

Amalgamation of Machine Learning and Slice-by-Slice Registration of MRI for Early Prognosis of Cognitive Decline

Manju Jain¹

University School of Information
Communication and Technology, Guru Gobind Singh
Indraprastha University, Delhi India
Meerabai Institute of Technology, New Delhi, India

C.S. Rai²

University School of Information, Communication and
Technology, Guru Gobind Singh Indraprastha University
Delhi India

Jai Jain³

Media Agility India Ltd
New Delhi, India

Deepak Gambhir⁴

Galgotia College of Engineering and Technology
Utter Pradesh, India

Abstract—Brain atrophy is the degradation of brain cells and tissues to the extent that it is clearly indicative during Mini-Mental State Exam test and other psychological analysis. It is an alarming state of the human brain that progressively results in Alzheimer disease which is not curable. But timely detection of brain atrophy can help millions of people before they reach the state of Alzheimer. In this study we analyzed the longitudinal structural MRI of older adults in the age group of 42 to 96 of OASIS 3 Open Access Database. The nth slice of one subject does not match with the nth slice of another subject because the head position under the magnetic field is not synchronized. As a radiologist analyzes the MRI image data slice wise so our system also compares the MRI images slice wise, we deduced a method of slice by slice registration by driving mid slice location in each MRI image so that slices from different MRI images can be compared with least error. Machine learning is the technique which helps to exploit the information available in abundance of data and it can detect patterns in data which can give indication and detection of particular events and states. Each slice of MRI analyzed using simple statistical determinants and Gray level Co-Occurrence Matrix based statistical texture features from whole brain MRI images. The study explored varied classifiers Support Vector Machine, Random Forest, K-nearest neighbor, Naive Bayes, AdaBoost and Bagging Classifier methods to predict how normal brain atrophy differs from brain atrophy causing cognitive impairment. Different hyper parameters of classifiers tuned to get the best results. The study indicates Support Vector Machine and AdaBoost the most promising classifier to be used for automatic medical image analysis and early detection of brain diseases. The AdaBoost gives accuracy of 96.76% with specificity 95.87% and sensitivity 87.37% and receiving operating curve accuracy 96.3%. The SVM gives accuracy of 96% with 92% specificity and 87% sensitivity and receiving operating curve accuracy 95.05%.

Keywords—Brain atrophy; registration; Freesurfer; GLCM; texture features; FDR; decision support system; SVM; AdaBoost; Randomforest Bagging; KNN; Naive Bayes; classification; hyperparameters; GridsearchCV; Sklearn; Python

I. INTRODUCTION

The brain tissues degenerate due to aging a visual difference between normal and atrophied brain shown in Fig.1. Besides age many other factors viz. social and occupational conditions and family history plays a major role in the degradation process of brain tissues resulting in the cognitive skills of the person nosedive.

This effect is measurable during clinical judgment trials in the form of Clinical Dementia Rating (CDR) score. The CDR value zero means the person is cognitive normal but more than zero means the person is with brain atrophy making him cognitive abnormal.

Another biomarker of brain atrophy is the deterioration of medial temporal lobe structure of the brain which is a volumetric detection using Magnetic Resonance Imaging (MRI) a pathological test. The goal of this study and experimentation is to find mapping of clinical findings and corresponding pathological finds using MRI scans. Medial temporal lobe is that anatomic and physiological part of the brain which is responsible for memory retention and retrieval of information. It is that part of the brain where our short-term memories become long term memories. In a way we can say its non-volatile memory of the brain which becomes volatile because of brain atrophy state. That's why we only remember only current events and forget as we lose the reference just as the computer's volatile RAM loses its contents after power is switched off.

Next to find the reasons of dimensional loss, the brain atrophy is characterized by deposits of plaques and neurofibrillary tangles (NFTs), which cause loss of neurons and synapses. The loss and deposits are a simultaneous process which makes it difficult to distinguish and identify.

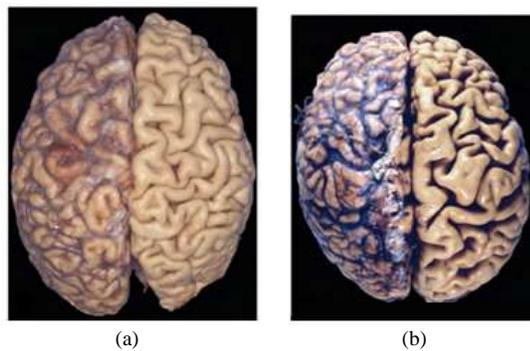


Fig. 1. (a) Normal Adult Brain, (b) Atrophied Brain with Narrow Gyri and Widened Sulci.

The extent of brain atrophy is determined by its anatomically distribution i.e. from stage I to stage VI [1], research and findings shows that major area affected as : stage I & II Entorhinal cortex a very small part behind hippocampus, stage III and IV hippocampus and amygdale stage V and VI neocortex., but the severity of disease is determined by NFTs. Hippocampus is a very compact area of the brain in the medial temporal lobe. It consists of cortical areas and main hippocampus. The cerebral cortex is highly folded as it has to be accommodated into a limited volume of brain skull.

Motivation to exploit the machine learning technology and computer based image processing is that radiologists sometimes find it very difficult to localize the degradation patterns because of many above said complicated and compact structures of the brain secondly Individuals show varied patterns. The MRI data itself is complicated 3D images. The 3D images consist of several slices of 2D images. It becomes very cumbersome for the radiologist to scan each slice and get the correlations. In this study we designed a computer aided decision support system of automatic detection using machine learning techniques which is useful for a radiologist for faster easy and accurate decisions.

II. LITERATURE REVIEW

The past few decades have proved to be promising in early experimentation and studies of detection of medical conditions using machine learning as a tool in combination of image processing.

The advancement in medical technology has led to providing data through various modalities of pathology like X Ray, MRI, fMRI, ultrasound scans and other advanced scans and availability of software to handle this data.

Image processing techniques play a significant role in the accuracy of a study. Some earlier studies used (VBM) voxel based morphometry [2][3][4]. These studies worked on T1 weighted MRI scans on very small groups of subjects, later they used voxel based relaxometry (VBR) on T2 weighted scans of same subjects. In VBM specific tissue templates were used to compare voxel by voxel and they segmented white, grey and cerebrospinal fluid by comparing with reference templates well defined by Montreal Neurologic Institute. The surface reconstruction was done voxel by voxel of size

1.5X1.5X1.5 mm each. But such procedures were too complicated and compromise accuracy.

Another voxel based morphometry study [5] used the comparisons of intensities of white matter, deep white matter and periventricular deep white matter voxel by voxel.

Another image processing technique, deformation based image analysis, was used in several studies [6][7][8]. These studies created a reference space and calculated the deformation required to transfer the individual image into reference space. The other deformation based studies[9] applied Jacobian determinant at each transformation to measure the volume change patterns. The study [7] applied Deformation based morphometry to detect brain changes, but they used the concept of longitudinal DBM where they tried to measure volume changes of same subjects over the period of study.

Tensor based morphometry is another image processing technique used in [10][11]. They designed 3D metrics of disease base differences in brain structures but again a very complicated and time consuming process. Other Tensor based morphometry [12][13] studies created difference tensors of diseased regions and a common anatomical template, at each pixel a colour coded Jacobin determinant calculated that gives a differential change in volumes at region of interest.

A study applied data mining [14] where millions of voxels are mined to select sufficient no of voxels to predict the hypothesis with high accuracy.

All the above studies performed on very small datasets, with changing lifestyle and growing no of cases in brain atrophy and other brain diseases, related data sets have increased manifolds giving researchers a wider domain to work on and yield better results in early detection of brain diseases using machine learning as a tool for both image processing and identification of diseases. The author in [15] applied Machine learning tools on ADNI (Alzheimer's Disease Neuroimaging initiative)database. They work on spatial patterns of abnormalities. It was a massive project and carried out on 16 CPU parallel processing as AD-PS scores computation needs overnight processing using parallel processors. It was extension study of earlier study [16].

