ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2017.0234

A ROBUST OPTIMIZED FEATURE SET BASED AUTOMATIC CLASSIFICATION OF ALZHEIMER'S DISEASE FROM BRAIN MR IMAGES USING K-NN AND ADA-BOOST

Rupali S. Kamathe[1] and Kalyani R. Joshi[2]
[1]*Department of Electronics and Telecommunication Engineering, College of Engineering, Pune, India*
[2]*Department of Electronics and Telecommunication Engineering, P.E.S's Modern College of Engineering, India*

Abstract

For individuals suffering from some cognitive impairment, treatment plans will be greatly help patients and medical practitioners, if early and accurate detection of Alzheimer's disease (AD) is carried out. Brain MR Scans of patients' with health history and supportive medical tests results can lead to distinguish between Healthy/ Normal Controls (NC), Mild Cognitive Impairment (MCI) and AD patients. However manual techniques for disease detection are labour intensive and time consuming. This work is towards the development of Computer Aided Diagnosis (CAD) tool for Alzheimer's disease detection and its classification into the early stage of AD i.e. MCI and later stage –AD. The paper is about selection of robust optimized feature set using combination of forward selection and/or backward elimination method with K-NN classifier and validation of results with features selected (using forward selection method); with Ada-boost for improved classification accuracy. The features are extracted on Gray Level Co-occurrence Matrix (GLCM). The experimentation is based on Public Brain Magnetic Resonance datasets named Open Access Series of Imaging Studies (OASIS) [7] with patients diagnosed with NC, MCI and AD. The four models considered for automatic classification are – i. Abnormal vs. Normal; ii. AD vs. MCI; iii. MCI vs. NC and iv. AD vs. NC. Feature set optimized using K-NN and validated with AdaBoost has given improved classification accuracy for each model. The output of developed CAD system is compared with Radiologists opinion for test images and has shown 100% match between the output of computer algorithm and experts opinion for some important models under consideration.

Keywords:

Feature Extraction, Feature Selection, Computer Aided Diagnosis, Mild Cognitive Impairment, Alzheimer's Disease

1. INTRODUCTION

Structural Magnetic Resonance (MR) imaging, due to its high resolution and high brain tissue contrast images, has become a valuable tool in diagnosis of different neurodegenerative diseases. However it is still challenging to detect the abnormality in MR brain images automatically, mainly due to the requirement of detection techniques with high accuracy within quick convergence time [1].

Alzheimer's disease (AD), the most common type of dementia, is a major cause of disability worldwide. Early and accurate detection of AD is essential to provide the appropriate treatments to patients and monitor its effectiveness. For diagnosing AD, the disease progression is considered as consisting of three stages: Normal or Healthy Controls (NC), Mild Cognitive Impairment (MCI), a transitional stage between normal ageing and dementia' and AD. MCI subjects have been considered to be at an increased risk of developing AD and hence distinguishing MCI individuals is very important.

Automated computer-aided classification techniques with statistical learning algorithms have much to offer to the physicians towards accurate diagnosis of such complex diseases. One of the very important steps in such systems to be highly accurate is the selection of most suitable features. The features should be such that, within the class the similarity should be maximized and between classes it should be minimized. A limited yet salient (optimized) feature set simplifies the pattern representation and in turn the performance of the classifiers. This paper presents the work towards finding the optimal feature set that increases the efficiency of the learning algorithm, K-NN and validating the optimization with AdaBoost for further improvement in classification accuracy of 4 models as described in the abstract.

