Informatics in Medicine Unlocked 14 (2019) 34–42

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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

Multistage classifier-based approach for Alzheimer's disease prediction and retrieval

K.R. Kruthika (Research Scholar)*, Rajeswari (Professor), H.D. Maheshappa (Professor), Alzheimer's Disease Neuroimaging Initiative[1]

Department of Electronics and Communication Engineering, Acharya Institute of Technology, Bangalore, India

ARTICLE INFO

Keywords:
Alzheimer's disease
Machine learning
Content-based image retrieval
Multistage classifier
PSO
Structural MRI
SVM
K-NN

ABSTRACT

The most prevalent and common type of dementia is Alzheimer's disease (AD). However, it is notable that very few people who are suffering from AD are diagnosed correctly and in a timely manner. The definite cause and cure of the disease are still unavailable. The symptoms might be more manageable and its treatment can be more effective, when the impairment is still at an earlier stage or at MCI (mild cognitive impairment). AD can be clinically diagnosed by physical and neurological examination, so there is an need for developing better and efficient diagnostic tools for AD. In recent years, content-based image retrieval (CBIR) systems have been widely researched and applied in many medical applications. Combining an automated image classification system and the radiologist's professional knowledge, to increase the accuracy of prediction and diagnosis, were the main motives. In this paper, a multistage classifier using machine learning, including Naive Bayes classifier, support vector machine (SVM), and K-nearest neighbor (KNN), was used to classify Alzheimer's disease more acceptably and efficiently. For this, MRI (Magnetic resonance imaging) scans were processed by FreeSurfer, a powerful software tool suitable for processing and normalizing brain MRI images. We also applied a feature selection technique - PSO (particle swarm optimization) to many feature vectors in order to obtain the best features that represent the salient characteristics of AD. The results of the proposed method outperform individual techniques in a benchmark database provided by the Alzheimer's Disease Neuroimaging Institute (ADNI).

1. Introduction

Alzheimer's Disease is a neurodegenerative brain disorder that is the most prevalent and common type of dementia. Although, other diseases and conditions of elder adults can also be the cause of dementia. AD is not currently a curable disease but the progress of the disease can be slowed if it is detected at an early or mild stage. AD causes pathological changes in the brain and these changes can be detected before clinical symptoms begin. From a bio-marker viewpoint, the chief pathologies of AD are: progressive buildup of the beta-amyloid protein fragments (plaques) outside the neurons, and the presence of twisted strands of the tau protein (tangles) inside neurons, in the brain. These changes eventually cause damage to, and death of, neurons.

All researchers of this domain focus on monitoring a patient's health change, clinical progression of disease, and reaction to the therapy. It is most challenging for them to find relevant bio-markers that represent AD and MCI well. Their goal is not only to diagnose this disease at an early stage, but also identify which people are most likely to develop AD. Magnetic Resonance Imaging (MRI) is a non-invasive medical tool that physicians use to diagnose patient disease or health conditions. The MRI techniques generally include a powerful magnetic field, radio-frequency pulses, and a computer to produce detailed pictures of all internal body structures [1]. Individual or combined structural MRI biomarkers such as shape and texture of the hippocampus, cortical measurements, and volume measurements are used as biomarkers for the multiclass classification of AD, mild cognitive impairment (MCI), and normal control (NC) [2,3].

Recently, several machine learning techniques such as SVM (support vector machines), KNN (K-nearest neighbor) algorithm, NN (Neural Network), Ensemble and regression models have been employed to classify AD, MCI and NC [1,7]. Sometimes, the classification by one specific classifier does not provide the desired result due to

* Corresponding author.
E-mail address: kr.kruthika@gmail.com (K.R. Kruthika).

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

https://doi.org/10.1016/j.imu.2018.12.003
Received 23 August 2018; Received in revised form 16 November 2018; Accepted 11 December 2018
Available online 13 December 2018

certain factors, so there were attempts to use an ensemble of classifiers and a multistage classifier for more correct classification results. The ensemble approach uses a voting technique principally, and the multistage classifier uses output to one stage classifier as input to the next stage classifier [8]. In this way, the advantages of individual classifiers can be combined and disadvantages can be overcome.

As mentioned above, it remains a major challenge to select best features that represent characteristics useful to discern between AD, MCI and NC. In recent years, a content-based image retrieval system and medical image classification were vigorously interactive to each other and were used for detecting AD or MCI [3].

CBIR systems that enable users to compare contents of a query image to a database directly are useful to obtain precise classification results by combining automated medical image classification techniques and the radiologists' professional knowledge and experiences. Similarly, classification is used as a reference for retrieval, to increase image retrieval performance and timeliness, considering large databases.

In general, the process of image retrieval involves two principal steps: the 1st step generates features that represent a given query image and the second step compares these generated features to those already stored in the database. The main issue is to find a satisfactory representation of the image content by using several image processing and operation methods. Moreover, it is important to provide adequate distributional distances between the features.

In this study, we will build a more effective classification system for Alzheimer's disease at an earlier stage by using the swarm intelligence feature selection technique and a multistage classifier. The AD MRI scans used in this study were obtained from the Alzheimer's Disease Neuroimaging Institute (ADNI) and processed via an MRI analyzing software – FreeSurfer. Thus, brain structural and volumetric measurements, and cortical thickness measurements, were used as feature vectors for AD detection. We used the PSO algorithm to select best biomarkers that represent AD or MCI and built a multistage classifier based on the KNN and SVM classifiers.

