

A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning

Yousra Asim¹ | Basit Raza¹ | Ahmad Kamran Malik¹ | Saima Rathore² | Lal Hussain³ | Mohammad Aksam Iftikhar⁴ 

¹Department of Computer Science, COMSATS Institute of Information Technology, Islamabad, Pakistan

²Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

³Department of Computer Science and Information Technology, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan

⁴Department of Computer Science, COMSATS Institute of Information Technology, Lahore, Pakistan

Correspondence

Mohammad Aksam Iftikhar, Department of Computer Science, Comsats Institute of Information Technology, Lahore, Pakistan.

Email: aksam.iftikhar@gmail.com

Abstract

The use of biomarkers for early detection of Alzheimer's disease (AD) improves the accuracy of imaging-based prediction of AD and its prodromal stage that is mild cognitive impairment (MCI). Brain parcellation-based computer-aided methods for detecting AD and MCI segregate the brain in different anatomical regions and use their features to predict AD and MCI. Brain parcellation generally is carried out based on existing anatomical atlas templates, which vary in the boundaries and number of anatomical regions. This work considers dividing the brain based on different atlases and combining the features extracted from these anatomical parcellations for a more holistic and robust representation. We collected data from the ADNI database and divided brains based on two well-known atlases: LONI Probabilistic Brain Atlas (LPBA40) and Automated Anatomical Labeling (AAL). We used baseline images of structural magnetic resonance imaging (MRI) and ¹⁸F-fluorodeoxyglucose positron emission tomography (FDG-PET) to calculate average gray-matter density and average relative cerebral metabolic rate for glucose in each region. Later, we classified AD, MCI and cognitively normal (CN) subjects using the individual features extracted from each atlas template and the combined features of both atlases. We reduced the dimensionality of individual and combined features using principal component analysis, and used support vector machines for classification. We also ranked features mostly involved in classification to determine the importance of brain regions for accurately classifying the subjects. Results demonstrated that features calculated from multiple atlases lead to improved performance compared to those extracted from one atlas only.

KEYWORDS

AAL atlas, Alzheimer disease, LPBA40 atlas, mild cognitive impairment, SVM classification

1 | INTRODUCTION

Alzheimer's disease (AD) is the most prevalent form of dementia in older people across the globe. It has been stated that the people affected with AD are expected to be double in the upcoming two decades, and every one person out of a total of eighty-five will have some form of dementia by 2050.¹ Therefore, it is highly desirable to diagnose AD as early as possible, especially in its early stage also known as amnesic mild cognitive impairment (MCI). The early detection is

important in the sense that it can help the physicians in accurately deciding the treatment plans.

AD brain shows atrophy and altered metabolic rate of glucose in several brain regions such as hippocampus and amygdala, as shown in Figure 1. Several classification methods have been proposed by the researchers in the past for automated discrimination of AD and MCI from cognitively normal (CN). The commonly used modalities for AD detection are structural MRI,^{2–5} functional MRI,^{6,7} ¹⁸F-fluorodeoxyglucose positron emission tomography (FDG-PET),^{8–10} single photon

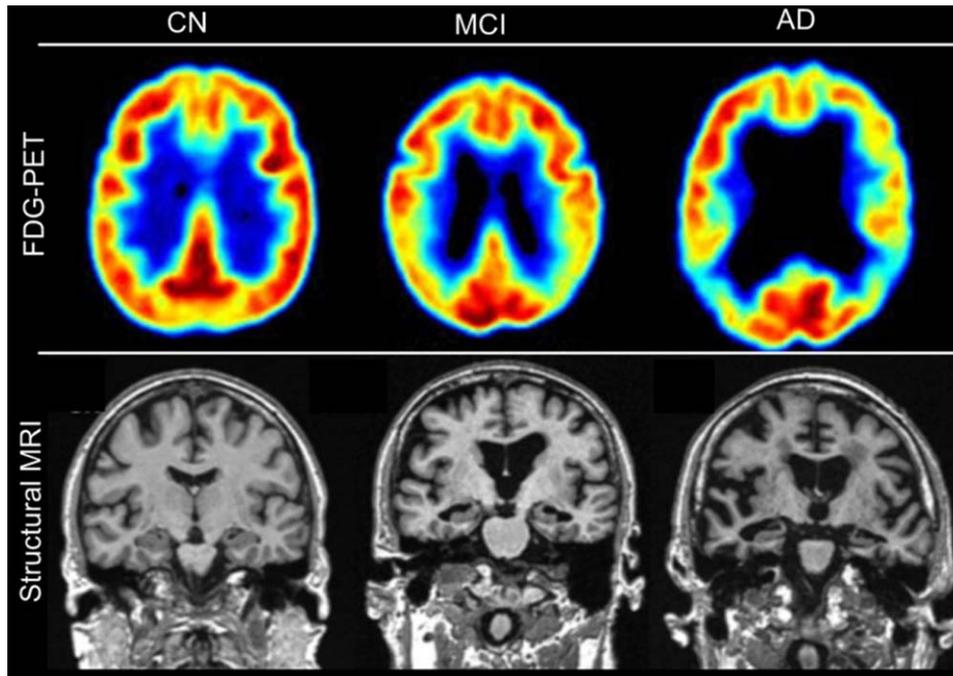


FIGURE 1 The atrophy in AD and MCI shown by structural MRI scans, and altered metabolic rate of glucose in AD and MCI shown by FDG-PET scans [Color figure can be viewed at wileyonlinelibrary.com]

emission computed tomography¹¹ and diffusion tensor imaging.^{12,13} A recent survey covers most of the techniques developed using these modalities.¹⁴ Some researchers have also combined the biomarkers from various modalities with genetic information and demographics to get better classification accuracy.^{15,16} A major subset of these techniques is in fact based on structural MRI, which can further be divided into two major categories: “*voxels as features-based methods*” and “*atlas based-methods.*” In “*voxels as features-based methods,*”^{2,3} all the voxels of brain are used as features, whereas in “*atlas based-methods,*” the brain is parcellated into several anatomical regions, and features such as voxel-based morphometry¹⁷ based gray-matter (GM) and white matter (WM) volumes, and cortical thickness are extracted from these anatomically defined regions. Several different atlases have been used in the past for brain parcellation, and each atlas divides brain into different anatomical regions. The atlas-based methods have been proven to be quite effective for accurate detection of AD and MCI, and can further be divided into two types; fixed and adaptive atlas-based methods. In “*fixed atlas based methods,*”^{4,18} brains are parcellated into anatomical regions based on fixed pre-defined atlases, and features are extracted from those regions. However, in “*adaptive atlas based methods,*”^{19,20} adaptive regions are calculated based on the subjects involved in one particular study, and features are calculated from the adaptively calculated regions.