Till the date lot of attention [1], [4], [6], [8]-[10], [12], [13], [15] and [17] has been received to automatic detection and classification of various progression stages of AD. Automated hippocampal segmentation (one of the anatomical structure indicating AD) using different machine learning algorithms such as: Hierarchical AdaBoost, SVM with manual and automated feature selection and a publicly available brain segmentation package is carried out by Morra et al. in [3]. Zhou et al. [6] proposed method to combine MRI data with a neuropsychological test - Mini Mental State Examination (MMSE), with classifier for different disease classes such as AD, MCI including amnestic MCI (aMCI) and non-amnestic MCI (naMCI). Westman et al. [10] combined raw cortical thickness measures with sub-cortical volumes normalized by intracranial volume to give the best prediction accuracy for separating AD cases form normal or healthy cases. Liu et al. [12] proposed a hierarchical ensemble classification algorithm to gradually combine the features and decisions (from number of low level classifiers) into a unified model (combination of high level classifiers) for more accurate classification. Fan [13] proposed an ordinal ranking based classification method for classes such as NC, MCI non-converter (MCI-NC), MCI converter (MCI-C) and AD. Zhanga et al. [14] proposed a new multimodal data fusion and classification method based on kernel combination for distinguishing between AD and MCI. Dessouky et al. [16] proposed CAD system using Linear SVM for binary classifications among AD patients with selected features based on intensity levels.

Though lot of work is going on towards the automatic classification of different progression stages of AD, the wide variety of number and type of classes is found to be considered with the choice of various learning algorithms. Feature sets considered for classification are also of different types; still a robust CAD system with 100% classification accuracy for different models based on class labels for different disease progression stages is an open challenge. This paper presents the

robust and optimized feature set selection and classification using K-NN and AdaBoost for four models:

i. Abnormal (MCI +AD) vs. Normal;
ii. AD vs. MCI;
iii. MCI vs. NC and
iv. AD vs. NC.

The Table.10 in Results and discussion section shows improvement in classification accuracy for every model with this work as compared to the previous work done by researchers.

This paper is organized as follows: section 2 and 3 are about K-NN and AdaBoost classifier respectively; section 4 presents feature extraction and selection followed by methodology. The details of data sets used for experimentation and quality assessment parameters are described in section 5. Section 6 covers results and discussions and section 7 presents conclusion.

2. K-NEAREST NEIGHBOR (K-NN) CLASSIFIER

K-NN algorithm is used in pattern recognition for classifying objects based on nearest training examples (based on value of K nearest neighbors by majority vote) in the feature space. In this method the unknown sample may be classified based on the classification of this nearest neighbor.

In K-NN the distance between input test instance and training set instance is computed using different 'distance functions'. We used the most common distance function for K-NN, the Euclidean distance:

$$d(x,y)=\left(\sum_{i=1}^{m}(x_i-y_i)^2\right) \quad (1)$$

where, x is the test feature vector and y is the training feature vector.

3. ADABOOST CLASSIFIER

AdaBoost, a machine learning algorithm is formulated by Freund and Schapire. AdaBoost is adaptive because the instances misclassified by previous classifier are reorganized into the subsequent classifier to improve the classification accuracy and is often used in conjunction with many other weak learning algorithms to boost their performance.

The boosting algorithm initially assigns equal weight to all instances in the training data. Based on classifiers output to this training data, the weight is reassigned for each instance. While reassigning the weights, the one for each correctly classified instance is decreased, and that of misclassified ones is increased.

In the next iteration, a classifier reworks on this reweighed data and tries to correctly classify the instances with more weight. Again the weights are updated based on the new classifier's output. Every time when the weights are updated, the normalization is carried out to maintain the sum of weights as it was before. After all iterations, the final hypothesis value is calculated. The final hypothesis is either +1 or -1, which is the prediction of the Strong Classifier.

4. FEATURE EXTRACTION AND SELECTION

One of the tasks of data pre-processing is to prepare the data for input into the machine learning algorithm is, 'feature selection'; in which only some of the features from the dataset are selected and used in the training process of the learning algorithm. In this process the aim is to find the optimal subset that increases the classification efficiency of the learning algorithm and decreases the computational burden.

Thus, the resulting classifier will be faster, more accurate and will use less memory. Feature selection (FS) can be done using Supervised Learning and Unsupervised Learning. FS algorithms can be broadly fall into three categories: the Filter Model, the Wrapper Model and the Hybrid Model. Features can be extracted on Gray-Level Co-occurrence Matrix (GLCM).