2. Related work and literature review

We sought papers on data mining/machine learning in the healthcare domain, addressing diagnosis of AD and MCI. Many research articles for detecting AD in its early stages were studied and the main issues involved building more effective bio-markers of brain MRI scans for AD detection. For selecting features relevant to AD, application of several machine learning classifiers and the usage of image retrieval systems have been a current field of interest.

Structural MRI is a powerful medical imaging technology that is used in AD detection, and extraction of effective structural MRI biomarkers of AD has become an active research area in this field, with several biomarkers having been proposed, analyzed, and researched [9–11]. In Ref. [12], there were attempts to use hippocampus volume as a structural MRI biomarker of AD [13]. Volume is generally used as a biomarker in several disease diagnoses, not only from the hippocampus but also many regions of interests (ROI) such as the amygdala [14], the ventricles [15], and whole brain [16], which have been also investigated. Another general type of biomarker for AD detection is morphological measurement, including cortical thickness measurement [17,18], shape [4,5], texture [19,20], and proximity of brain structures [6].

Some MRI bio-markers have different information for diagnosis and they complement each other. For example, biomarkers such as hippocampal volume provide diagnostic information independent of hippocampal shape and texture [4,20].

Moreover, the biomarkers of the different regions of the brain would be sensitive and effective to different stages of this disease - AD, MCI and NC. It was found previously that the hippocampus represents the early stage of AD well and the cortex represents later stage well [21]. Hippocampal volume measurements seem better to separate mild cognitive (MCI) and normal controls (NC) – from early stages of Alzheimer's [22]. Cortical thickness measurements were shown to separate MCI from AD [23]. Some investigators believed that a combination of complementary biomarkers may have more efficient and better information for diagnostics of AD, MCI and NC, and proposed a combination of volume and cortical thinning biomarkers in their studies [10].

Considering different MRI biomarkers for AD detection, the standardized benchmark dataset would be requisite for comparisons of approaches, and to understand the performance of different biomarkers and their relationships.

The use of CAD in dementia analysis has been widely investigated [12]. The medical image retrieval for AD based upon structural MRI measures have been studied [3]. This study focused chiefly on improvement of the performance of image retrieval by using the smallest number of features. The feature vector consisted of the volume and thickness measurements of the brain, and then subsets of the feature vector were selected by using the Correlation-based Feature Selection (CFS) method to remove any irrelevant, possibly noisy, and redundant data. The volumes and thickness measurements of the brain structure obtained from Open Access Series of Imaging Studies (OASIS) were analyzed in other research [1,2]. Modern machine learning techniques such as SVM, KNN, and Back-Propagation Neural Network (BP-NN) were used to separate AD and MCI from NC.

The work in Ref. [24] was modeled based on detail coefficients of 2-level DWT, by generalizing the auto-regressive conditional heteroscedasticity (GARCH) statistical model and the parameter limits of the GARCH model, considered as the primary feature vector. The KNN and SVM models were used to derive the results. In another study [25] a detection system was designed for breast cancer using Naive Bayes and KNN. In their work, the authors reported that their suggested method provided 96% accuracy for breast cancer categorization.

3. Proposed work

The proposed method is introduced with a purpose that is twofold: (i) to find a model technique which efficaciously distinguishes AD from MCI and NC; and (ii) to develop a multistage CAD system and study its performance. In this work, we aimed to build an effective classification system for AD by using the particle swarm intelligence feature selection technique and a multistage classifier. The AD MRI scans used in this study were obtained from the Alzheimer's Disease Neuroimaging Institute (ADNI) and processed via a processing and analyzing software - FreeSurfer; thus brain structural volumetric measurements and cortical thickness measurements were developed and used as feature vectors for AD detection. We selected the best biomarkers that represent AD or MCI by using swarm intelligent algorithm - PSO and based on them, built a multistage classifier of Naive Bayes, KNN and SVM for AD detection and image retrieval.

3.1. Data and methodology

3.1.1. Dataset

Data used in the preparation for this study were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The ADNI was collectively launched by six nonprofit organizations in 2003: The National Institute139 on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and is available at adni.loni.ucla.edu. ADNI has 3 phases: ADNI 1, ADNI GO, ADNI2 that vary in their goals and cognitive stages. The stages provided in the dataset are

Table 1
Sample size of each cognition stage for train and test.

Class	Number of Samples	Train/Test
Alzheimer's Disease (AD)	178	70:30
Mild Cognitive Impairment (MCI)	160	70:30
Normal Cognition (NC)	137	70:30

the normal control (NC), significant memory concern (SMC), early mild cognitive impairment (EMCI), mild cognitive impairment (MCI), late mild cognitive impairment (LMCI), and Alzheimer's disease (AD). In this research, MCI, NC, and AD data sets from ADNI were studied, with their quantities highlighted in Table 1.