It has been shown in the past that different brain parcellations using different anatomical atlas templates lead to different classification success rate.^{21,22} In fact, each brain atlas divides brain into unique set of regions. The parcellation

achieved using one anatomical atlas template differs in terms of the number of regions, and the location and size of these regions in the brain. So, each atlas template captures the features of brain from a unique perspective. The atrophy features extracted from multiple atlases provide complementary information. Here, we hypothesize that higher classification accuracy can be achieved if we combine the atrophy patterns extracted based on the parcellation of brain using more than one atlas.

In this work, we have divided the brains into two well-known anatomical atlas templates that is Automated Anatomical Labeling (AAL)²³ and LONI Probabilistic Brain Atlas (LPBA40),²⁴ and extracted regional features from structural MRI and FDG-PET modalities. We have combined the features extracted from both the atlas templates to utilize the complementary information provided by these templates. Discriminating regions of both the atlas templates were also analyzed.

The rest of the article is structured in the following sequence. Proposed method is given in Section 2 of the manuscript. Sections 3 and 4, respectively, present performance measures and experimental results in detail. Conclusion is provided in Section 5.

2 | PROPOSED METHODOLOGY

In this article, we have proposed an AD and MCI classification technique that comprises several steps such as image acquisition, pre-processing, brain parcellation, feature extraction, feature reduction and classification. The sequential flow

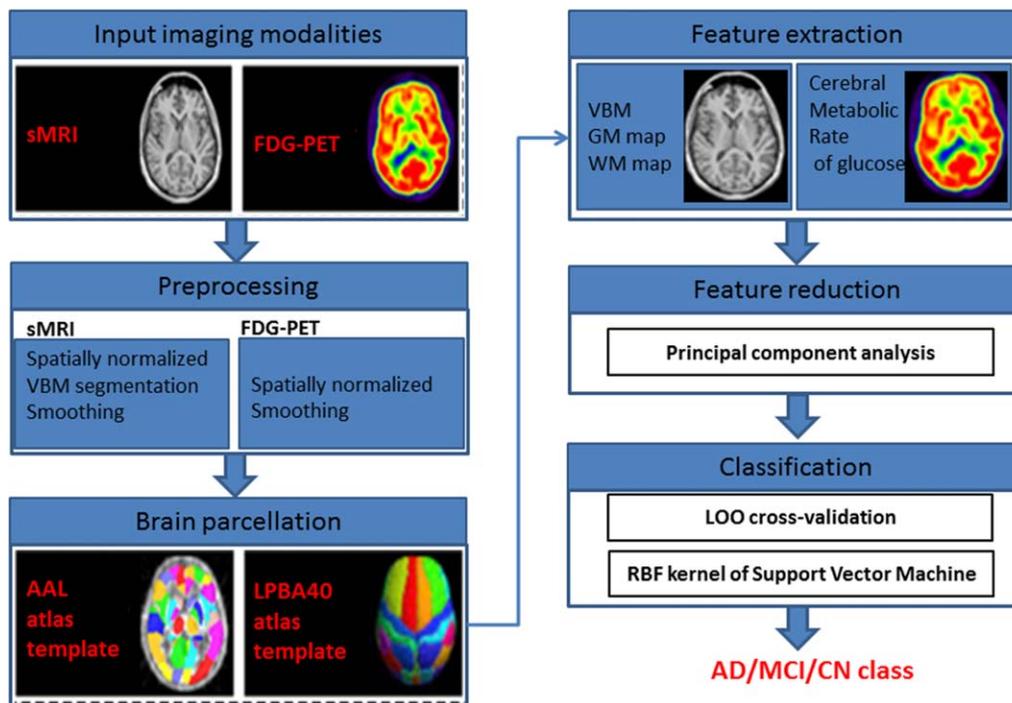


FIGURE 2 Multi-modal, multi-atlas classification method for AD and MCI detection [Color figure can be viewed at wileyonlinelibrary.com]

of these steps is described in Figure 2, and the following text describes these steps in detail.

2.1 | Image acquisition

The data used in this research work (structural MRI, FDG-PET and clinical diagnosis of AD, MCI or CN) was downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database,^{25,26} which is public database hosting more than 800 images. The baseline three-dimensional structural MRI images were acquired by using T1 weighted gradient echo sequences on 1.5 Tesla and 3.0 Tesla scanners. To acquire FDG-PET images, venous injection of 18F-FDG was given to patients and images were acquired after 40–60 minutes in a resting-state dark room.

2.2 | Preprocessing

The preprocessing comprises segmentation, spatial harmonization and smoothing of structural MRI images. Briefly, we do anterior commissure (AC)-posterior commissure (PC) correction on all the images, and use the N3 bias correction method to correct the intensity inhomogeneity. Next, we do skull-stripping on structural MR images using a brain extraction tool.²⁷ After removal of cerebellum, FAST in the FSL package is used to segment structural MR images into three different tissues: gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF).

FDG-PET images are normalized using an existing FDG-PET template, prepared from MRI and FDG-PET scans of 50

cognitively normal individuals used in this study. Smoothing of 8-mm full-width half-max is applied on the resultant images.