4.1 GRAY-LEVEL CO-OCCURRENCE MATRIX (GLCM)

GLCM also known as the gray-level spatial dependence matrix is the basis for the Haralick's texture features [19]. This matrix is square with dimension N_g, where N_g is the number of grey levels in the image. Element [i, j] of the matrix is generated by counting the number of times a pixel with value i is adjacent to a pixel with value j and then dividing the entire matrix by the total number of such comparisons made. Each entry is therefore considered to be the probability that a pixel with value i will be found adjacent to a pixel of value j and orientation θ.

$$G=\begin{bmatrix} p(1,1) & p(1,2) & \cdots & p(1,N_g) \\ p(2,1) & p(2,2) & \cdots & p(2,N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,1) & \cdots & p(N_g,N_g) \end{bmatrix} \quad (2)$$

5. METHODOLOGY

5.1 DATASETS

We used the Open Access Series of Imaging Studies (OASIS) [7]; publicly available dataset for experimentation which consists of total 416 subjects aged 18 to 96. Out of the complete data, 201 subjects (107 NCs, 68 MCIs, and 26 ADs) are labelled. For each subject, 3 individual T1-weighted 3D MR images as axial, coronal and sagittal images obtained in a single scan session are included. For experiment only axial images are selected. The subjects include both men and women. Subjects with Clinical Dementia Rating (CDR) of 0, 0.5, and 1 are considered NC, MCI, and AD, respectively.

In this study, four datasets are used shown in the Table.1 containing T1 weighted axial brain MR images.

For entire experimentation Haralick's [18] features based on GLCM are used. GLCM is calculated for distance d = 1 with angles θ = 0°, 45°, 90° and 135°. Each element in normalised GLCM is labelled as (r, c). Each element represents the joint probability occurrence of pixel pairs with a defined spatial relationship having grey level values 'r' and 'c' in the image.

Table.1. T1 images for Model I-IV, Ab -Abnormal

Model No.	Model Name	Total No of Images
I	Ab vs. NC	201 (Ab = 94; AD = 26, MCI = 68, NC = 107)
II	AD vs. MCI	94 (AD= 26, MCI= 68)
III	AD vs. NC	133 (AD= 26, NC= 107)
IV	MCI vs. NC	175(MCI=68, NC =107)

5.2 PERFORMANCE EVALUATION PARAMETERS

The performance of classifier (K-NN or AdaBoost) at every stage of feature selection is measured and compared with the parameters like Accuracy (*Acc*), Error Rate (*ER*), Sensitivity (*Sen*), Specificity (*Spe*) and Profile Time. For this True Positive (*TP*), True Negative (*TN*), False Positive (*FP*) and False Negative (*FN*) image counts based on the true class of the test sample and assigned class by the classifier at every stage during cross validation process are taken into account. The evaluation parameters are calculated as follows:

Acc or Correct Rate: Correctly Classified samples / Total samples i.e.

$$Acc = (TP+TN) / (TP+TN+FP+FN) \quad (3)$$

ER: Incorrectly Classified samples / Total samples i.e.

$$Er = (FP+FN) / (TP+TN+FP+FN) \quad (4)$$

Sen: Correctly Classified Positive samples / True Positive samples i.e. true positive fraction

$$Sen = TP / (TP+FN) \quad (5)$$

Spe: Correctly Classified Negative samples / True Negative samples i.e. true negative fraction.

$$Sp = TN / (TN+FP) \quad (6)$$

Also the 'Profile Time' i.e. the CPU time at each step of feature selection is recorded.

5.3 FEATURE SET OPTIMIZATION AND VALIDATION

The work is done in two parts:

- ***Part A***: Selection of optimized Features with forward selection or backward elimination with K-NN Classifier
- ***Part B***: Validating the feature set selection done in Part A for improving classification accuracy with Adaboost Classifier.

