedher L, Ramırez J, Gorriz JM, Brahim A, Segovia F, Initiative ADN. Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. Neurocomputing 2015;151:139–50. Elsevier.
[58] Jha D, Kwon GR. Alzheimer disease detection in MRI using curvelet transform with KNN. J Korea Inst Info Tech 2016;14(8):121–9.
[59] Sudeb D, Manish C, Kundu MK. Brain MR image classification using multi-scale geometric analysis of ripplet. Prog Electromagn Res 2013;137:1–17.
[60] Siddiqui MF, Reza AW, Kanesan J. An automated and intelligent medical decision support system for brain MRI scans classification. Plos One 2015;0135875:1–16.
[61] Wolfers T, Buitelaar JK, Beckmann C, Franke B, Marquand AF. From estimating activation locality to predicting disorder: a review of pattern recognition for neuroimaging-based psychiatric diagnostics. Neurosci Biobehav 2015;57:328–49.
[62] Kim YK, Na KS. Application of machine learning classification for structural brain MRI in mood disorders: a critical review from a clinical perspective. Prog Neuropsychopharmacol Biol Psychiatry 2017;80:71–80. Elsevier.
[63] Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia—meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009;119:252–65.
[64] Keith AJ, Nick CF, Reisa AS, William EK. Brain imaging in Alzheimer disease, Cold Spring Harb (CSH). Perspect Med 2012;2:1–23.
[65] Ramesh KL, Jeonghwan G, Jeong-Seon P, Sang WL. Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features. J Healthcare Eng 2017:1–11. Hindawi. no. 5485080.
[66] Qian MS, David A, Elizabeth P, Weizhao Z, Jason A, Maria T, et al. Volumetric and visual rating of MRI scans in the diagnosis of amnestic MCI and Alzheimer's disease. Alzheimers Dement 2011;7(4):1–17.
[67] Muhammad FS, Ahmed WR, Jeevan K. An automated and intelligent medical decision support system for brain MRI scans classification. Plos One 2015;10(8):1–16.