The author in [17] used machine learning SVM (Support Vector Machine) combined with voxel based morphometry for early detection of brain atrophy using ADNI database. The classifier is used as an iterator to find the weights associated with each voxel. Voxels with particular weight values were selected as features rest are dropped hence voxels as features are redetermined at every training level. This study finds that study accuracy depends on number of subjects in the database.

Texture analysis may be defined as "the feel, appearance or consistency of a surface or a substance". In our study of Biomedical Image analysis image texture provides information about micro and macro structural changes in the tissues and cells. Radiologist with time train themselves to drive a relationship between visual patterns indicating molecular and cellular properties of tissues. Radiologist face many problems in evaluating and inference the biomedical images:

- Diversity in diseases and anatomy.
- Complicated operational physics behind acquisition tools and dependence on technical staff.
- Non Uniformity image acquisition, interpretation and Reporting.

Computer aided mathematical biomedical image texture analysis provides an aid to radiology by interpreting the image in terms of statistical features and signal variation algorithms giving a quantitative definition of image. List of latest texture based studies [18]-[24] on Brain atrophy MRI are listed in Table 1A.

Limitations of above studies are:

1) These were constrained to very small datasets subject numbers below 200 subjects except few. Most of the studies

are on ADNI1 and ADNI2, OASIS1, OASIS2, the latest published data set OASIS 3 a potential data to be explored.

2) Most of the studies used cross sectional MRI Database than longitudinal, while brain atrophy is a longitudinal study.

3) Most of the studies are ROI (Region of Interest) based. But such studies need a prior and in depth knowledge of the under study disease, means it becomes necessary that one of co-researcher must be from a medical background. Even when we segment the image to get ROI, the classification accuracy will depend on the accuracy of segmentation. Most studies used SPM or free surfer software to get ROI. Most of the above studies consider only the shrinkage of the hippocampus and cerebral cortex and enlargements of ventricles. But brain atrophy is not localized to some segments of the brain but it affects the brain as a whole, hence the whole brain MRI needs to be analyzed slice by slice as most Radiologists do.

TABLE I. (A) EARLIER STUDIES ON THE BRAIN DEGENERATION DISEASES CLASSIFICATION USING TEXTURE ANALYSIS FEATURES

Reference	Dataset	No. of Subjects	Method	Accuracy	Sensitivity	Specificity
Olfa Ben Ahmed 2014 [18]	ADNI	AD218 NC250	Content Based Visual Features from Hippocampus ROI SVM, 1.5 T1 Weighted	87%	Not Available	Not Available
	Bordeaux	AD16 NC21		85%		
Amulya E.R. 2017[19]	OASIS 2	235	Texture Base GLCM, SVM	75.5%	Not Available	Not Available
Tooba Altaf S Anwar, Feb 2018 [20]	ADNI		Hybrid features Texture + Clinical Data ROI and Complete Brain, KNN AdaBoost	79%	79%	92%
				97.8%	95.65%	100%
Loris Nanni May 2019 [21]	ADNI Salvator	AD 137 NC 162	Texture plus Voxel Based, ROI SVM, 1.5 T1 Weighted	78.8%	78.8%	77.4%
				87.6%	84%	90.3%
K W Kim June 2019 [22]	ADNI2		Texture Based GLCM, GLRLM, ROI, SVM 3T1 weighted	73%	65%	100%
Jia-Hui Cai Jan 2020 [23]	ADNI		ROI, Texture Based GLCM, GLRLM	Not Available	Not Available	Not Available
M. Gattu Feb 2020 [24]	ADNI	1167	Cortical Thickness Measurements left and right hippocampal	75%	Not Available	Not Available

III. DATA PRE-PROCESSING

The baseline of sustainable research and development is the infrastructure, data, software and algorithms. This work used the best image analysis environment which provided computational tools and facilitated the reproducible research and data. The Jupyter notebook is used to provide a flexible and well documented workflow. The Python 3.0 gives the very interesting and useful library modules, which make image processing implementation work very easy, like SimpleITK [25] and Nibable, Sklearn.

The study used OASIS-3 latest release December 2019 MRI dataset. Its retrospective data over the period of 15 years consists of 1098 subjects and more than 2000 sessions. The

link to the data is www.oasis-brains.org. The dataset is accompanied with clinical and cognitive assessments. The Table 1B lists the Demographic Details of the Subjects.

In our study we took the patients CDR status at a particular time stamp, and tried to classify for early prognosis of brain atrophy causing cognitive impairment which may lead to Alzheimer.

TABLE I (B) DEMOGRAPHIC DETAILS

	Female Subjects	Male subjects	Total
Number	487	611	1098
Average Age	43-95	42 – 91	

Machine learning approach is data based approach accuracy of study strongly based on data clarity and details because data is the building block of such studies. Besides data acquisition process is not perfect, the MRI scanning results into images which have to pre-processed to improve the accuracy of final results, because the MRI scanning process got affected due to static magnetic field strength, coil variations, tissues penetration difference, eddy currents etc. in MRI machine. The study used Freesurfer [26] open access specialised software for neuroimaging analysis and interpretation of Brain MRI data. The study performed a set of scripts using Freesurfer software to implement preprocessing pipeline procedures as described in Fig. 2.

A. Skull Stripping

It is a process to remove non-skull tissues from the brain MRI Images to improve accuracy of brain image processing to be used for early diagnosis and prognosis of various brain related diseases. Many techniques of brain stripping are used in biomedical image studies.

- **Mathematical Morphometric Method:** This method uses edge detection and thresholding criteria to remove non skull tissues from brain MRIs. It is highly dependent on initial parameters like threshold values.
- **Intensity based Method:** This method uses the intensity of the basic feature of image that is pixel to differentiate non brain tissues from brain tissues by using histogram or region growing method.
- **Deformable surface based Method:** An active contour which works like self growing contour based on energy components of a desirable brain mask is used to separate out brain tissues. It's a very robust method.

B. Inhomogeneity Correction

Inhomogeneity means similar tissues of brain have different pixel intensity during MRI scan of brain, while similar tissues of brain should have approximate same pixel intensities hence this problem is known as inhomogeneity. It is because during MRI scanning process signal intensity is not uniform because different tissues of brain require different magnitude of signal to penetrate so signal is not kept uniform throughout the scan, but this change in signal may result into spikes and inhomogeneity in pixel intensities of same tissues, to correct it signal is convolved with a bias signal using two models additive or multiplicative model. This process is called inhomogeneity correction. If $T(x)$ is the observed image signal with bias field $b(x)$ and noise $n(x)$.

Then two models to represent the observed image signal are:

I Additive Model

$$T(x) = S(x) + b(x) + n(x) \quad (1)$$

II Multiplicative Model

$$T(x) = S(x).b(x) + n(x) \quad (2)$$

$$T(\widehat{x}) = \log S(x) + \log b(x)$$

$$\widehat{T}(x) = \widehat{S}(x) + \widehat{b}(x) \quad (3)$$

(multiplicative model transferred to logarithmic signal).

Inhomogeneity Corrections methods used in this study are:

1) *Modified fuzzy C means:* Modified Fuzzy C means segments the brain into three segments background, white matter and gray matter. To improve the quality of segmentation it adds two more parameters that is Spatial coherence of tissue classes t , tissues can be white matter, Gray matter, cerebrospinal fluid muscle, fat skin or skull or background (as signal penetration depends on type of tissue). And bias field \widehat{b} used to smooth the output image signal. Fuzzy C means jointly segments and estimate the bias field to minimize the inhomogeneity and the joint objective function is written as under.

$$O(k) = \sum_{k=1}^t \sum_{x \in \text{gridpoint}} S_{kx}^n |\widehat{T}(x) - \widehat{b}(x) - t_k|^2 + \frac{\alpha}{\#N_x} \sum_{k=1}^t \sum_{x \in \text{gridpoint}} S_{kx}^n \left(\sum_{r \in N_x} |\widehat{T}(r) - \widehat{b}(r) - t_k|^2 \right) \quad (4)$$

' t ' is the number of tissue classes, α is the neighbourhood influence and N_x is the number of neighbours, S_{kx} is the voxel X belonging to k^{th} tissue class. The parameters to be estimated for the minimization of $O(k)$ are the class centres $\{t_k\}$ and biasfield estimates $\{b_x\}$.