3.2. MRI pre-processing

Datasets may contain unusual data combinations, missing values, and redundant information which can lead to misleading results. Therefore, the quality of the data must be improved prior to running an analysis. After obtaining the dataset from ADNI, it was processed further using FreeSurfer. The preprocessing steps involved head motion correction, compensation for slice-dependent time shift, smoothing, and normalizing. Finally, we obtained a raw 66 volumetric and 72 thickness measurements by FreeSurfer. Brain volume measurements, including cerebral cortex and white-matter, 3rd and 4th ventricle, inferior lateral ventricle, lateral ventricles, cerebellum cortex, caudate, putamen, pallidum, cerebellum white-matter, hippocampus and the amygdala were normalized by head size - intracranial volume. Fig. 1 represents MRI preprocessing and normalization of images in axial, sagittal and coronal planes. Obtained right and left thickness features are as follows:

Superior temporal, middle temporal, inferior temporal, entorhinal, temporal pole, lateral orbitofrontal, para-hippocampal, medial orbitofrontal, pars orbitalis, superior frontal, rostral middle frontal, inferior parietal, supramarginal, caudal middle frontal, postcentral, precuneus, pars opercularis, pars triangularis, precentral, paracentral, frontal pole, superior parietal, transverse temporal, posterior, anterior, isthmus and caudal cingulate, etc.

3.3. Feature selection

We have three high dimensional feature sets - the first 66 volumetric measurements, second 72 thickness measurements, and the total of them; these may include noisy (not very useful data in AD or MCI detection) and redundant information. Therefore, it was required to remove noisy or redundant information and represent the dataset by a smaller number of effective subset features to obtain better performance and cut the calculation time cost.

Many research and clinical experiences were studied using statistical analysis as mentioned in Ref. [20] for developing results in order to select a feature rich data representation for AD or MCI. We analyzed the effect of all of the biomarker features for AD and MCI detection and integrated certain regions of interest. Then we selected and used only a small numbers of features for AD and MCI detection. The optimum number of features were extracted using the Particle Swarm Optimization - PSO algorithm.

3.4. PSO algorithm

The Particle Swarm Optimization algorithm developed by Kennedy and Eberhart in 1995 [26] is an evolutionary computational technique.

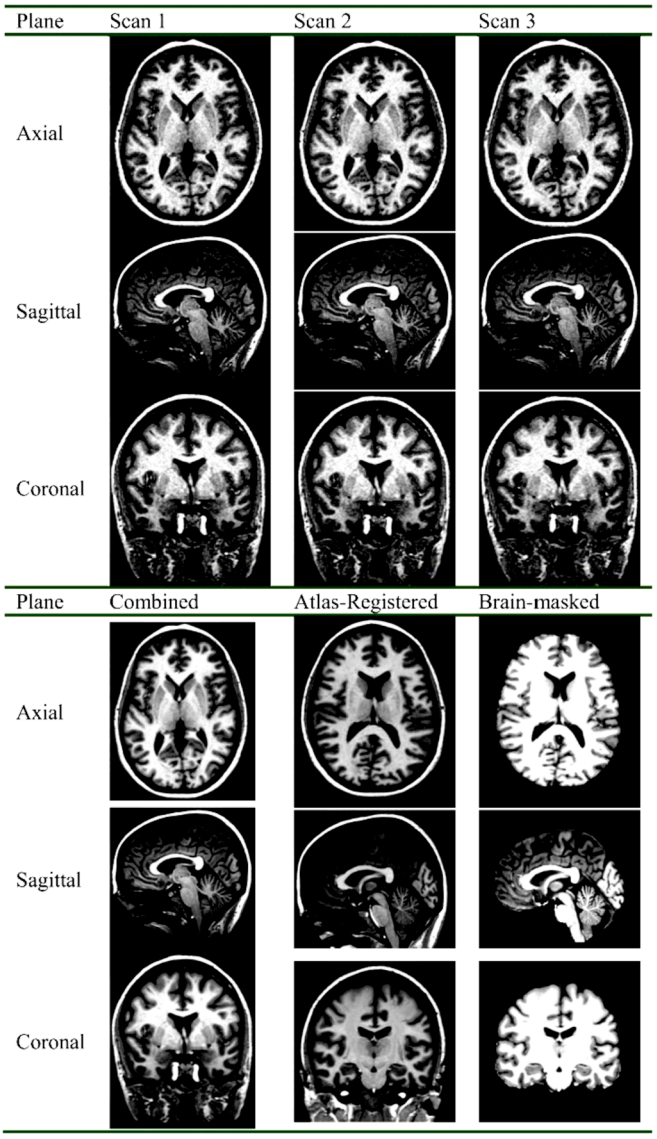

Fig. 1. MRI preprocessing and normalization, axial, sagittal, coronal planes.

PSO has shown satisfactory performance in many recent optimization scheme projects. The algorithm was developed according to the co-ordinate movement dynamics of a group of animals such as flocking of birds. This method optimizes a problem by trying to improve a candidate solution (the current location) regarding a given measurement of quality (the fitness function). PSO performs searches using a population (so-called swarm) of agents (or particles) [27]. Each particle i has a current position $loc_i = (loc_{i,1}, loc_{i,2}, \cdots, loc_{i,d})^t$ and a current flying velocity$vel_i = (vel_{i,1}, vel_{i,2}, \cdots, vel_{i,d})^t$, where d is the problem dimension. To discover the optimal solution, each particle moves in the direction of its previous best position (p_best) and its best global position (g_best), as in the following equations:

$$vel_{i,j}^{t+1} = w \cdot vel_{i,j}^{t+1} + c_1 \cdot r_1 (p\_best_{i,j} - loc_{i,j}^{t}) + c_2 \cdot r_2 (g\_best_{i,j} - loc_{i,j}^{t})$$

$$loc_{i,j}^{t+1} = loc_{i,j}^{t} + vel_{i,j}^{t+1}$$

Algorithm 1: Standard PSO

for *i* from 1 to SN **do**

$P_{vel} \leftarrow$ *random Vel*

$P_{loc} \leftarrow$ *random Loc*

$P_{p\_best} \leftarrow P_{loc}$

$P_{g\_best} \leftarrow$ *Compare (Fit* (P_{p_best}), *Fit* (P_{g_best}))

end for

while $t <$ *max_iteration* **do**

for *i* from 1 to SN **do**

$P_{vel}^{t+1} \leftarrow$ *Update Vel* (P_{vel}^{t}, P_{p_best}, P_{g_best})