The steps followed in Part A and B, are listed below:

Part A: K-NN based features set selection

i. *Starting point*: 4 GLCM Features: Contrast, Correlation, Energy and Homogeneity [18] are selected based on verifying the similarity of feature vectors within the class and between the classes with the help of Normalized Cross Correlation (NCC) (Refer Table.2).

ii. Performance of K-NN for Model I is assessed with different parameters, listed in section 5.2 (Refer Table.3).

iii. New feature is added (forward Selection) and step ii is repeated.

iv. Step iii is continued till the addition of feature in training data set has not shown any significant improvement in classifying the samples as normal or abnormal.

v. The feature not showing significant improvement in classification accuracy is then removed (backward elimination) followed by step ii.

vi. Step v is repeated till the classification accuracy value is reached at its maximum for different *K*-folds in the cross validation process and number of neighbors (*K*).

Following steps i-vi, the optimal feature set for Model I for highest possible classification accuracy is obtained. The same feature set is used for Model II-IV with cross-validation for *K*-fold = 3, 5, 7, 9, 10, 12, 15 and *K* (Number of nearest neighbour to be considered in K-NN) =1, 3, 5, 7 and 9.

Part B: AdaBoost based validation of feature set selected from Part A

i. *Starting point*: Out of initial 4 GLCM features used in step i of Part A, 'Contrast' feature is used for training AdaBoost.

ii. The performance of AdaBoost is assessed with different evaluation parameters for Model I.

iii. Next feature is added from the optimized feature set of Part A and step ii is repeated.

iv. Step iii is repeated for entire set of features obtained from Part A.

v. To ensure that the optimal features are selected for highest possible classification accuracy even with the AdaBoost, few other features from Haralicks's feature [19] are added one by one and are tested. For entire experimentation with AdaBoost the number of weak classifiers are considered till the training error is zero (preferred) or reached to minimum possible value.

vi. The performance of each feature set is evaluated based on the parameters like no of iterations (number of weak classifiers), classification accuracy, time required for classification etc.

The final optimal feature set is used for testing the performance of Model II-IV.

The above fully automatic CAD system (with K-NN and Adaboost) is tested for images from OASIS database (without CDR rating) i.e. the expert ratings are not available for these images; and also for images from the different hospitals. The class labels assigned are compared with that of the Radiologists' rating/ labels. The performance is presented in Section 6 with the help of Receiver Operating Characteristics (ROC) curves as shown in Fig.2.

6. RESULTS AND DISCUSSION

***Part A*:**

Selection of starting point for feature selection (for Model I) based on correlation:

Table.2 shows that first 4 GLCM features are highly uncorrelated between the two classes under consideration. Hence these 4 features are selected as base features (starting point: Part A).

Features considered (in the order of their use):- Contrast, Correlation, Energy, Homogeneity, Inverse difference moment (IDM), Entropy, Symmetrical Feature, Spatial Frequency, Information Measure of Correlation, Absolute Value and Maximum probability [19].

Table.2. Correlation Result for 6 Features

Features	Contrast	Correlation	Energy	Homogeneity	IDM	Entropy
Contrast	1.0000	-0.2247	-0.2479	-0.3406	-0.3093	0.2636
Correlation	-0.2247	1.0000	0.1026	0.0947	0.0907	-0.1316
Energy	-0.2479	0.1026	1.0000	0.8105	0.8428	-0.7462
Homogeneity	-0.3406	0.0947	0.8105	1.0000	0.9952	-0.8335
IDM	-0.3093	0.0907	0.8428	0.9952	1.0000	-0.8181
Entropy	0.2636	-0.1316	-0.7462	-0.8335	-0.8181	1.0000

Table.3. Results for Forward Selection Process

K-fold	K	TP	TN	FP	FN	Acc	Error Rate	Sensitivity	Specificity
4 GLCM features (Contrast+ Correlation+ Energy+ Homogeneity)									
3	3	11	25	10	20	0.5455	0.4545	0.3548	0.7143
9	5	3	5	2	3	0.6154	0.3846	0.5000	0.7143
5 GLCM features (Contrast+ Correlation + Energy+ Homogeneity + IDM)									
5	3	7	16	6	12	0.5610	0.4390	0.3684	0.7273
7	3	7	11	4	7	0.6207	0.3793	0.5000	0.7333
6 GLCM features (Contrast+ Correlation+ Energy+ Homogeneity + IDM+Energy)									
9	3	6	9	3	5	0.6522	0.3478	0.5455	0.7500
10	5	5	9	2	4	0.7000	0.3000	0.5556	0.8182