2) *Non parametric non uniform intensity normalization (N3):* Freesurfer scripts uses N3 method of inhomogeneity correction. N3 is a histogram based non uniform intensity correction method. If $S = (s_1, s_2, \dots, s_N)^T$ be intensities of N voxels of a MRI scan and $b = (b_1, b_2, \dots, b_N)^T$ are the corresponding bias field. The histogram of S will be blurred version of actual true image due to convolution of bias part b . The objective of this algorithm is to minimize this blurriness by de-convolution method using an iterative way to estimate a smooth bias model. The metric to be estimated is known as

$$CJV = \frac{\sigma_1 + \sigma_2}{|\mu_1 - \mu_2|} \quad (5)$$



Fig. 2. Data Preprocessing Pipeline.

where (μ_1, σ_1) and (μ_2, σ_2) are the mean and standard deviation of two different tissue types. This metric will be optimized if the standard deviation with in one class of tissues is minimum, hence the objective that one type of tissues should approximately should have same intensity values. It is done iteratively in particular value of bin $K = 200$, we try to estimate the CJV for the values.

$$\mu_1 = \min(S - b) \text{ to } \mu_k = \max(S - b) \quad (6)$$

C. Co-Registration

Registration is the most crucial stage of pre-processing because it helps to control the changes in data acquisition because of rotational transformational changes in brain position and even the size of brain may be different in different subjects. It helps to quantify the anatomical and morphometric alterations related to an individual (longitudinal studies) and a group of individual (both longitudinal as well as cross sectional studies). A common reference space or template is used to compare the source image and the template by applying optimal geometric transformations. The template can be the brain image of the same subject in case of longitudinal studies or common available templates.

D. Normalization

A technique to have uniform intensity distribution throughout the group of MRI images of a group to improve the accuracy of study using histogram equalization method.

E. Smoothing

It is a technique to remove unwanted noise from the MRI image which may result in incorrect results and affects accuracy of the study.

IV. PROPOSED METHOD

But during study we observed after applying Freesurfer scripts of registration, the slices of inter subjects does not

contain the similar information, means the slices of different subjects are not exactly parallel as shown in Fig 3, as our study is slice by slice study the Nth slice of X subject should contain almost same contents as the Nth slice of Y subject. Even the brain size of all subjects not same. We deduced a method to synchronize the inter subject slices. The steps of this method are listed below:

Mid_Slice_brainsize_Equalization_Method:

- Find the actual slice number of data acquisition, means first nonempty slice the actual start of MRI scan.
- Find the actual slice number of data acquisition ends, means first empty slice of MRI scan.
- Take the mid of first non-empty slice number and first empty slice of MRI scan., that is actual mid slice of each MRI scan, also calculate the length of scanning in each MRI scan, means Number of Nonempty slices in each MRI scan.
- From Mid Slice and actual size of brain which is actually the Number of Non empty slices we synchronize the Nth slice of X subject to the Nth slice of y subject as shown in Fig. 3.

A. SWMA Slice Wise Multivolume Analyse (SWMA) Design

Multivariate Approach considering Whole Brain Slices instead of Region of Interest (ROI). Earlier studies used ROI because of small sample size. As our sample set is sufficiently large so our study experimented with whole brain slices without compromising loss of information due to segmentation and approximation. Each MRI image is a volumetric representation which is flattened to 256 slices. In computation each slice is a two dimensional matrix of order 256X256. Slice Wise Multivolume Analysis described in Fig. 4.

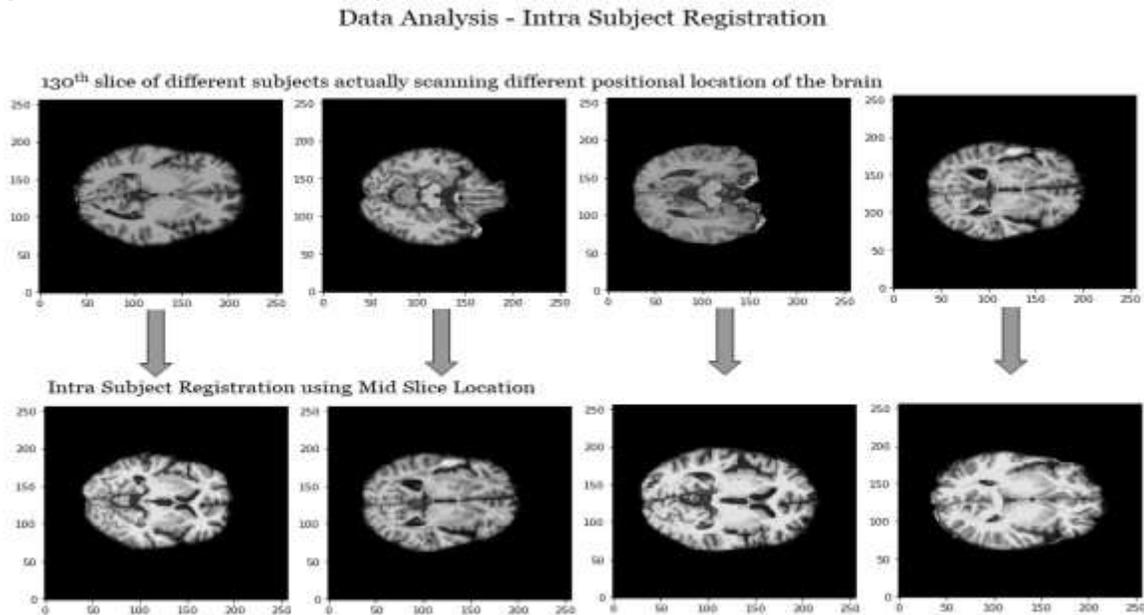


Fig. 3. Mid Slice Brain Size Equalization Method.

256 slices in each MRI Image for each subject

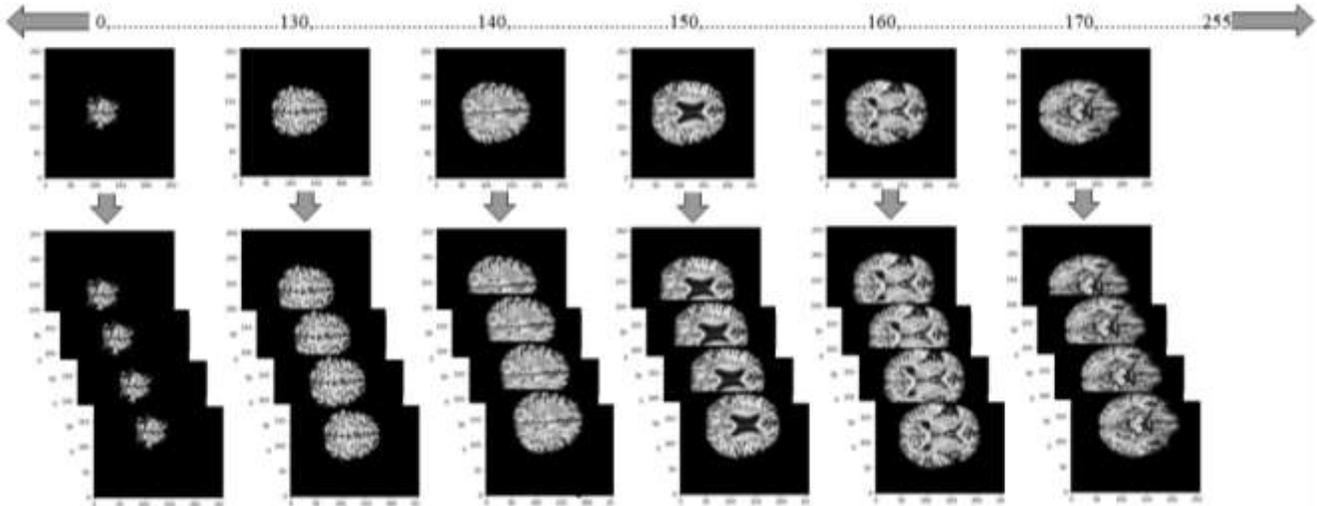


Fig. 4. Slice Wise Multivolume Analysis.