2.3 | Brain parcellation

To analyze the effect of atlas (template)-based region of interest (ROI) analysis on the classification performance of AD and MCI, we utilized two brain templates in this study: AAL and LPBA40 templates. Both these are well-known atlases and have been utilized in various research studies in the last decade for brain parcellation into different anatomical regions.

The AAL atlas comprises 116 ROIs, which include 26 cerebellar regions and 90 cerebral regions. The LPBA40 atlas, however, comprises 56 ROIs, which include brainstem and cerebellum regions along with 54 cerebral regions.

2.4 | Feature extraction

Using each brain atlas, we calculated two types of features that is features from structural MRI and features from FDG-PET. From the structural MRI, we calculated the average GM value from each ROI, and from FDG-PET we calculated average relative cerebral metabolic rate for glucose in each ROI.

There is a rationale behind using ROI-based features. Whole brain-based features suffer the drawback of dimensionality, as the numbers of features are typically much larger than the size of the available subjects. When the number of features is high relative to the number of subjects in the training set, it is possible that classification rules yielding high accuracy on the training set were originated only by

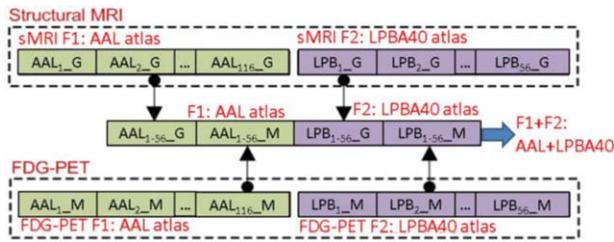


FIGURE 3 Combination of features extracted from multiple modalities and multiple atlases [Color figure can be viewed at wileyonlinelibrary.com]

chance. This can lead many algorithms to select classification rules that could fail to generalize to new data.²⁸ Consequently, features are either reduced using supervised or unsupervised feature-reduction methods, or are extracted from pre-defined atlases and adaptive regions to reduce dimensionality. We adopted the second approach and extracted features from ROIs. These features were combined in the end to get final feature set for each atlas. Another feature set was obtained by combining the MRI and FDG-PET features extracted from both the atlases. This process led us to a total of three feature sets (two feature sets from individual atlas, and one combined feature set from both the atlases). Figure 3 shows the feature concatenation process. In the figure, AAL_{i_G} and AAL_{i_M}, respectively, shows the regional GM volume and regional metabolic rate of glucose calculated via AAL atlas. Similarly, LPB_{i_G} and LPB_{i_M}, respectively, show the regional GM volume and regional metabolic rate of glucose calculated via LPBA40 atlas. The feature sets from individual atlases have also been used for the classification to prove the hypothesis that the feature set obtained from two atlases is better than the feature sets obtained from either of the individual atlases.

2.5 | Feature reduction

Feature reduction is an important component of a classification framework.²⁹ Main purpose of feature reduction is to reduce the dimensions of the input feature space and utilize a smaller optimal subset of feature space, almost having similar classification ability to discriminate the subjects belonging to different classes as the original, complete feature set. Major benefit of feature reduction is that it finds non-redundant, optimal and smaller feature set, which has reduced noise, and also leads to reduced computational cost.

In this research study, principal component analysis (PCA) was used as a feature reduction procedure. The theoretic principles of PCA were put forward by Pearson. PCA is a mathematical transformation, which transforms the correlated variables of a feature set into uncorrelated variables.³⁰ The resultant uncorrelated variables are termed as principal components of the feature set. The PCA transformation (orthogonal)

makes sure that principal component 1 (PC1) bears maximum variance of the data, and each following principal component (PC2, PC3, ...) preserves orthogonality with preceding PCs and also holds maximum possible variance of the feature set. It has been extensively used in the past for reducing dimensionality of medical data.³¹ We started from one PC and kept on adding subsequent PCs until there was no increase in the performance and the remaining PCs were simply discarded.

2.6 | Classification

SVMs³² are the most extensively used classification methods for medical images.^{33–36} SVM has also been the most popular choice of the researchers for AD and MCI prediction.¹⁶ To focus on the performance comparison of features extracted from individual atlases and the combined features of both the atlases, we used radial basis function (RBF) kernel of SVM. The optimal value of the parameters of the RBF SVM that is C (cost of constraint violation) and gamma (width) has been determined through grid search. The values of these parameters have been varied in a suitable range to determine the optimal value for these parameters.

We used leave-one-out (LOO) cross-validation, which is the most widely used case of K-fold cross-validation. In one particular iteration of LOO cross-validation, one subject was left out as a test subject, and the classifier trained on the remaining n-1 subjects was applied on the hold out test subject.

2.7 | Software's availability

This algorithm has been implemented in Matlab. Upon publication of this article, code will be made publicly available.

3 | PERFORMANCE EVALUATION MEASURES

We evaluated the classification capability of the proposed framework via following performance metrics.

3.1 | Classification accuracy

Classification accuracy is a measure of the overall classification ability of the classification framework. It can be calculated by

$$\text{Accuracy} = \frac{\text{TRP} + \text{TRN}}{\text{TRP} + \text{FLP} + \text{TRN} + \text{FLN}} \times 100,$$

where

TRP = # of correctly identified positive patients

TRN = # of correctly identified negative patients

FLP = # of incorrectly identified positive patients

FLN = # of incorrectly identified negative patients

3.2 | Sensitivity (specificity)

Sensitivity (specificity) is a measure of the percentage of positive (negative) samples, which are correctly identified by the classification framework. These measures can be formulated by the following formulas.

$$\text{Sensitivity} = \left(\frac{\text{TRP}}{\text{TRP} + \text{FLN}} \right) \times 100$$

$$\text{Specificity} = \left(\frac{\text{TRN}}{\text{TRN} + \text{FLP}} \right) \times 100$$

3.3 | Receiver operating characteristic (ROC) curve

ROC curve provides a graphical method to visualize (characterize) the classification framework over its entire operating range. In ROC curve, true positive rate is plotted on the x -axis, whereas false positive rate is plotted on the y -axis.