Table.4. Forward selection results continued till 9 features

No of features	K-fold= 10		For K-fold= 15	
	K	Acc (%)	K	Acc (%)
4	7	66.67	5	61.54
5	9	65.00	7	69.23
6	5	70.00	3	64.29
7	9	76.19	9	76.92
8	3	76.19	9	92.31
9	3	75.00	3	78.57

As seen in Table.4, the accuracy of the classifier goes on increasing with sequential addition of features. The maximum accuracy 92.31% with K-fold = 15 for 8 features (Average including different K values is less). But no further improvement is seen with any new feature, so the approach of backward elimination process has been adopted.

Table.5. Backward elimination

No. of features	Features	K-fold	K	Acc
8	Contrast + Correlation + Energy + Homogeneity + IDM+ Entropy + Information Measure of Correlation + Absolute Value	10	9	0.7500
		15	9	0.7143
7	Contrast + Correlation + Energy + Homogeneity + IDM+ Entropy + Information Measure of Correlation	10	5	0.8000
		15	3	0.7692
6	Contrast + Correlation + Energy + Homogeneity + Information Measure of Correlation + Absolute Value	10	3	0.7368
		15	3	0.7692

Based on experimentation and the improvement in accuracy (average) the 6 features are selected as listed in Table.5.

Table.6. K-NN performance for Model II-IV with 6 (optimized) features

Model	K-fold	K	Acc	Error Rate	Sensitivity	Specificity
II	10	1	0.8889	0.1111	0.5000	1.0000
	10	3	0.8000	0.2000	0.3333	1.0000
	15	7 , 9	0.8333	0.1667	0.0000	1.0000
III	10	3,5,7,9	0.8462	0.1538	0.0000	1
	15	3,5,7,9	0.8750	0.1250	0.0000	1.0000
IV	10	9	0.7647	0.2353	0.7143	0.8000
	15	9	0.8333	0.1667	0.6000	1.0000

***Part B*:**

Table.7. Forward Selection Using Adaboost Classifier (Model I)

Feature Set	No. of Iterations (T)	TP	TN	FP	FN	Acc	Error Rate	Profile Time (s)
1	352	69	80	27	25	0.7412	0.2590	7.134
2	475	74	96	11	20	0.8457	0.145	13.199
3	198	75	91	16	19	0.8258	0.1730	8.289
4	483	84	96	11	10	0.8955	0.0990	20.896
5	184	76	93	14	18	0.8407	0.1500	11.587
6	196	82	93	14	12	0.8700	0.1200	13.096

As seen in Table.7, the 6 features selected by K-NN (Part A) showed optimal performance in terms of classification accuracy and number of iterations per weak classifiers and CPU time required for the process with AdaBoost classifier.