B. Feature Extraction

This study uses biomedical texture analysis for feature extraction. Texture analysis is a way of extracting image signatures pixel by pixel in terms of intensities, intra and inters pixel relationship and spectral properties. These can be calculated using mathematical statistical tools. Image analysis using this gives consistent, fast and accurate results. The features generated using texture based statistical distribution of pixel intensities give quantitative measures of image which are easily differentiable from each other hence helping image comparison easily. Each element of the matrix is the value of intensity at a particular pixel. We calculated the simple central tendencies statistics of these image slice matrices. These gross values are very much helpful in providing wide characteristics of image slice contents.

1) *Mean*: it gives a measure of concentration of data around the central distribution of data. But it is affected by extreme observations.

2) *Standard Deviation*: It is the measurement of how well the mean is able to represent the whole dataset. It gives the dispersion of the data.

3) *Skewness*: It is the measure of lack of symmetry. It helps us to determine the concentration of observation towards the higher and lower side of the observed data.

4) *Kurtosis*: It measures the convexity of the distribution curve.

These statistics give only intensity based information. These do not provide repeating nature of pixel values.

Gray Level CO-occurrence Matrix (GLCM) gives texture analysis of the image by measuring the spatial relationship among the pixels. At each pixel value we calculate a Gray Level co-occurrence matrix around it which calculates the number of pixels having the same pixel value. The GLCM matrix is calculated in four major directions. The directions are horizontal, vertical, diagonal up and diagonal down (at angles 0°, 90°, 45°, 135°, respectively).

Steps to create GLCM:

- Let x is the pixel under consideration.
- Let M is the set of pixels surrounding pixel x , which lie under the considered region M .
- Define each element mn of the GLCM as the number of times two pixels of intensity m and n occur in specified spatial relationship. Sum all the values with the specified intensity around that pixel x .
- To get symmetric GLCM make a transpose copy of GLCM and then add it to itself.
- Normalize the GLCM, divide each element by the sum of all elements.

If we have a slice of 256X256, GLCM will be too much data, we use some descriptive quantities from GLCM matrices. Each descriptor is calculated in four directions.

$$Energy = \sum_{m,n=0}^{N-1} (X_{mn})^2$$

$$Contrast = \sum_{m,n=0}^{N-1} X_{mn}(m-n)^2$$

$$Homogeneity = \sum_{m,n=0}^{N-1} \frac{X_{mn}}{1+(m-n)^2}$$

$$Correlation = \sum_{m,n=0}^{N-1} X_{mn} \frac{(m-\mu)(n-\mu)}{\sigma^2}$$

$$Entropy = \sum_{m,n=0}^{N-1} -\ln(X_{mn})X_{mn}$$

X_{mn} is the element of the normalized symmetrical GLCM

N is the number of gray levels

$$\mu = \sum_{m,n=0}^{N-1} iX_{mn} \sigma^2 = \sum_{m,n=0}^{N-1} X_{mn}(m - \mu)^2$$

Total Number of features from Texture analysis are 28. The most impotent and unique property of these statistical and GLCM features is that these are invariant to geometrical transformations of surfaces like translation horizontal or vertical, rotation, etc. The features should follow the rule of invariance. Features are volumetric signatures of microscopic structures of Brain: The most affected microstructures of the brain are hippocampus, amygdale and temporal horn. Studies show the volume of these structures decline with age but if the rate of change of the volumes over a certain time is more than normal change, it indicates some non-cognitive developments may cause brain diseases in future.

C. Feature Selection

Feature extraction and selection and classification share very thin line boundaries, a good feature extractor and selection technique surely makes the classification very easy and correct, but a good classifier would not need a good feature extractor or selection technique. As the features are the input to the classifiers so either we should have the best features so the classification should be with least error or the classification algorithm should be such that even the features provide least information but the algorithm is smart enough to extract the correct piece of information with least classification error.

Every classifier works on a discriminate function $F_{ci}(X)$, the classifier as described in Fig. 5 will assign a feature vector X to a said class c1 if $F_{ck}(X) > F_{cj}(X)$ for all $k <> j$.

Objective of this function is that create a boundary or hyper plane in feature space which distinguishes the n No of classes. The hyper plane can be represented with the equation

$$F_{ci}(X) = w_i^t X + w_{i0} \tag{7}$$

Where $w = \mu_j - \mu_k$

but the classifier function’s discriminability gets affected by decision bias degrading Classification accuracy and other scores. The variance σ is also biased. The means the variance of a sample feature is not as expected.

Theoretically when we extract features we hope that each feature help up to some extent to the discrimination function means all are independent but practically it’s not true many times. Table II shows discriminatory performance of basic statistical features in the concerned study and Table III shows the discriminatory performance of GLCM Features. The classification accuracy also depends on dimensionality. After applying a set of feature the accuracy performance may be inadequate we may think to add more no of features to improve the performance at the cost of computational cost but practically as we add the new features generally it increase the performance but up to some extent only after a point as we increase the features the performance decreases. Our study applied Fisher Linear Discriminant It is based on simple criteria if the mean of two sample space features differ than its

variance then it will definitely provide better discrimination ability to classify two sets of classes. The vector w in decision function is a scalar dot product with X as in equation vii, results into a vector the direction of this vector is important, not the magnitude. The FLD employs the linear function $W_i^t X$ such that

$$J(w) = \frac{|m_1 - m_2|^2}{\sigma_1^2 + \sigma_2^2} \tag{8}$$

Should be maximum where m_1 and m_2 are mean of the feature in two different classes and σ_1 and σ_2 are the standard deviation of features in two classes of the same feature. This is called Feature Discrimination Ratio (FDR). FDR is applied in each classifier, by keeping on adding the features if the classifier shows improved accuracy, if the accuracy or other scores decrease stop adding the features. By applying FDR on our extracted features we find that Mean, standard deviation, skewness, homogeineity in two directions and energy in all four directions are the best FDR values by adding other features the accuracy and specificity sensitivity decrease. But it’s not true in all the classifiers. The AdaBoost, Randomforest and Bagging Classifier based on ensemble techniques are more efficient classifiers and almost give similar accuracy with or without feature selection but SVM and K neighbours accuracy increase a lot after applying FDR.

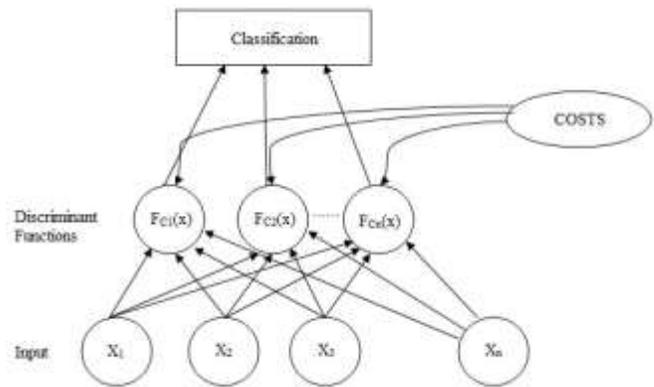


Fig. 5. A Generic Classifier.