3.4 | F-Measure

F-Measure makes use of precision-rate and recall-rate to calculate effectiveness of classification.

$$\text{Precision-Rate} = \frac{\text{TRP}}{\text{TRP} + \text{FLP}},$$

$$\text{Recall-Rate} = \frac{\text{TRP}}{\text{TRP} + \text{FLN}}$$

F-Measure can be calculated via following equation. It ranges [0 1], where 0 is the least and 1 is the maximum possible value.

$$F\text{-measure} = 2 * \left(\frac{\text{Precision-Rate} \times \text{Recall-Rate}}{\text{Precision-Rate} + \text{Recall-Rate}} \right)$$

3.5 | Kappa statistic

Kappa is a rate of agreement between the two raters. The rater can be anything, a computerized method or a human expert. In this search study, Kappa has been used to measure the agreement between the ground truth status of a patient and the status calculated by the classification framework.

4 | EXPERIMENTAL RESULTS

4.1 | Dataset

The data used in this research work was downloaded from ADNI database^{25,26} as we have already discussed. The general eligibility criteria for ADNI subjects are given at www.adni-info.org. Briefly summarizing the data, subjects have ages ranging from 55 to 90 years, having a companion who

TABLE 1 Baseline characteristics of 300 participants

Baseline characteristics of participants			
Characteristic	AD	MCI	CN
No. of subjects	100	100	100
Male	64	58	49
Female	36	42	51
Average age (years)	71.4	70.5	69.5
Age (SD)	06.7	06.8	03.8
Average education (years)	12.3	11.8	12.5
Education (SD)	03.3	03.1	04.2

was able to provide an independent assessment of functioning abilities.

General inclusion and exclusion criterion for the subjects from ADNI are as follows:

- Cognitively normal: Mini-Mental State Examination (MMSE) scores ranging from 24 to 30, a Clinical Dementia Rating (CDR) of 0, non-demented, non-depressed and non-MCI.
- MCI patients: MMSE scores ranging from 24 to 30, a CDR of 0.5, having considerable memory loss evaluated by Wechsler Memory Scale Logical Memory II scores which were education adjusted, no other significant cognitive impairment, essentially performing well in daily routine tasks, and no evidence of dementia.
- AD patients: MMSE scores ranging from 20 to 26, a CDR of 0.5/1.0, and meets the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for AD.

In this paper, we focused only on those ADNI subjects, which have structural MRI and FDG-PET baseline data available. We randomly picked 300 subjects including 100 AD, 100 MCI and 100 CN subjects from the subjects having both the modalities available. More details about the dataset are provided in Table 1.

4.2 | Classification of AD, MCI and CN

Here, we have three classes, so we have divided the problem into the three different classification problems:

- AD versus CN
- AD versus MCI
- MCI versus CN

For each of these three classification problems, the features from each modality that is structural MRI and FDG-PET

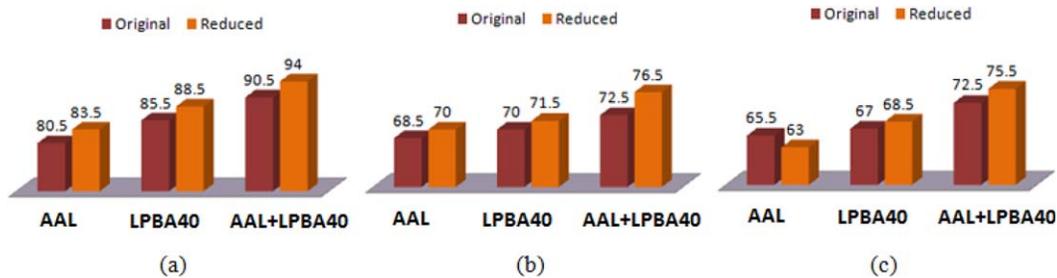


FIGURE 4 Classification accuracy of the individual (AAL, LPBA40) and hybrid (AAL + LPBA40) feature sets: (A) CN versus AD, (B) CN versus MCI, (C) MCI versus AD [Color figure can be viewed at wileyonlinelibrary.com]

were extracted by using two atlas templates that is LPBA40 and AAL. These features were combined to generate a hybrid feature set. The features were reduced using PCA as already described in the methods section, and were given as input to the classification algorithm.

We hypothesized that hybrid feature set (AAL + LPBA40) will produce better performance than individual atlas-based features (AAL, LPBA40). Therefore, to compare their performance, individual and hybrid features were given as input to classifiers. In addition, we also hypothesized that PCA-based feature reduction helps in better performance when compared with raw feature set. Therefore, to analyze the effect of PCA-based feature reduction on classification performance, the original and reduced features were used for the classification.

Figure 4 shows the overall performance of the feature sets extracted from individual and hybrid atlases for different problems. It also shows the performance of original feature set and the feature set reduced by PCA (with optimal number of principal components). We obtained 94%, 76.5% and 75.5% classification success rate for AD versus CN, CN versus MCI and MCI versus AD subjects, respectively. Overall the results are very encouraging leading to the conclusion

TABLE 2 Performance measures for individual and hybrid atlas-based features for different datasets by using optimal number of PCA components

	Accuracy	Sensitivity	Specificity
CN versus AD			
AAL	83.5	0.79	0.88
LPBA40	88.5	0.90	0.86
AAL + LPBA40	94.0	0.95	0.93
CN versus MCI			
AAL	70.0	0.70	0.70
LPBA40	71.5	0.72	0.71
AAL + LPBA40	76.5	0.78	0.75
MCI versus AD			
AAL	63.0	0.64	0.62
LPBA40	68.5	0.70	0.67
AAL + LPBA40	75.5	0.71	0.80

that classification performance improves once features extracted from different atlases are combined. Different atlases have different parcellation mechanisms and they parcellate the brain into different number of anatomical regions spanning slightly varying regions on the brain. Therefore, when the features are combined from multiple regions, they reinforce each other and classification performance increases. Another fact that we can conclude from Figure 4 is that the reduced feature set, after applying PCA, proves to be very helpful by eliminating unimportant features from the given feature set, and thus increases the classification accuracy. In all the classification experiments (ie, AD/CN, CN/MCI, MCI/AD) shown in Figure 4, the classification accuracy obtained by using optimal number of PCA components is significantly better compared to the classification accuracy obtained by using all the features.