Table.8. Adaboost performance for Model I-IV with 6 (optimized) features

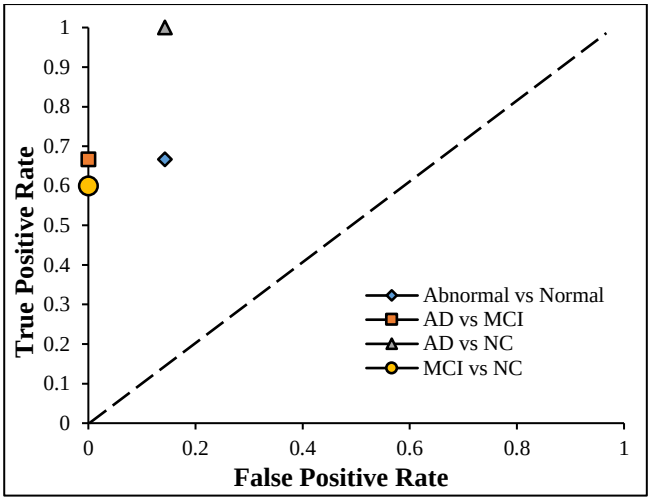
Model	Feature Set	No of Iterations	TP	TN	FP	FN	Acc	Error Rate	Time Required (s)
Model I- Abnormal vs. Normal	1	212	69	80	27	25	0.7412	0.2590	7.134
	2	475	74	96	11	20	0.8457	0.145	13.199
	3	198	75	91	16	19	0.8258	0.1730	8.289
	4	483	84	96	11	10	0.8955	0.0990	20.896
	5	184	76	93	14	18	0.8407	0.1500	11.587
	6	196	82	93	14	12	0.8700	0.1200	13.096
Model II AD vs. MCI	1	326	26	68	0	0	1.0000	0.0000	7.325
	2	185	26	68	0	0	1.0000	0.0000	6.054
	3	117	26	68	0	0	1.0000	0.0000	5.409
	4	102	26	68	0	0	1.0000	0.0000	5.722
	5	111	26	68	0	0	1.0000	0.0000	7.086
	6	129	26	68	0	0	1.0000	0.0000	8.889
Model III AD vs. NC	1	441	26	107	0	0	1.0000	0.0000	9.123
	2	272	26	107	0	0	1.0000	0.0000	7.965
	3	154	26	107	0	0	1.0000	0.0000	6.633
	4	125	26	107	0	0	1.0000	0.0000	6.573
	5	139	26	107	0	0	1.0000	0.0000	8.334
	6	131	26	107	0	0	1.0000	0.0000	9.399
Model IV MCI vs. NC	1	296	40	93	14	28	0.7600	0.2420	11.347
	2	1185	54	104	3	14	0.9028	0.1030	34.045
	3	681	54	104	3	14	0.9028	0.1030	31.229
	4	507	54	104	3	14	0.9028	0.1030	38.430
	5	266	56	102	5	12	0.9028	0.1000	25.004
	6	391	58	100	7	10	0.9028	0.0970	34.893

Table.9. Comparison of K-NN and AdaBoost performance

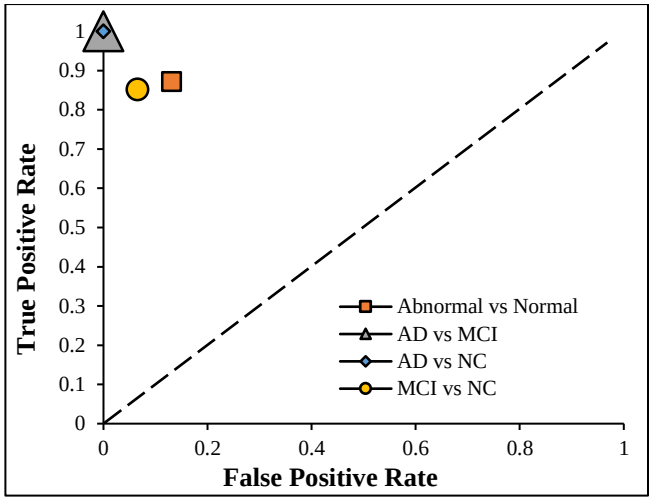
Model Name	Accuracy (%)	
	K-NN	AdaBoost
Abnormal vs. Normal	76.92	87
AD vs. MCI	92.31	100
AD vs. NC	92.75	100
MCI vs. NC	83.33	90.28

The Fig.1 shows that the K-NN and Adaboost perform extremely well with optimized 6 features for all four Models for classification between different progression stages of AD.

The Fig.2 shows the comparison between the Radiologist's class labels and the output of our CAD system with i. K-NN and ii. AdaBoost; which performed well for Models [I-IV] under consideration.



(a)



(b)