TABLE II. BASIC STATISTICS SHOWING HIGH DISCRIMINARY PERFORMANCE

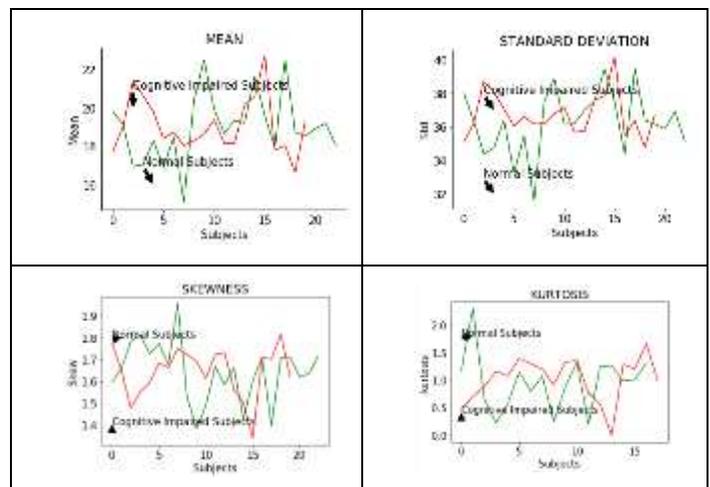
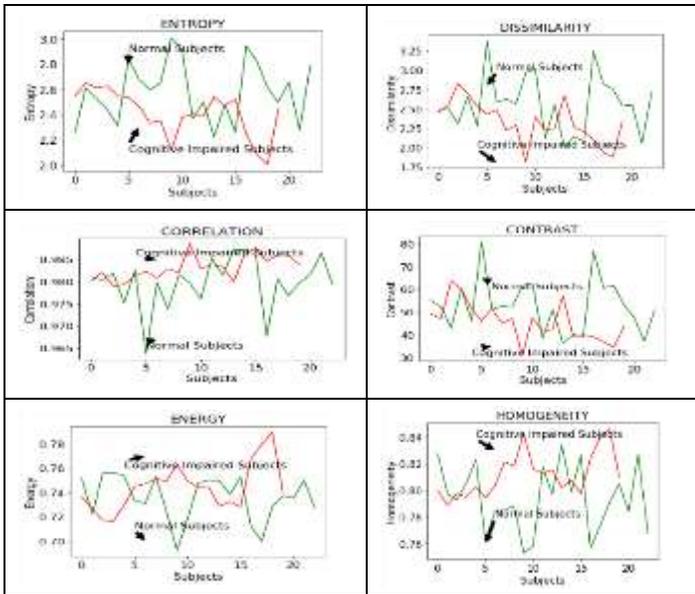


TABLE III. GLCM SHOWING HIGH DISCRIMINATORY PERFORMANCE



V. CLASSIFICATION

A. Support Vector Machine

As the objective of a classifier is to find a hyperplane which divides the sample space into desired set of classes with least error, SVM tries to find this hyperplane by processing the input data transferring into higher dimension plane using suitable kernel function so that sample data can be easily classified which cannot otherwise classified in lower dimension plane. The solution vector hyperplane may not be unique. The objective is to find the optimal hyperplane.

If L is the optimal hyperplane and two hyperplanes S and T passing through the nearest vectors in two classes from the optimal hyperplane. Then the distance between the optimal hyperplane L and S or L and T is called margin. The points on the hyperplane S and T are called support vectors, as shown in Fig. 6. These are the vectors which are the most informative for the classifier. The algorithm implements such that the controlling parameters are C and gamma and the kernel. Kernel is the function which converts the input features from lower dimensional plane to higher dimensional plane. C is a regularity parameter which changes the width of margin and gamma decides how much stringent is the classifier to the outliers. The training the data with SVM is that we want the hyperplane margin big enough to generalize the classifier. The C is the costing factor also, if C is large then it gives a large penalty and margin will be small but if C is small less penalty hence margin will be big. But the behavior change also depends on the particular size of sample set, the hyperparameter tuning results vary from model to model. The hyperparameter tuning do have limitations like, hyperparameters values change from dataset to datasets. The best parameters for one dataset may not work perfectly with other datasets. Moreover it is a time consuming process. But Data Processing and classification model evaluating scores really affected by hyperparameter tuning. It gives practical experience of algorithms. The classifier behaviour under various parameters gives an insight of its design. Fig. 7A

depicts the hyperparameter tuning C and Gamma to optimize accuracy, Fig. 7B depicts the hyperparameter tuning to optimize specificity and Fig. 7C depicts the hyperparameter tuning to optimize sensitivity.

1) SVM classification with full features: First the experimentation was carried out with full features, Table IV shows the results of GridsearchCV method, which internally applies 10 fold cross validation under a given set of parameters. The best value of accuracy is 92.95% with specificity 84.22% and sensitivity 79.28%. The results are again checked with 10 fold cross validation with hold out data, the results are comparable with receiving operating curve area showing accuracy as shown in Fig. 8.

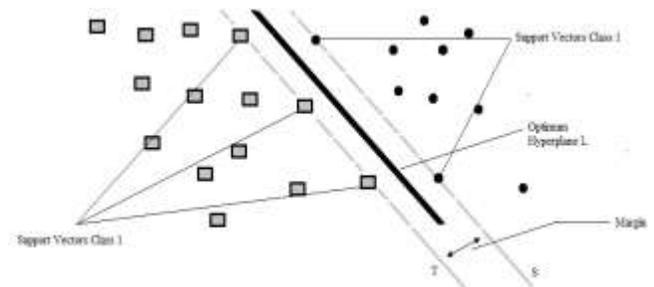
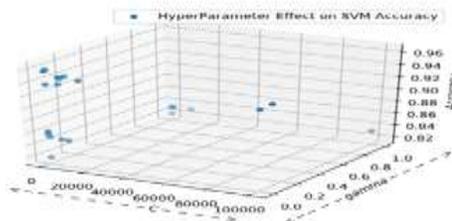
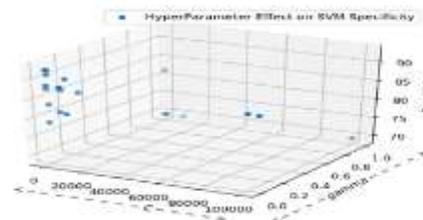


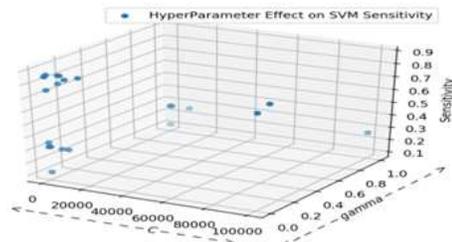
Fig. 6. SVM Hyperplane.



(a) Effect of C and Gamma in SVM Classifier Accuracy



(b) Effect of C and Gamma on SVM Classifier Specificity



(c) Effect of C and Gamma on SVM Classifier Sensitivity

Fig. 7. (a) Hyper Parameter Tuning to Optimize SVM Accuracy, (b) Hyper Parameter Tuning to Optimize SVM Specificity, (c) Hyper Parameter Tuning to Optimize Sensitivity.

2) *SVM Classification with FDR Selected Features:* The Table V are results of Gridsearch CV exploring SVM under varying C and gamma, using a subset of features after applying FDR. The highest value of Accuracy is 96.09% with specificity 92.63% and sensitivity 87.21%. The results are again checked with 10 fold cross validation with hold out data, the results are comparable with receiving operating curve area showing accuracy as shown in Fig. 9.

B. Random Forest

The Random Forest algorithm is a meta-process which internally works on N no of decision trees to keep the information. Unlike decision tree the result is based on a multiple decision trees, here the algorithm based on divide and conquer approach means it divides the samples among N no of decision trees randomly and then enumerates the decision of all these trees to give the final result. Its way of taking advice of N experts rather than single. It's an ensemble approach hence time consuming but because today the technology is advanced to handle parallel processing so mean time to fit is not that important criteria to evaluate a classifier. One more important thing the study observed, Feature selection process does not much affect accuracy as Random forest itself chooses both sample divides as well as feature vector divides. The results with FDR or without FDR are almost the same. The Random Forest classifier is a very stable classifier which the study found during the GridsearchCV method. The Accuracy range does not change much even after tuning hyper parameters.

1) *Randomforest classification with full features:* Table VI are results of GridsearchCV with all features, the best accuracy is 89.98% with specificity 88.23 and sensitivity 56.39%. The results are again cross validated with hold out data and compared with receiving operating accuracy as shown in Fig. 10.