We have also measured the performance of the features extracted from the individual and the hybrid templates in terms of sensitivity and specificity. The results are given in Table 2, where we see nearly equal sensitivity and specificity in all the cases, which shows that the extracted features are not biased toward any class and are equally capable of distinguishing the subjects of all the classes that is AD, MCI and CN.

We have also calculated the performance of the extracted features in terms of Kappa statistics and F-Measure. Corresponding results are given in Figure 5. Overall these results are also promising leading us to the same conclusion.

In all the cases, we see better classification performance of AD versus the CN subjects. This is inspiring and is consistent with the literature as well where we see better separation among CN and AD subjects compared to other subject groups. The reason of better classification performance for this subject group is that the AD subjects have much more atrophy compared to the normal aging therefore show higher variation in terms of the gray matter volume and relative cerebral metabolic rate for glucose in different parts of the brain compared to CN. MCI group, however, is very heterogeneous and shows a stage on a continuum of progression from CN to AD group. Some of the subjects in MCI group are the stable MCIs (sMCI), which do not progress to AD for a certain period of time and show atrophy similar to CN subjects. Similarly, some MCI subjects are progressive MCIs (pMCI) and they

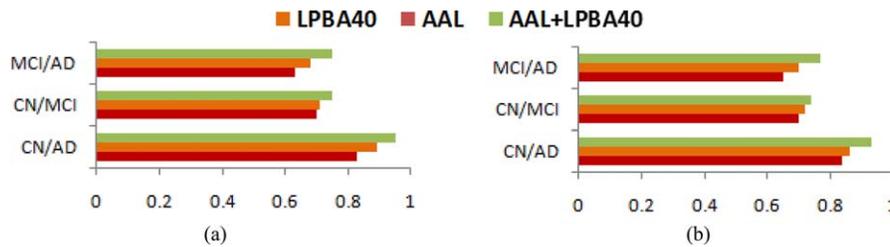


FIGURE 5 Performance measures for individual and hybrid atlas-based features for different datasets at optimal number of PCA components, (A) F-Measure, (B) Kappa statistics [Color figure can be viewed at wileyonlinelibrary.com]

transform to AD group within the length of follow-up period. Therefore, we believe that when all the MCI subjects are grouped as one entity, classification accuracy between CN and MCI group decreases due to sMCI subjects, and decreases between AD and MCI classification due to pMCI subjects.

4.3 | Performance analysis in terms of ROC curves

ROC curves are another very common way of reporting results for a binary classifier. The ROC curve plots the sensitivity against “1-specificity” by changing the discrimination threshold and therefore provides a complete picture of classifier’s performance. The ROC curve is usually summarized by the area under the curve (AUC), which is a number between 0 and 1. The ROC curves are shown in Figure 6 for different data subsets. The ROC curves here are quite promising and show that the classifiers have been properly trained in all the cases, and do not lead to any bias in any case. The ROC curves for all the cases are well above the chance, especially for the classification between AD and CN groups, which is again consistent with all the performance measures we are calculating here.

4.4 | Influence of PCA on the classification rate of different patient groups

Inspired by existing literature on the use of PCA for medical images,³⁷ we have also analyzed the performance of the

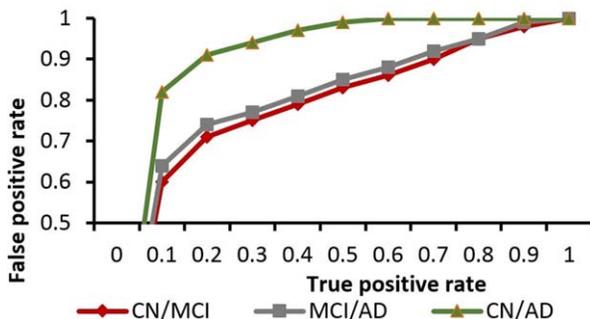


FIGURE 6 ROC curve for the hybrid atlas-based features (AAL + LPBA40) at the optimal number of PCA components for AD versus MCI, AD versus CN and CN versus MCI subjects [Color figure can be viewed at wileyonlinelibrary.com]

individual and hybrid atlas-based features in terms of the number of principal components selected by the PCA. In fact, the number of components has been varied from minimum of 2 to a maximum value of 15, and the performance of different datasets extracted from different individual and a hybrid atlas has been investigated. The results for this experiment are given in Figure 7.

These figures show that the classification performance increases with an increase in the number of principal components, and then reaches to a maximum value at generally 10–11 number of components in all the cases. The classification performance either remains the same after adding more components, or it deteriorates, thereby showing that 10–11 numbers of principal components are enough in these cases and addition of more components leads to noisy feature in the feature set. An analysis of the variance captured by various principal components is given in Figure 8.

4.5 | Performance evaluation of the proposed method using other classification models

We have also compared the performance of the RBF classifier against other kernels of SVM such as linear, polynomial and sigmoid (see Table 3). The performance using RBF kernel is indeed better than other kernels of SVM. The better performance of RBF kernel may be attributed to the better generalization capability of RBF kernel on unseen populations. In addition, we have also shown the performance of diseased versus healthy groups by combining AD and MCI in one group and classifying them against CN group.