Fig.1. ROC for (a) K-NN Classifier and (b) Adaboost Classifier

Table.10. Comparison of our method with previous results

Classifier	No. of Images	Features	Image database	Results
Saliency based pattern extraction [1]	NC=66; MCI=20	Intensity, Contrast (18 Features)	OASIS-MIRIAD	Acc = 86.05%; Sen = 85%; Spe = 86.36%
Adaboost, SVM [4]	NC=10; MCI=10; AD=10	Intensity, Mean (100 features)	ICBM	AdaSVM: R = 78, L = 85; SVM: R = 36, L = 75
SVM[7]	NC=59; MCI=123; AD=127	Statistical features	Private	Acc = 92.40%; Sen = 84%; Spe = 96.10%
SVM [9]	NC=75; AD=75	Intensity, Standard Deviation	ICBM	Acc = 92 %
SVM [10]	NC= 16; AD= 22	Mean	Onset Dating	Acc = 94.5%; Sen = 91.5%; Spe = 96.6%
Multi-variate Analysis [11]	NC=255; MCI=287; AD=187	Shape based features	ADNI	AD vs. NC = 91.5% MCI vs. AD = 75.9%
Single, Multilayer Classifier [13]	NC= 16 AD=22	Correlation context features	ADNI	CL: Single, Multi Acc: 86%, 92% Sen: 84%, 90% Spe: 88%, 93%
Ordinal Ranking [14]	NC=55; MCI=101; AD=51	Rank based features	ADNI	MCI vs. NC = 9% AD vs. NC = 2%
LSVM [15]	NC=52; AD=51; MCI=99	Shape based features	ADNI	Acc = 93.2%; Sen = 93%; Spe = 93.3%
K-NN [This work]	NC= 107; Abnormal (Ab)=94; MCI=68; AD=26	6 Features	OASIS	Ab vs. NC = 76.92% AD vs. MCI = 92.31% AD vs. NC = 92.75% MCI vs. NC = 83.33%
AdaBoost [This work]				Ab vs. NC = 87% AD vs. MCI = 100% AD vs. NC = 100% MCI vs. NC = 90.28%

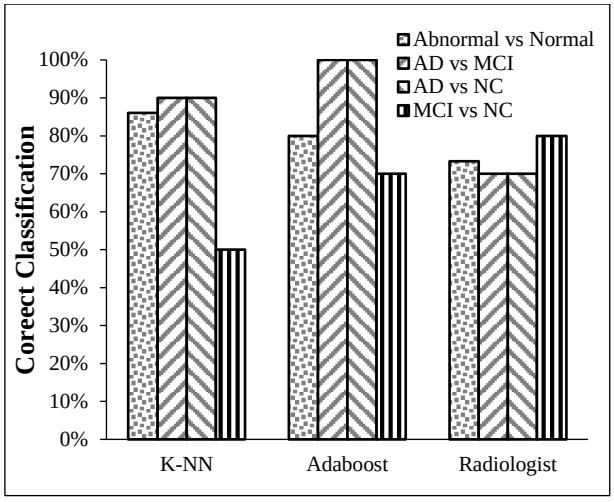


Fig.2. Comparison between classifier Results and Radiologist's feedback

7. CONCLUSION

In this work, we have developed a computerized support system for radiologists/neurologists with three class sets as NC, MCI and AD for MR Images. The system is based on minimum number of features (optimum) through exhaustive experimentation with two classifiers- K-NN and AdaBoost. The system has shown notable performance as compared to previous work done and also as compared to radiologist's feedback in assigning the labels such as Normal, MCI or AD. The improved classification accuracy and reduced computation time for automatic detection of Alzheimer's disease and even the progression stage of AD i.e. MCI are the important characteristics of this system. This CAD system can be proved be useful in clinical diagnosis of stages of AD from brain MR scans and may also be used for training new readers to help them to understand the disease progression based on MR scans by correlating it with medical examinations and providing some additional biomarkers.

ACKNOWLEDGEMENTS

Authors are thankful to Dr. Sangolkar, Nizam's Institute of Medical Sciences, Hyderabad, India for giving us valuable feedback during the testing phase. Because of his support we could perform comparison of developed CAD system with the performance of one of the probable end user of such system-the Neuro Radiologist. We are also thankful to Dr. Geetanjali Kadam, Deenanath Mangeshkar Hospital, Pune, India for showing interest in this system.

REFERENCES

[1] R. Andrea and R. Eduardo, "Extracting Salient Brain Patterns for Imaging-Based Classification of Neurodegenerative Diseases", *IEEE Transactions on Medical Imaging*, Vol. 33, No. 6, pp. 1262-1274, 2014.