$P_{loc}^{t+1} \leftarrow$ *Update Vel* (P_{loc}^{t}, P_{vel}^{t+1})

$P_{p\_best} \leftarrow$ *Compare (Fit* (P_{p_best}), *Fit* (P_{loc}^{t+1}))

$P_{g\_best} \leftarrow$ *Compare (Fit* (P_{p_best}), *Fit* (P_{g_best}))

end for

$t \leftarrow t+1$

$val \leftarrow$ *Fit* (P_{g_best})

end while

In the above formula, w is the weight inertia constant that is used to balance the goal exploration and local exploration, c_1, c_2 are personal and social learning factors, and r_1 and r_2 are random numbers between 0 and 1. The PSO algorithm showed satisfactory performance in the optimization of linear and non-linear complicated systems with different complex constraints, in many applications.

4. Multistage classifier

There were attempts to use several classifiers such as single model, including KNN, SVM, MLP, and ensemble model, including Voting, Bagged Decision Tree and Gradient Boosting in many previous research articles.

4.1. Gaussian Naive Bayes classifier

Bayesian learning is a successful method to learn the structure of data in different applications. Bayesian methods provide several structural learning algorithms. They provide models of causal influence, and allow us to explore causal relationships, perform explanatory analysis, and make predictions. Finally, Bayesian networks provide a way to visualize results. Moreover, in a binary classification problem, we have two classes, and each class has an associated feature set.

The Gaussian Naive Bayes (GNB) classifier is a simple probabilistic classifier based on applying the Bayes theorem. GNB considers each feature variable as an independent variable with each class (feature or observation) assumed to be distributed according to a Gaussian distribution. This classifier can be trained very efficiently in supervised learning, and can be used in complex real-world situations. The main advantage of GNB is that it requires a small amount of training data which is necessary for classification. For the respective feature set for each class, a mean and variance is calculated where a normal distribution is then parameterized:

$$p(x = v|c) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{-\frac{(v-\mu_c)^2}{2\sigma_c^2}} \quad (1)$$

where c is a class, v is an observation, σ_c^2 is the variance for the selected class, and μ_c is an associated mean of the class.

In the training stage, using class probabilities and conditional probabilities, the class label of the testing data point is estimated. For a two class data set, the data point is classified with respect to which class probability is higher.

4.2. K- Nearest Neighbor (KNN)

The KNN classifier is well-known and the simplest machine learning classifier. For training, a labeled database is given, and then an unlabeled data point is classified based on the label of K data points nearest the unlabeled points of the neighborhood. Here, K is the key parameter for the algorithm. In this method, the distance between testing data point x and training data points x_i, $I = 1, \ldots,n$ are calculated.

$$d_E(x, x_i) = \sum_{i=1}^{n} |x - x_i| \quad (2)$$

$$x: d_E(x, x_i) < d_E(x, x_i),\ i \neq j \quad (3)$$

The nearest k points are determined. Testing data points are classified with respect to specified k nearest neighbors.

4.3. Support vector machine (SVM)

SVM is a supervised learning model, and one of the most well-known classification algorithms; its usage has been beneficial in a large number of applications including the prediction of disease from structural MRI scans. Fig. 2 represents the SVM classification graph. SVM

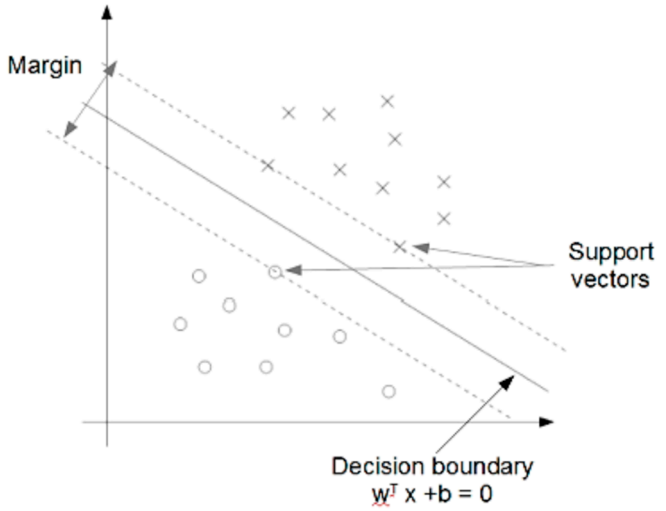

Fig. 2. Support Vector Machine classification.

classifies the data by a hyperplane that is defined as $w^T \cdot x + b = 0$ of a very high-dimensional feature space, where b is the bias for the input vector x, and w is the weight vector. A good separation is said to be the nearest training data point belonging to any class. The more the distance of the margin, the lower will be the generalization error of the classifier. Assuming N training data samples $\{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$are given, $x_i \in R^d$ is a set of feature vectors and $y_i \in \{-1,1\}$ is the class label. Then the classification problem can be expressed by the following minimization problem:

$$\min_{w,b}\left(\frac{w^T w}{2}\right) \tag{4}$$

$$s.\,t.: \quad y_i(w^T \cdot x_i) + b > 1, \quad i = 1, \cdots n \tag{5}$$

A new data object x can be classified with the following function$g(x) = \mathrm{sgn}(w^T x + b)$.