4.6 | Important brain regions involved in AD and MCI classification

In the end, we have also analyzed the important regions pertaining to AD and MCI detection. The classification process was repeated 100 times, and the features which were mostly ranked on top by the SVM classifier were noted down. Among these features, the features extracted from hippocampus and amygdala seems to have better performance compared to features of other regions. Figure 9 lists the regions which were mostly used by the SVM classification. The x -axis in the figure shows the number of times any feature was

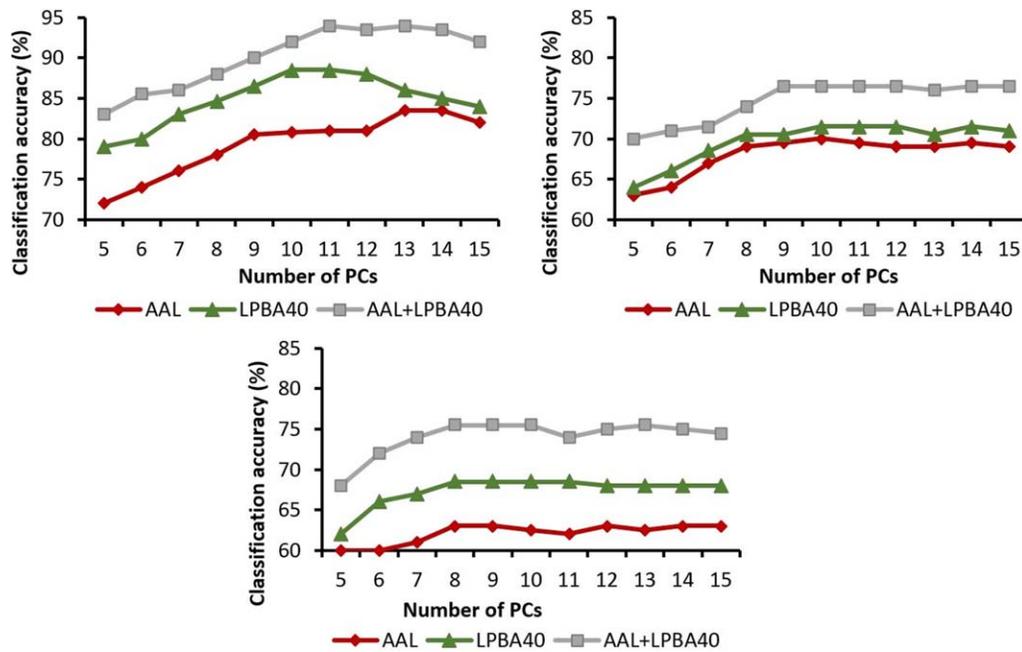


FIGURE 7 Classification performance of individual and hybrid atlas-based features from different datasets at different number of principal components selected by PCA. (A) AD versus CN, (B) CN versus MCI, (C) MCI versus AD [Color figure can be viewed at wileyonlinelibrary.com]

selected amongst the top 5 by SVM classifier, and the y-axis shows the name of the regions. Out of the 116 regions of AAL atlas, 28 regions were selected amongst the top five, and out of the 56 regions of LPBA40 atlas, 24 regions were selected. The left hippocampus was in the highest rank in AAL, which is consistent with the results of the VBM analysis. In LPBA40, the right middle frontal gyrus was ranked second following the left inferior occipital gyrus.

The results indicate that LPBA40 template yields better classification success rate compared to AAL in all the cases despite having more number of regions in AAL compared to LPBA40. The possible explanation for this phenomenon can be that AAL atlas is based on single-subject brain, whereas the LPBA40 atlas has been developed based on a population,

therefore a single subject-based atlas may not be sufficient enough to divide brain into more meaningful regions.

4.7 | Performance comparison of the proposed method with existing methods

We have compared the performance of the proposed multi-atlas, multi-modal classification framework with existing techniques of AD and MCI classification. We selected two techniques, each from voxel as features based methods, fixed atlas-based and adaptive atlas-based methods.

Amongst the selected voxel as features based methods, Klöppel et al.³ and Casanova et al.² used GM density map of the entire brain, respectively, together with SVM and large-scale regularization approach. Amongst the fixed atlas-based methods,^{4,18} Magnin et al., used AAL to parcellate the brain image into 116 regions, and then used the relative weight of the GM, compared to that of the WM and CSF, for each parcellated region to develop a feature vector for SVM-

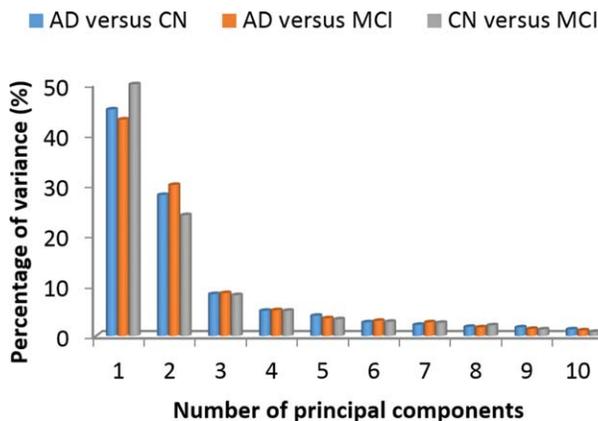


FIGURE 8 Percentage variance captured by the principal components for the three classification problems [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Performance evaluation of the proposed method (hybrid features) in terms of classification accuracy using various kernels of SVM by using optimal number of PCA components in each case

	RBF	Linear	Sigmoid	Polynomial
CN versus AD	94.0	89.5	88.5	88.5
CN versus MCI	76.5	72.5	70.5	71.0
MCI versus AD	75.5	70.0	68.5	69.0
MCI+AD versus CN	85.5	83.5	83.5	80.0