[2] H.Y. Dong, M.P. Kilian and D. Christos, "Semi-Supervised Pattern Classification: Application to Structural MRI of Alzheimer's Disease", *Proceedings of IEEE International Workshop on Pattern Recognition in Neuroimaging*, pp. 101-105, 2011.

[3] H.M. Jonathan, T. Zhuowen, G.A. Liana, E.G. Amity, W.T. Arthur and M.T. Paul, "Comparison of AdaBoost and Support Vector Machines for Detecting Alzheimer's disease through Automated Hippocampal Segmentation", *IEEE Transactions on Medical Imaging*, Vol. 29, No. 1, pp. 30-43, 2010.

[4] H.A. Mark, "Correlation-Based Feature Selection for Machine Learning", PhD Dissertation, Department of Computer Science, University of Waikato, 1999.

[5] A.F. Peter and C.F. Nick, "MR Image Texture Analysis Applied to the Diagnosis and Tracking of Alzheimer's Disease", *IEEE Transactions on Medical Imaging*, Vol. 17, No. 3, pp. 475-479, 1998.

[6] Z. Qi, G. Mohammed, C. Mercedes, W. Jin, B. Warren, A.L. David, D. Ranjan and A. Malek, "An Optimal Decisional Space for the Classification of Alzheimer's Disease and Mild Cognitive Impairment", *IEEE Transactions on Biomedical Engineering*, Vol. 61, No. 8, pp. 2245-2253, 2014.

[7] Open Access Series of Imaging Studies, Available at: http://www.oasis-brains.org/

[8] D. Simon, et al., "MRI-Based Automated Computer Classification of Probable Ad Versus Normal Controls", *IEEE Transactions on Medical Imaging*, Vol. 27, No. 4, pp. 509-520, 2008.

[9] Benoit Magnin, et al., "Support Vector Machine-based Classification of Alzheimer's Disease from Whole-Brain Anatomical MRI", *Neuroradiology*, Vol. 51, No. 2, pp. 73-83, 2008.

[10] E. Westman, C. Aguilar, J.S. Muehlboeck and A. Simmons, "Regional Magnetic Resonance Imaging Measures for Multivariate Analysis in Alzheimer's Disease and Mild Cognitive Impairment", *Brain Topography*, Vol. 26, No. 1, pp. 9-23, 2012.

[11] J.E. Iglesias, J. Jiang, C.Y. Liu and Z. Tu, "Classification of Alzheimer's Disease using a Self-Smoothing Operator", *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, pp. 59-65, 2011.

[12] Manhua Liu, Daoqiang Zhang, Pew-Thian Yap and Dinggang Shen, "Hierarchical Ensemble of Multi-level Classifiers for Diagnosis of Alzheimer's disease", *Proceedings of International Workshop on Machine Learning in Medical Imaging*, pp. 27-35, 2012.

[13] F. Yong, "Ordinal Ranking for Detecting Mild Cognitive Impairment and Alzheimer's Disease based on Multimodal Neuroimages and CSF Biomarkers", *Proceedings of International Workshop on Multimodal Brain Image Analysis*, pp. 44-51, 2011.

[14] Z. Daoqiang, W. Yaping, Z. Luping, H. Hong and S. Dinggang, "Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment", *Neuroimage*, Vol. 55, No. 3, pp. 856-867, 2011.

[15] S. Kioppel et al., "Automatic classification of MR scans in Alzheimer's disease", *Brain: A Journal of Neurology*, Vol. 131, No. 3, pp. 681-689, 2008.

[16] Mohamed M. Dessouky, Mohamed A. Elrashidy and Hatem M. Abdelkader, "Selecting and Extracting Effective Features for Automated Diagnosis of Alzheimer's disease",

International Journal of Computer Applications, Vol. 81, No. 4, pp. 17-28, 2013.

[17] Zhen Nanning and Xue Jianru, "Statistical Learning and Pattern Analysis for Image and Video Processing", Springer, 2009.

[18] R.M. Haralick, K. Shanmugam and D. Itshak, "Texture Feature for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 3, No. 6, pp. 610-621, 1973.