4.3.1. 2 stage classifiers

Recently, research on multistage classifiers were widely studied in order to increase performance of clustering and classification or prediction, because sometimes the direct approach to classification doesn't provide desired results. Several types of multistage classifiers were proposed. The front stage classifier can be built, a classifier which uses small dimension features and of which calculation cost (time) is low, and a back stage classifier can use more features than the previous stages, and its performance is high even though its calculation cost (time) is also high. A multistage classifier for detecting AD, MCI and NC classes is shown in Fig. 3.

If the front stage classifier classifies an input object belonging to any class at a reasonable level, the next stage classifier is not needed and the classification procedure terminates. But if the classification result of the front stage is inadequate, further input stages are needed to obtain more precise classification results.

Here we used the 2-stage classifier:

a) 1st stage classifier - Gaussian Naive Bayes Classifier, classifies the probability of whether the object belongs to AD/MCI/NC class or uncertain (reject).
b) 2nd stage classifier - SVM classifier and KNN classifier, classify object on the performance basis of 1st stage classifier.

If the first stage classifies the object as an AD class with a high-level of confidence, the procedure stops generating output results. If the first stage classifies object as MCI and NC class with high-levels, we input this object to a 2nd classifier (binary SVM classifier) and then analyze the classification results generated by this 2nd classifier.

If the result of the first stage is uncertain, this object is transferred to a 2nd classifier (KNN classifier in this case). Thus, Gaussian Naive Bayes Classifier is trained as a binary classifier (for detection of AD, MCI and NC) and the SVM classifier is also trained as a binary classifier (for classification of MCI or NC) whereas, while the KNN classifier is trained as a multiclass classifier (for classification of all AD, MCI and NC).

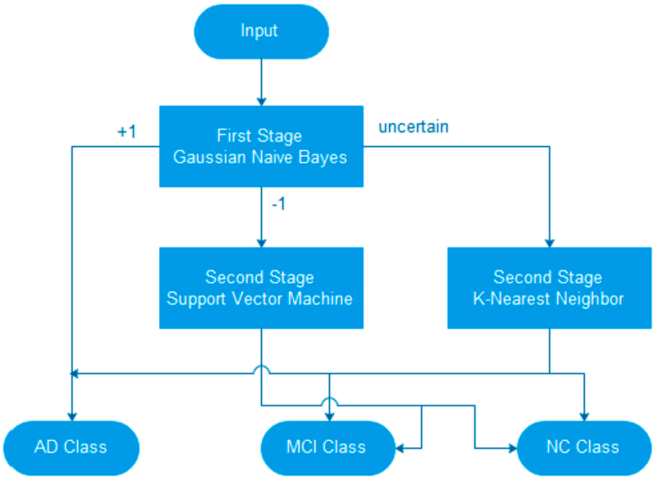

Fig. 3. Multistage classifier for detecting AD, MCI and NC classes.

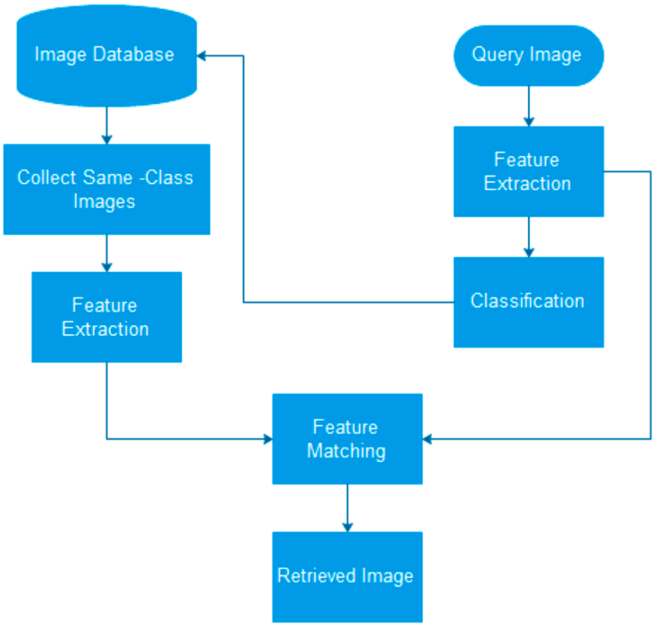

Fig. 4. Image retrieval system with classification.