2) *Randomforest classification with FDR selected features:* The random forest hyperparameters tuning after applying FDR, results are listed in Table VII, with maximum accuracy 90.6% with specificity 87.13% and sensitivity 61.55% with criterion entropy max_depth None and No of estimators 100. The results are cross validated on hold out data and results are comparable for receiving operating area accuracy using 10 fold cross_validation algorithm shown in Fig. 11.

C. AdaBoost

Boosting is a process which is designed to deal with the problem of weak learning classifiers. Weak learning results in higher detection errors and low decision accuracy of the classifier. Weak classifiers are the moderate classifiers which give a bit better insight of the problem than random guesses. AdaBoost is a classifier which deals with a set of weak classifiers iteratively. Logic of using same weak classifiers on same data does not lead to a better results, but AdaBoost is designed in such a way that during each iteration the weak classifiers work with subsets of data, not full data as whole, these subsets of data may give different results with weak classifiers, initially all the classifiers are assigned equal

weights, but after each iteration the classifiers are judged on the basis of classification error, the classifiers with less error is given higher weight. AdaBoost is a kind of greedy algorithm with the objective of minimizing the classification error by improving the learning model after each iteration. AdaBoost is an adaptive boosting algorithm because it has no error bound and no bounds on the number of weak classifiers.

1) *AdaBoost classification with full features:* The AdaBoost algorithm works better with full features. Table VIII shows results of AdaBoost with all parameters GridsearchCV results with maximum average accuracy 96.76% with specificity 95.87% and sensitivity 87.37% using learning rate 1 and No of estimators 150. AdaBoost wins over all the classification method. The results are cross validated on hold out data using ROC curves shown in Fig. 12.

2) *AdaBoost Classification with FDR Slected Features:* The FDR degrades the accuracy of AdaBoost. Table IX shows AdaBoost with Gridsearch CV results With 10 features the best accuracy is 91.6% with specificity 86.15% and sensitivity 68.59% using no of estimators 150, learning rate 1. The results are cross validated on hold out data using ROC curves shown in Fig. 13.

D. Bagging Classifier

It is also an ensemble technique classifier very similar to random forest classifier, as in such classifiers the subsets of samples are randomly chosen in random forest, in which the previously selected samples are replaced with new samples. This is also used to improve the accuracy and other performances of decision tree classifiers.

1) *Bagging classification with full features:* Gridsearch CV results for different parameters are tabulated in Table X. The best accuracy is 86.86% with specificity 87.25% and sensitivity 38.95% which is using maximum samples selected from the bag are 200 and No of estimators 200, which are cross verified using hold out data using receiving operating curve accuracy as shown in Fig. 14.

2) *Bagging classification with FDR slected features:* Table XI lists the results of GridsearchCV using FDR selected features the accuracy is 86.1% with accuracy sensitivity 38.95 and specificity 85.9%, the results are cross verified on hold out data using Receiving Operating Curve accuracy as shown in Fig. 15.

E. Nearest Neighbours

KNN is a non parametric classifier, it is a lazy algorithm but very simple. Like to predict a vector X, it will look k Vectors which are nearest to X, the distance is generally calculated using Euclidean or Manhattan metrics which measure the distance between two observations X_s and X_t for j features.

$$\sqrt{\sum_{j=1}^p (x_{sj} - x_{tj})^2} \text{ Euclidean Distance}$$

$$\sum_{j=1}^p |x_{sj} - x_{tj}| \text{ Manhattan Distance}$$

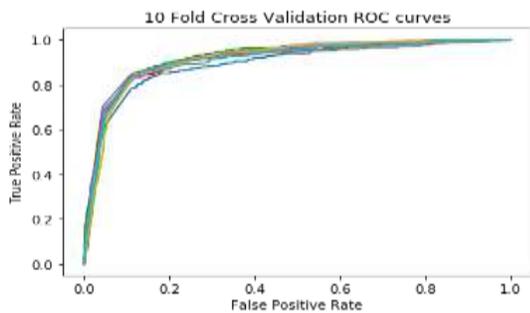
SVM Parameter Tuning

TABLE IV. SVM GRIDSEARCH CV: BEST RESULTS WITH C=1, GAMMA = 10, ACCURACY 92.9% WITH SPECIFICITY 84.22% AND SENSITIVITY 79.28% WITH ALL FEATURES

param_C	Param gamma	mean Specificity	mean_Sensitivity	mean_Accuracy
1	0.01	83.344442332466	0.2578847438557	0.8432073329718
1	0.1	83.853987796136	0.7178562638444	0.9168477040299
1	1	83.607909582329	0.7727646707218	0.9250390969494
1	10	84.223396084093	0.7928993354664	0.9295074469075
1	100	86.272333827734	0.6822210892725	0.9156894483427
10	0.001	80.241227828115	0.2457156921345	0.8390699034284
10	0.01	82.878571005636	0.7677437502197	0.9226396671760
10	0.1	83.314746871451	0.7845188284518	0.9262807503327
10	100	80.511885530581	0.7245385183362	0.9108060955636
100	0.0001	74.584867980930	0.1567930100910	0.8226860220531
100	0.001	81.966054078167	0.7610386413979	0.9195780542180
100	0.01	82.705260558913	0.7882880348792	0.9254527988212
100	100	78.643503364618	0.7148939910692	0.9051800376864
1000	1E-05	66.975494228749	0.1136176646390	0.8133371266275
1000	0.0001	80.486585304024	0.7350462360676	0.9122973257960
1000	0.001	82.692441319188	0.7891318870644	0.9255359226332
1000	100	78.568888719043	0.7148939910692	0.9050144747725
10000	1E-06	65.472364610713	0.1224359199746	0.8132545505836
10000	1E-05	77.250521741845	0.6331598748285	0.8905360047984
10000	0.0001	81.697496074876	0.7861836784923	0.9228044769088
10000	10	81.119868761615	0.8155233641573	0.9257005954239
10000	100	78.568888719043	0.7148939910692	0.9050144747725
100000	1E-06	72.383866822978	0.4713178158292	0.8597560925509
100000	1E-05	80.296232630667	0.7656569740867	0.9164344129842
100000	10	80.506045797683	0.8079726451249	0.9231350549685
100000	100	78.568888719043	0.7148939910692	0.9050144747725

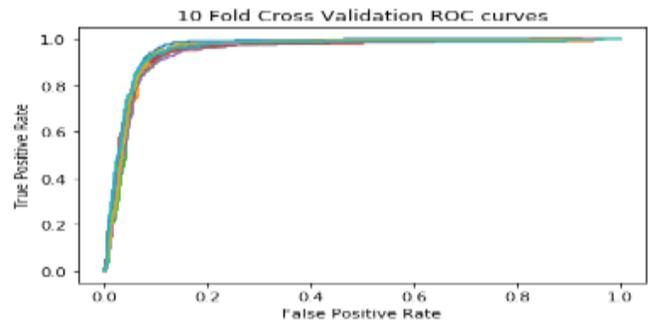
TABLE V. SVM GRIDSEARCH: BEST RESULTS WITH C 1000, GAMMA = 0.0001, ACCURACY 96.09%, SPECIFICITY 92.63% AND SENSITIVITY 87.21% AFTER APPLYING FDR FEATURE SELECTION