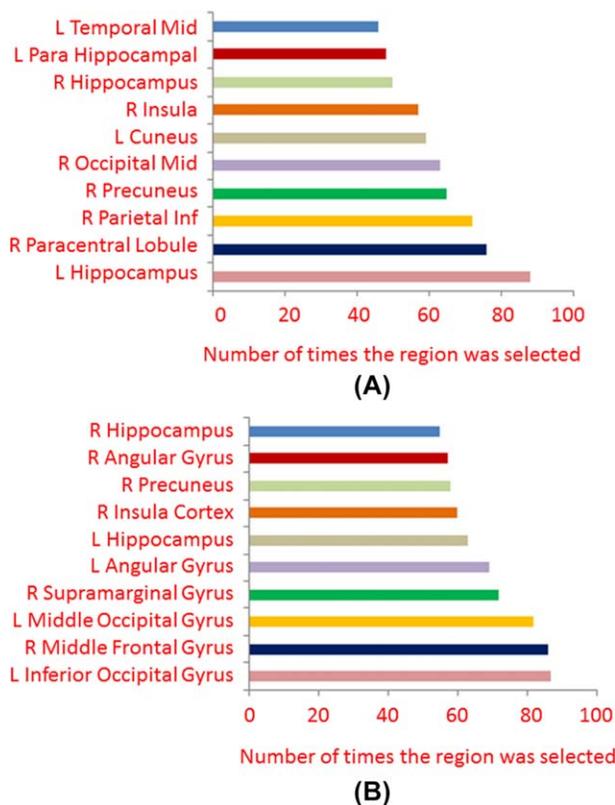


FIGURE 9 Final bar plots of the features of the individual brain atlases calculated by SVM. (A) AAL atlas, (B) LPBA40 atlas [Color figure can be viewed at wileyonlinelibrary.com]

based AD classification.⁴ Dai et al. used regional GM volumetric measures and functional measures as features.¹⁸ They trained separate maximum uncertainty LDA classifiers on the structural and functional measures, and combined the output of the classifiers via weighted voting.

In addition, two methods from adaptive atlas-based methods were also chosen for comparison. First, Min et al.

TABLE 4 Comparison of the classification capability of the proposed method with existing AD detection techniques

Classification techniques	CN versus AD	CN versus MCI	MCI versus AD
Voxels as features-based methods			
Klöppel et al. [3]	74.12	62.25	71.87
Casanova et al. [2]	73.44	59.12	71.56
Fixed atlas-based methods			
Magnin et al. [4]	88.98	57.15	81.78
Dai et al. [18]	87.45	62.56	83.12
Adaptive atlas-based methods			
Min et al. [19]	93.20	66.15	86.56
Liu et al. [20]	90.45	67.25	89.45
Proposed method (hybrid features + PCA)	94.00	76.50	75.50

derived multiple atlases from the non-overlapping clusters of subjects,¹⁹ obtained using affinity propagation.³⁸ They registered subjects to the atlases and adaptively calculated a set of ROIs and volumetric features in each atlas space. The top-most K discriminating features calculated from each atlas space were combined for SVM-based classification. Subsequently, Liu et al. argued that the features extracted from K sets of adaptive ROIs are different representations of the same subject,²⁰ and should not be concatenated, as in a previous study.¹⁹ To resolve this, Liu et al. registered subjects to different selected atlases and extracted features from adaptive regions of each atlas-registered image, viewing that image as the main source, and all other atlas registered-images as adjunctive sources.²⁰ SVM was separately trained on features extracted from each set and the results of multiple sets were combined using majority voting.

We evaluated the performance of these techniques on ADNI dataset described in Table 1. Table 4 provides a performance comparison of our method with those of existing techniques.

Our proposed approach has produced highest classification accuracy (94.00%) for AD versus CN classification, which is 4.80% higher than the highest accuracy yielded by Min et al.¹⁹ and 20.56% higher than the lowest accuracy yielded by Casanova et al.² Similarly, there is a significant increase (9.25%) in classification accuracy of CN versus MCI compared to Liu et al.²⁰ The performance of the proposed approach is boosted due to the proposed hybrid features computed from multiple atlases. It can be concluded reasonably that the better performance of the proposed method is attributed to the fact that previous schemes take one atlas-based features for classification. These features represent brain atrophy from one perspective and do not capture atrophy from multiple perspectives. Contrary, our proposed AD and MCI classification framework uses rich feature set captured from multiple atlases wherein each individual feature set, calculated from one atlas, captures potentially exclusive atrophy pattern of the disease. The results suggest that when features extracted from individual atlases are combined, they reinforce each other and produce better classification compared to the classification achieved by features of one atlas.

5 | CONCLUSION

This paper presents a new method that first quantifies the atrophy in the brain of AD and MCI subjects by parcellating brain images into several anatomical regions based on two atlas templates and extracting features of those regions. Later, it combines the quantitative features extracted from both the templates. PCA-based feature reduction has been applied to select the discerning features from the dataset.

Working with ADNI dataset, we show that the proposed hybrid atlas-based method quite effectively captures the atrophy in diseased subjects when compared with the atrophy patterns extracted from one atlas only, and hence leads to effective classification results. Furthermore, we have also separately analyzed the classification capabilities of the features extracted from each atlas template separately to determine the performance advantage achieved due to hybrid atlas-based features. We have also analyzed the performance advantage we get by applying PCA when compared with original features. The results show that multiple atlases capture the brain atrophy from different perspectives, and hence lead to rich and effective feature set that in turn leads to better classification accuracy. This work can be extended in numerous possible ways. For instance, ensemble of various different classifiers can be developed for classification. Similarly, features from more than two atlases can be combined to get better and more effective feature set.