4.4. Image retrieval

Content-based image retrieval - CBIR systems aim to retrieve similar images to query an image from a database by comparing the similarities of image features extracted from the query image against another in the database instead of using meta-tags or index. In medical prediction, CBIR system plays a key role to combine automated image processing and classification techniques with the radiologist's professional knowledge and previous experience, to obtain more accurate clinical predictions. In general, CBIR systems consist of two parts, such as extraction of features from image and measurement similarities, though it

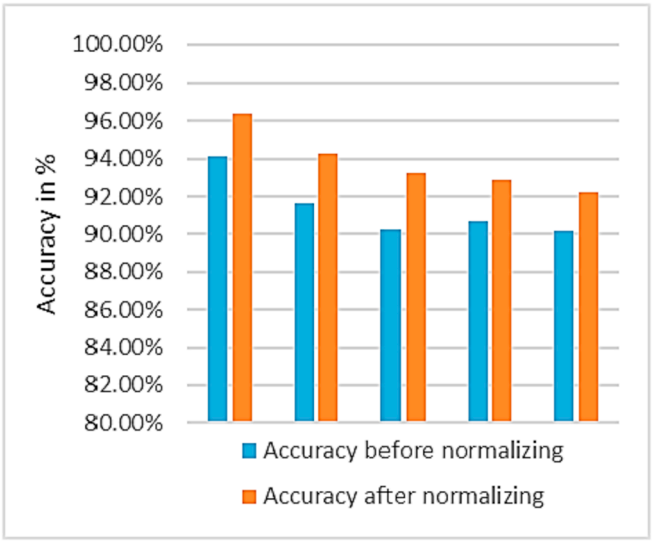

Fig. 5. Accuracy before and after normalization.

may require an inordinate amount of time if the database's size is large. Consistently, the CBIR systems with feedback as classification results were proposed for various MRI based disease detection problems. Fig. 4 shows the Image retrieval system.

In this system, the query image is first classified by trained classifiers, and only then, compared with images belonging to the same class with a query image in the database, that is, to speed up retrieval time and increase performance.

5. Experiments and results

The current paper used a system comprising a core i3 processor with 2.3 GHz speed, 8 GB RAM and the Windows operating system. The data was obtained from ADNI, which was normalized prior to gain selective features. Normalization improves the accuracy, which is shown in Fig. 5.

We used 124 AD objects, 112 MCI objects and 96 NC objects for training model, and 54 AD objects, 48 MCI objects and 41 NC objects for evaluating the performance of prediction by model. The names of the volume and thickness features selected by PSO are shown in Appendix 1.

In order to evaluate the efficiency of feature selection by the PSO algorithm, we applied our proposed approach - 2 stage classifier to three original feature sets (volume, thickness, and the total of them) and then applied it to three subsets of selected features. 5-fold cross validation was applied for all classification tasks to assess the average and general performance. The results are represented in the following Table 2.

As seen in Table 2, thickness features represent AD or MCI states better than volume features; also, we can see the selected features by PSO improved the performance of AD or MCI detection.

Table 3 shows all of the details of classification performances of our approach, using subsets of selected volume and thickness features for 5-fold cross validation.

Table 3 also shows that classification with thickness features provide best performance and our approach for AD detection is reasonably improved. But as seen in Table 4, the fault alarm rates for MCI and NC stages are larger than for AD stages. This means that extracted cortical thickness and volume features, and selected features by PSO, are more suitable for AD detection than MCI and NC detection. We examined the performance of the image retrieval scheme on the current database: a reasonable speed up and improved retrieval accuracy were shown. The total time taken for retrieval of 475 query images, and retrieval accuracy on several experimental cases, are represented in the following Table 5. Table 6 shows a performance comparison of the different ensemble methods used in the study.

We derive the accuracy score by comparing the true class label, y^*, to the predicted class label, $\hat{y}^*$, using the symmetric threshold of η =

Table 4
Confusion Matrix (a,b,c), rows are predicted and columns are true class.

(a) Thickness

	NC	AD	MCI
NC	36	2	4
AD	6	45	3
MCI	9	2	37

(b) Volume

	NC	AD	MCI
NC	26	8	8
AD	11	38	5
MCI	6	10	32

(c) Total

	NC	AD	MCI
NC	36	3	3
AD	5	45	4
MCI	7	3	38

Table 2
The accuracy of classification using original and selected features.

Thickness features		Volume features		All features	
Original (66)	Selected (50)	Original (72)	Selected (38)	Original (138)	Selected (88)
0.627	0.802	0.544	0.713	0.723	0.813

Table 3
Classification performances of 2-stage classifier.

Features	Accuracy			Sensitivity			Specificity			Precision		
	AD	MCI	NC	AD	MCI	NC	AD	MCI	NC	AD	MCI	NC
Thickness	0.880	0.869	0.855	0.788	0.787	0.838	0.935	0.910	0.862	0.882	0.815	0.721
Volume	0.811	0.818	0.798	0.755	0.708	0.666	0.844	0.872	0.852	0.746	0.735	0.651
Total	0.886	0.884	0.856	0.788	0.800	0.861	0.944	0.927	0.854	0.894	0.847	0.711

Table 5
Evaluation of the retrieval performance on several feature sets (total 475 query images).
i) Accuracy:

	Features Scenario					
	All thickness	All volume	All combination	Selected thickness	Selected volume	Selected combination
Time (Sec)	0.910	0.905	1.182	0.841	0.804	1.104
Accuracy	0.787	0.675	0.787	0.816	0.728	0.823

Table 6
Comparison of different methods.