C	Gamma	Specificity	Sensitivity	Accuracy
1	0.01	78.829923692182	0.11613339896628	0.81879816935894
1	0.1	80.532799121772	0.26162054780071	0.84146940167288
1	1	81.817542579539	0.11196863682711	0.81979236849456
10	0.001	88.608313139021	0.32035793396856	0.85760425396721
10	0.01	91.338500106760	0.75266692451039	0.93703631428743
10	0.1	86.861630821078	0.76898843219296	0.93124339254706
10	1	69.457820713316	0.26581343834605	0.83137526772166
100	0.0001	87.785798941696	0.35392039661052	0.86256915572499
100	0.001	92.636598407176	0.85619703948525	0.95805075345504
100	0.1	87.117762600668	0.8310273900355	0.942247780169698
100	1	69.4578207133165	0.2658134383460	0.831375267721668
1000	1E-05	83.781206927729	0.32413768854822	0.85396365010763
1000	0.0001	92.634766545677	0.8721229914559	0.96094570796290
1000	0.1	86.32964031705	0.8314387679758	0.9405924933857
1000	1	69.457820713316	0.2658134383460	0.83137526772166
10000	1E-06	81.859571838812	0.3203614500193	0.85156380950815
10000	1E-05	90.934324474569	0.8540979571745	0.95424397044243
10000	0.1	86.32964031705	0.8314387679758	0.9405924933857
10000	1	69.457820713316	0.2658134383460	0.83137526772166
100000	1E-06	88.029537238048	0.8021166625646	0.93918527708850
100000	0.1	86.32964031705	0.83143876797581	0.9405924933857
100000	1	69.457820713316	0.26581343834605	0.83137526772166



Receiving operating Curve Accuracy: 0.8937792926314483
 Receiving operating Curve Accuracy: 0.9179476564187485
 Receiving operating Curve Accuracy: 0.9226346488033448
 Receiving operating Curve Accuracy: 0.9157707603773151
 Receiving operating Curve Accuracy: 0.9162423970273819
 Receiving operating Curve Accuracy: 0.9149871321207529
 Receiving operating Curve Accuracy: 0.9027465236824136
 Receiving operating Curve Accuracy: 0.913242903607333
 Receiving operating Curve Accuracy: 0.912467964860967
 Receiving operating Curve Accuracy: 0.9134321596900

Fig. 8. The Cross Validation of Table IV Results with Receiving Operating Curve with Hold out Data with SVM and All Parameters.



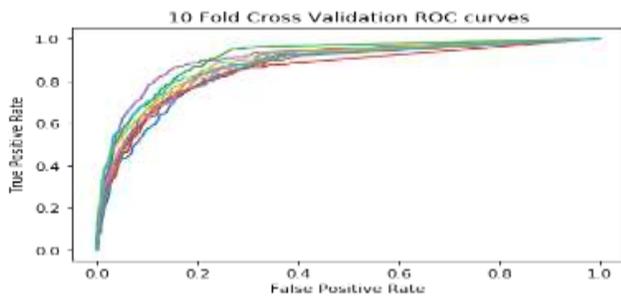
Receiver operating Curve accuracy:0.9617590965184548
 Receiver operating Curve accuracy: 0.942635941021788
 Receiver operating Curve accuracy: 0.949475534894809
 Receiver operating Curve accuracy: 0.947978539771402
 Receiver operating Curve accuracy: 0.9452122348088005
 Receiver operating Curve accuracy: 0.9480958157900996
 Receiver operating Curve accuracy: 0.9542638074464368
 Receiver operating Curve accuracy: 0.9523513244133015
 Receiver operating Curve accuracy: 0.9523604360217566
 Receiver operating Curve accuracy: 0.9549160237865637

Fig. 9. The Cross Validation of Table V Results with Receiving Operating Curve with Hold Out Data with SVM with Selected Features.

Random Forest Parameter Tuning

TABLE VI. RANDOM FOREST GRIDSEARCHCV RESULTS MAXIMUM ACCURACY IS 89.98% WITH SPECIFICITY 88.23% AND SENSITIVITY 56.39% USING CRITERION ENTROPY AND MAX_DEPTH NONE AND N ESTIMATORS 100

criterion	depth	estimator	specificity	sensitivity	Accuracy
gini	5	20	84.888824121	0.230188679245	0.839400708997
gini	5	30	85.850144850	0.234800838574	0.840972734829
gini	5	50	84.784164752	0.231027253668	0.839648984840
gini	5	100	84.905496091	0.218867924528	0.837828477838
gini	15	20	86.667233743	0.522012578616	0.889706767578
gini	15	30	87.691345970	0.540880503144	0.894257419101
gini	15	50	88.032210687	0.539203354297	0.894505626502
gini	15	100	88.341393086	0.543815513626	0.895664190068
gini	20	20	86.353249455	0.516142557651	0.888217694282
gini	20	30	87.101982059	0.529140461215	0.891527411466
gini	20	50	87.815480695	0.540880503144	0.894423118844
gini	20	100	88.034856531	0.538364779874	0.894340440079
gini	25	20	86.502917560	0.517400419287	0.888631430318
gini	25	30	87.233662182	0.529559748427	0.89177565308
gini	25	50	87.989689597	0.540880503144	0.894671360466
gini	25	100	88.215312778	0.538364779874	0.894588613258
gini	None	20	86.502917560	0.517400419287	0.888631430318
gini	None	30	87.233662182	0.529559748427	0.891775653088
gini	None	50	87.989689597	0.540880503144	0.894671360466
gini	None	100	88.215312778	0.538364779874	0.894588613258
entropy	5	20	86.565016992	0.179454926624	0.832533409441
entropy	5	30	85.625514909	0.198322851153	0.835015688774
entropy	5	50	85.087555336	0.198322851153	0.834850160138
entropy	5	100	85.445716966	0.189517819706	0.833443389172
entropy	15	30	88.583992309	0.566037735849	0.899966497305
entropy	25	20	88.823140395	0.563941299790	0.899883715876
entropy	None	20	88.823140395	0.563941299790	0.899883715876

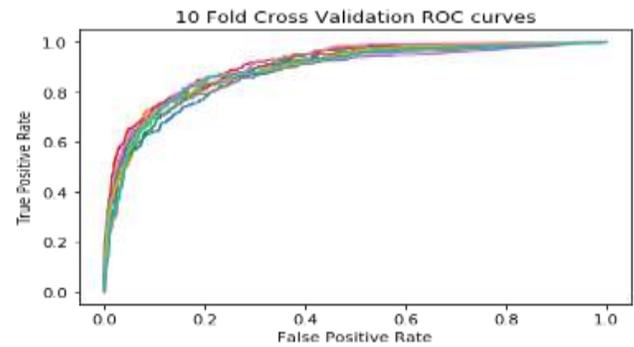


Receiving Operating Curve accuracy: 0.8635070069526353
 Receiving Operating Curve accuracy: 0.8768584331740022
 Receiving Operating Curve accuracy: 0.9081541453199207
 Receiving Operating Curve accuracy: 0.8475329655992669
 Receiving Operating Curve accuracy: 0.9046812480763313
 Receiving Operating Curve accuracy: 0.8626087454212453
 Receiving Operating Curve accuracy: 0.8753501140194329
 Receiving Operating Curve accuracy: 0.8639699552341596
 Receiving Operating Curve accuracy: 0.8887315511961502
 Receiving Operating Curve accuracy: 0.8852172129377727

Fig. 10. Random Forest the Cross Validation of Table VI Results with Hold out Data Results Comparable with ROC Area Accuracy.