ORCID

Mohammad Aksam Iftikhar  <http://orcid.org/0000-0002-5324-2041>

REFERENCES

- [1] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. *Alzheimer's Dement.* 2007;3:186–191.
- [2] Casanova R, Wagner B, Whitlow CT, et al. High dimensional classification of structural MRI Alzheimer's disease data based on large scale regularization. *Front Neuroinform.* 2011;5:22.
- [3] Klöppel S, Stonnington CM, Chu C, et al. Automatic classification of MR scans in Alzheimer's disease. *Brain.* 2008;131:681–689.
- [4] Magnin B, Mesrob L, Kinkingnéhun S, et al. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology.* 2009;51:73–83.
- [5] Cuingnet R, Gerardin E, Tessieras J, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *NeuroImage.* 2011;56:766–781.
- [6] Chen G, Ward BD, Xie C, et al. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology.* 2011;259:213–221.
- [7] Koch W, Teipel S, Mueller S, et al. Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease. *Neurobiol Aging.* 2012;33:466–478.
- [8] Gray KR, Wolz R, Heckemann RA, et al. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. *NeuroImage.* 2012;60:221–229.
- [9] Gray KR, Wolz R, Keihaninejad S, et al. Regional analysis of FDG-PET for use in the classification of Alzheimer's disease. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro;* Chicago, IL, USA, 2011:1082–1085. <https://doi.org/10.1109/ISBI.2011.5872589>.
- [10] Padilla P, Lopez M, Gorriz JM, et al. NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease. *IEEE Trans Med Imaging.* 2012;31:207–216.
- [11] Ramirez J, Chaves R, Gorriz J, et al. Computer aided diagnosis of the Alzheimer's disease combining SPECT-based feature selection and random forest classifiers. Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE; 2009:2738–2742.
- [12] Haller S, Missonnier P, Herrmann F, et al. Individual classification of mild cognitive impairment subtypes by support vector machine analysis of white matter DTI. *Am J Neuroradiol.* 2013;34:283–291.
- [13] Haller S, Nguyen D, Rodriguez C, et al. Individual prediction of cognitive decline in mild cognitive impairment using support vector machine-based analysis of diffusion tensor imaging data. *J Alzheimer's Dis.* 2009;22:315–327.
- [14] Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage.* 2017;155:530–548.
- [15] Xu L, Wu X, Chen K, Yao L. Multi-modality sparse representation-based classification for Alzheimer's disease and mild cognitive impairment. *Comput Methods Programs Biomed.* 2015;122:182–190.
- [16] Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage.* 2011;55:856–867.
- [17] Ashburner J, Friston KJ. Voxel-based morphometry—the methods. *NeuroImage.* 2000;11:805–821.
- [18] Dai Z, Yan C, Wang Z, et al. Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). *NeuroImage.* 2012;59:2187–2195.
- [19] Min R, Wu G, Cheng J, Wang Q, Shen D. Multi-atlas based representations for Alzheimer's disease diagnosis. *Hum Brain Mapp.* 2014;35:5052–5070.
- [20] Liu M, Zhang D, Shen D. View-centralized multi-atlas classification for Alzheimer's disease diagnosis. *Hum Brain Mapp.* 2015;36:1847–1865.
- [21] Ota K, Oishi N, Ito K, Fukuyama H. A comparison of three brain atlases for MCI prediction. *J Neurosci Methods.* 2014;221:139–150.
- [22] Ota K, Oishi N, Ito K, Fukuyama H. Effects of imaging modalities, brain atlases and feature selection on prediction of Alzheimer's disease. *J Neurosci Methods.* 2015;256:168–183.
- [23] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage.* 2002;15:273–289.
- [24] Shattuck DW, Mirza M, Adisetiyo V, et al. Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage.* 2008;39:1064–1080.
- [25] Weiner MW, Veitch DP, Aisen PS, et al. 2014 Update of the Alzheimer's disease neuroimaging initiative: a review of papers

- published since its inception. *Alzheimer's Dement.* 2015;11:e1–e120.
- [26] Weiner MW, Veitch DP, Aisen PS, et al. The Alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimer's Dement.* 2013;9:e111–e194.
- [27] Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp.* 2002;17:143–155.
- [28] Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw.* 1999;10:988–999.
- [29] Rathore S, Iftikhar MA, Hussain M. A novel approach for automatic gene selection and classification of gene based colon cancer datasets. Islamabad, Pakistan, 42–47. <https://doi.org/10.1109/ICET.2014.7021014>.
- [30] Pearson K. On lines and planes of closest fit to systems of points in space. *Philos Mag.* 1901;2:559–572.
- [31] Rathore S, Hussain M, Khan A. GECC: gene expression based ensemble classification of colon samples. *IEEE/ACM Trans Comput Biol Bioinform.* 2014;11:1131–1145.
- [32] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol.* 2011;2:27.
- [33] Rathore S, Hussain M, Iftikhar MA, Jalil A. Ensemble classification of colon biopsy images based on information rich hybrid features. *Comput Biol Med.* 2014;47:76–92.
- [34] Rathore S, Hussain M, Khan A. Automated colon cancer detection using hybrid of novel geometric features and some traditional features. *Comput Biol Med.* 2015;65:279–296.
- [35] Tahir M. Pattern analysis of protein images from fluorescence microscopy using gray level co-occurrence matrix. *J King Saud Univ Sci.* 2016. <https://doi.org/10.1016/j.jksus.2016.12.004>.
- [36] Tahir M, Khan A. Protein subcellular localization of fluorescence microscopy images: employing new statistical and texton based image features and SVM based ensemble classification. *J Inform Sci.* 2016;345:65–80.
- [37] Tahir M, Khan A, Kaya H. Protein subcellular localization in human and hamster cell lines: employing local ternary patterns of fluorescence microscopy images. *J Theor Biol.* 2014;340:85–95.
- [38] Frey BJ, Dueck D. Clustering by passing messages between data points. *Science.* 2007;315:972–976.

How to cite this article: Asim Y, Raza B, Malik AK, Rathore S, Hussain L, Iftikhar MA. A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning. *Int J Imaging Syst Technol.* 2018;00:1–11. <https://doi.org/10.1002/ima.22263>