Algorithm	Accuracy	Sensitivity	Specificity	Precision
SVM + KNN	89.22 ± 1.89	82.42 ± 1.26	78.63 ± 1.46	72.31 ± 1.12
BN + SVM + KNN	90.47 ± 1.24	88.62 ± 1.62	90.15 ± 1.84	80.26 ± 2.12
BN + SVM + KNN + PSO	96.31 ± 1.22	91.27 ± 1.44	89.90 ± 1.14	96.05 ± 1.21

Precision Rate: P = TP/(TP + FP).

0.5. If the true and predicted class labels match, we assign a score of +1, otherwise a score of 0. The final score is calculated by averaging the accuracy score over the predictions made. Thus, the total accuracy is-

$$AC = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

where TP, FP, TN and FN denote the number of true positive, false positive, true negative and false negative predictions.

ii) Sensitivity and specificity:

The accuracy is a weighted average of the sensitivity and specificity scores of the classifier. The sensitivity score is defined as

$$SS = \frac{TP}{TP + FN} \quad (7)$$

which measures the accuracy of the classifier at detecting 'disease' state (i.e. $y = +1$) subjects. The specificity score is defined as

$$SC = \frac{TN}{TN + FP} \quad (8)$$

which measures the accuracy of the classifier at detecting 'healthy' or 'control' state (i.e. $y = -1$) subjects. The specificity is equal to one minus the false positive rate.

Normalization increases the accuracy of the results, as can be seen in Fig. 5.

6. Discussion and conclusion

The selection of effective and better biomarkers (features) of brain MRI scans for AD, and the multistage classification model for AD detection and image retrieval, were investigated in this paper. The swarm intelligent technique - PSO for feature selection used in this study was performed to reflect the brain structural change, which is related to the clinical progress of AD. The feature selection technique was examined using several feature sets from MRI scans: cortical thickness features, volume features, as well as a combination of thickness and volume.

The multistage classifier used in this thesis produced a good performance for AD detection as compared with previous individual machine learning approaches, such as SVM and KNN. Also, the image retrieval scheme, following the proposed method for AD classification, produced good results. Significant improvements were observed in retrieval speed and accuracy during the implementation.

However, despite this potential improvement, there are some problems for discriminating MCI and NC states versus AD. As seen in Table 4 (confusion matrix), the fault alarm rate for the AD class was smaller than the other classes - NC and MCI. This means that advanced biomarkers and biochemical information are needed to combine structural MRI biomarkers for better performance of the diagnosis between MCI and NC classes. Additionally, the selection of a first stage classifier is important in the building of multistage classifier. The final purpose of AD detection is to diagnose the disease at an earlier or mild stage. Therefore, research for extraction and selection of effective biomarkers that represent MCI and NC stages should be further studied in future work.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research Development, LLC.; Johnson, Johnson Pharmaceutical Research Development, LLC.; Lumosity; Lundbeck; Merck Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research are providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California.

Appendix 1. Selected Thickness and Volume features by PSO algorithm

	Thickness	Volume
1	lh_bankssts_thickness	Left-Inf-Lat-Vent
2	lh_caudalanteriorcingulate_thickness	Left-Cerebellum-White-Matter
3	lh_caudalmiddlefrontal_thickness	Left-Cerebellum-Cortex
4	lh_fusiform_thickness	Left-Thalamus-Proper
5	lh_inferiorparietal_thickness	Left-Caudate
6	lh_isthmuscingulate_thickness	Left-Putamen
7	lh_lateralorbitofrontal_thickness	Left-Pallidum
8	lh_lingual_thickness	Brain-Stem,
9	lh_medialorbitofrontal_thickness	Left-Hippocampus
10	lh_middletemporal_thickness	CSF
11	lh_parahippocampal_thickness	Left-Accumbens-area
12	lh_paracentral_thickness	Left-VentralDC
13	lh_parsorbitalis_thickness	Left-vessel
14	lh_parstriangularis_thickness	Left-choroid-plexus
15	lh_posteriorcingulate_thickness	Right-Lateral-Ventricle
16	lh_precuneus_thickness	Right-Cerebellum-White-Matter
17	lh_rostralanteriorcingulate_thickness	Right-Hippocampus
18	lh_rostralmiddlefrontal_thickness	Right-Amygdala
19	lh_superiorfrontal_thickness	Right-choroid-plexus
20	h_superiorparietal_thickness	5th-Ventricle
21	lh_superiortemporal_thickness	non-WM-hypointensities
22	lh_supramarginal_thickness	Left-non-WM-hypointensities
23	lh_temporalpole_thickness	Right-non-WM-hypointensities
24	lh_transversetemporal_thickness	CC_Posterior
25	lh_insula_thickness	CC_Central
26	lh_MeanThickness_thickness	BrainSegVolNotVentSurf
27	rh_caudalanteriorcingulate_thickness	lhCortexVol
28	rh_cuneus_thickness	rhCortexVol
29	rh_entorhinal_thickness	CortexVol
30	rh_fusiform_thickness	rhCerebralWhiteMatterVol
31	rh_inferiorparietal_thickness	SubCortGrayVol
32	rh_inferiortemporal_thickness	TotalGrayVol
33	rh_isthmuscingulate_thickness	SupraTentorialVol
34	rh_lateraloccipital_thickness	SupraTentorialVolNotVent
35	rh_lateralorbitofrontal_thickness	SupraTentorialVolNotVentVox
36	rh_lingual_thickness	MaskVol
37	rh_medialorbitofrontal_thickness	MaskVol-to-eTIV,
38	rh_middletemporal_thickness	lhSurfaceHoles
39	rh_parsorbitalis_thickness	
40	rh_pericalcarine_thickness	
41	rh_postcentral_thickness	
42	rh_precentral_thickness	
43	rh_precuneus_thickness	
44	rh_rostralanteriorcingulate_thickness	
45	rh_rostralmiddlefrontal_thickness	
46	rh_superiorfrontal_thickness	
47	rh_superiorparietal_thickness	
48	rh_superiortemporal_thickness	
49	rh_insula_thickness	
50	rh_MeanThickness_thickness	

References

[1] Adaszewski S, Dukart J, Kherif F, Frackowiak R, Draganski B. Alzheimer's Disease Neuroimaging Initiative. How early can we predict Alzheimer's disease using computational anatomy? Neurobiol Aging 2013;34(12):2815–26.

[2] Demirhan A. Classification of structural MRI for detecting Alzheimer's disease. International Journal of Intelligent Systems and Applications in Engineering 2016;4(1):195–8.

[3] Trojacanec Katarina, et al. Image retrieval for Alzheimer's disease based on brain atrophy pattern. International conference on ICT innovations springer, cham. 2017. p. 165–75.

[4] Achterberg HC, et al. Hippocampal shape is predictive for the development of dementia in a normal, elderly population. Hum Brain Mapp 2014;35(5):2359–71https://doi.org/10.1002/hbm.22333.

[5] Gerardin E, Chételat G, Chupin M, Cuingnet R, Desgranges B, Kim HS, Niethammer M, Dubois B, Lehéricy S, Garnero L, Eustache F. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 2009;47(4):1476–86. https://doi.org/10.1016/j.neuroimage.2009.05.036.

[6] Lillemark Lene, et al. Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. BMC Med Imaging 2014;14(1):21https://doi.org/10.1186/1471-2342-14-21.

[7] Klöppel Stefan, et al. Automatic classification of MR scans in Alzheimer's disease. Brain 2008:681–9. https://doi.org/10.1093/brain/awm319. 131.3.

[8] Kirill T, et al. Multi-stage classifier design. JMLR: Workshop and Conference Proceedings 2012;25:459–74. https://doi.org/10.1007/s10994-013-5349-4.

[9] Cuingnet Rémi, 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(2):766–81.

[10] Falahati F, et al. Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. J Alzheim Dis 2014;41:685–708.

[11] Ramani A, et al. Quantitative MR imaging in Alzheimer disease. Radiology 2006;241:26–44.

[12] Bron EE, et al. Standardized evaluation of algorithms for computer-aideddiagnosis of dementia based on structural MRI: the CADDementia challenge. Neuroimage 2015;111:562–79.

[13] Hill DLG, et al. Coalition against major diseases/European medicines agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease. Alzhe imers Demen t 2014;1 0.

[14] Poulin SP, et al. For the Alzheimer's Disease Neuroimaging Initiative, Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. Psychiatr Res 2011;194:7–13.

[15] Tanabe JL, et al. Tissue segmentation of the brain in Alzheimer disease. AJNR Am J Neuroradiol 1997;18:115–23.

[16] Weiner, M.W., et al., For the Alzheimer's disease neuroimaging initiative, 2012, the Alzheimer's disease neuroimaging initiative: a review of papers published since its

inception, Alzheimer's Dementia 8, S1–S68.
[17] Eskildsen, S.F., et al., For the Alzheimer's Disease Neuroimaging Initiative, 2013, Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning, Neuroimage 65, 5 11–521.
[18] Singh V, et al. Spatial patterns of cortical thinning in mild cognitive impairment and Alzheimer's disease. Brain 2006;129:2885–93.
[19] Chincarini A, et al. For the Alzheimer's Disease Neuroimaging Initiative, Sep.Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 2011;58:469–80.
[20] Sørensen, L., et al., For the Alzheimer's Disease Neuroimaging Initiative and the Australian Imaging bio-markers and Lifestyle flagship study of ageing, 2016, Early detection of Alzheimer's disease using MRI hippocampa l texture. Hum Brain Mapp 37, 1148–1161.
[21] Braak H, Braak E. Neuropathological staging of Alzheimer related changes. Acta Neuropathol 1991;82:239–59.
[22] Colliot O, et al. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. Radiology 2008;248:194–201.
[23] Wyman, B.T., et al., For the Alzheimer's Disease Neuroimaging Initiative, 2013, Standardization of analysis sets for reporting results from ADNIMRI data, Alzheimer's Dementia 9, 332–337.
[24] Kalbkhani H, Shayesteh MG, Zali-Vargahan B. Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series. Biomed Signal Process Control 2013;8(6):909–19. https://doi.org/10.1016/j.bspc.2013.09.001.
[25] Güzel C, Mahmut Kaya M, Yıldız O. Breast cancer diagnosis based on naïve Bayes machine learning classifier with KNN missing data imputation. In 3rd world conference on innovation and computer sciences. 2013.
[26] Eberhart Russell, Kennedy James. A new optimizer using particle swarm theory." Micro Machine and Human Science. Proceedings of the sixth international symposium on. IEEE, 1995. 1995. p. 39–43. MHS'95.
[27] Poli Riccardo, Kennedy James, Blackwell Tim. Particle swarm optimization. Swarm intelligence 2007:33–57. 1.1.