TABLE VII. RANDOM FOREST CLASSIFIER WITH FDR FEATURES USING GRIDSEARCHCV, BEST ACCURACY 90.6% WITH SPECIFICITY 87.13% AND SENSITIVITY 61.55% USING ENTROPY CRITERION MAX_DEPTH NONE AND NO OF ESTIMATORS 100

criterion	depth	estimators	Specificity	Sensitivity	Accuracy
gini	5	20	82.405557662	0.281761006289	0.846102822131
gini	5	30	81.753935627	0.268343815513	0.843537761369
gini	5	50	82.324952592	0.278825995807	0.845440810251
gini	5	100	83.218901214	0.272955974842	0.84527555538
gini	15	20	85.672104091	0.594549266247	0.900131854834
gini	15	30	86.027370483	0.596226415094	0.900959224243
gini	15	50	86.273793287	0.592452830188	0.900793866714
gini	15	100	86.892389272	0.600838574423	0.903110377861
gini	20	20	85.894661091	0.596226415094	0.900793661385
gini	20	30	86.037359156	0.598742138364	0.901373234050
gini	20	50	86.312500959	0.594549266247	0.901207568529
gini	20	100	86.680282575	0.598742138364	0.90244829753
gini	25	20	85.894661091	0.596226415094	0.900793661385
gini	25	30	86.037359156	0.598742138364	0.901373234050
gini	25	50	86.312500959	0.594549266247	0.901207568529
gini	25	100	86.688653578	0.599161425576	0.902531044746
gini	None	20	85.894661091	0.596226415094	0.900793661385
gini	None	30	86.037359156	0.598742138364	0.901373234050
gini	None	50	86.312500959	0.594549266247	0.901207568529
gini	None	100	86.688653578	0.599161425576	0.902531044746
entropy	5	20	86.4218333770	0.220125786163	0.838656326345
entropy	5	30	84.4299042596	0.218448637316	0.837332644799
entropy	5	50	83.3893286936	0.222641509433	0.837415084015
entropy	5	100	82.6401550351	0.244444444444	0.840228009963
entropy	15	20	86.3301678223	0.599580712788	0.90195201962
entropy	15	30	86.7762836186	0.60377358490	0.90344116136
entropy	15	50	87.1423000822	0.607547169811	0.904682369472
entropy	15	100	87.3657096503	0.615094339622	0.906336766076
entropy	20	20	86.1350772246	0.603773584905	0.90236568721
entropy	20	30	86.4527911908	0.600838574423	0.902448160654
entropy	20	50	86.9335626954	0.607547169811	0.904351483307
entropy	20	100	87.1309161889	0.615513626834	0.906006051017
entropy	25	20	86.1350772246	0.603773584905	0.902365687217
entropy	25	30	86.4527911908	0.600838574423	0.902448160654
entropy	25	50	86.9335626954	0.607547169811	0.904351483307
entropy	25	100	87.1309161889	0.615513626834	0.906006051017
entropy	None	20	86.1350772246	0.603773584905	0.902365687217
entropy	None	30	86.4527911908	0.600838574423	0.902448160654
entropy	None	50	86.9335626954	0.607547169811	0.904351483307
entropy	None	100	87.1309161889	0.615513626834	0.906006051017



Area under the ROC curve: 0.8772840343735866
 Area under the ROC curve: 0.889214912760619
 Area under the ROC curve: 0.8947546991251121
 Area under the ROC curve: 0.9128083521162034
 Area under the ROC curve: 0.883718402186543
 Area under the ROC curve: 0.8945758258258258
 Area under the ROC curve: 0.9118374548334127
 Area under the ROC curve: 0.8821687953919359
 Area under the ROC curve: 0.8938542616531675
 Area under the ROC curve: 0.8887535609191084

Fig. 11. GridsearchCV Results of Random Forest Verified with Hold Out Data Results Verification with FDR Features.

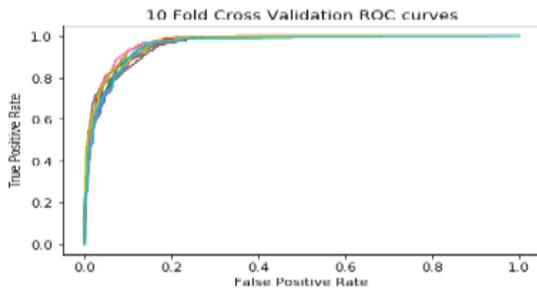
AdaBoost Parameter Tuning

TABLE VIII. ADABOOST WITH ALL PARAMETERS GRIDSEARCHCV RESULTS WITH MAXIMUM AVERAGE ACCURACY 96.76% WITH SPECIFICITY 95.87% AND SENSITIVITY 87.37% USING LEARNING RATE 1 AND NO OF ESTIMATORS 150

Learning rate	Estimators	Specificity	Sensitivity	Accuracy
0.001	20	100	0	0.80266
0.001	30	100	0	0.80266
0.001	50	100	0	0.80266
0.001	100	100	0	0.80266
0.001	150	100	0	0.80266
0.005	20	100	0	0.80266
0.005	30	100	0	0.80266
0.005	50	100	0	0.80266
0.005	100	100	0	0.80266
0.005	150	100	0	0.80266
0.01	20	100	0	0.80266
0.01	30	100	0	0.80266
0.01	50	100	0	0.80266
0.01	100	100	0	0.80266
0.01	150	100	0	0.80266
0.2	20	88.73902	0.05073	0.81094
0.2	30	79.53308	0.12075	0.81996
0.2	50	77.35336	0.19078	0.82914
0.2	100	78.42884	0.24780	0.83791
0.2	150	80.73031	0.29392	0.84660
0.3	20	76.08379	0.12872	0.81996
0.3	30	75.20920	0.18449	0.82682
0.3	50	76.16444	0.23438	0.83419
0.3	100	79.68608	0.31614	0.84883
0.3	150	84.43293	0.40084	0.86704
0.5	20	75.60548	0.19706	0.82873
0.5	30	75.98205	0.24235	0.83493
0.5	50	79.10773	0.30818	0.84726
0.5	100	86.60799	0.46289	0.87970
0.5	150	90.92653	0.58742	0.90692
0.9	20	74.20717	0.30734	0.84147
0.9	30	78.71854	0.36520	0.85454
0.9	50	86.59273	0.49853	0.88574
0.9	100	91.11120	0.70021	0.92727
0.9	150	93.94760	0.81426	0.95284
1	20	75.97621	0.34214	0.84759
1	30	81.68857	0.40545	0.86249
1	50	89.30735	0.53627	0.89566
1	100	91.18001	0.76394	0.93861
1	150	95.87176	0.87379	0.96765

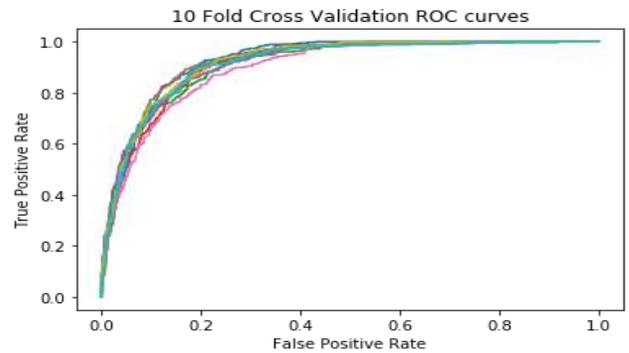
TABLE IX. ADABOOST WITH GRIDSEARCHCV RESULTS WITH 10 FEATURES THE BEST ACCURACY IS 91.6% WITH SPECIFICITY 86.15% AND SENSITIVITY 68.59% USING NO OF ESTIMATORS 150, LEARNING RATE 1

Learning rate	Estimators	Specificity	Sensitivity	Accuracy
0.001	20	100	0	0.80266
0.001	30	100	0	0.80266
0.001	50	100	0	0.80266
0.001	100	100	0	0.80266
0.001	150	100	0	0.80266
0.005	20	100	0	0.80266
0.005	30	100	0	0.80266
0.005	50	100	0	0.80266
0.005	100	100	0	0.80266
0.005	150	100	0	0.80266
0.01	20	100	0	0.80266
0.01	30	100	0	0.80266
0.01	50	100	0	0.80266
0.01	100	100	0	0.80266
0.01	150	100	0	0.80266
0.2	20	87.06042	0.05912	0.81193
0.2	30	76.39592	0.14046	0.82169
0.2	50	74.62386	0.19161	0.82749
0.2	100	75.31635	0.24570	0.83485
0.2	150	77.70321	0.29811	0.84420
0.3	20	75.59483	0.13962	0.82054
0.3	30	73.17302	0.20377	0.82790
0.3	50	75.07925	0.22683	0.83220
0.3	100	77.68116	0.30650	0.84528
0.3	150	79.09073	0.36646	0.85562
0.5	20	73.78344	0.20042	0.82798
0.5	30	75.12703	0.24277	0.83452
0.5	50	76.43679	0.28889	0.84180
0.5	100	80.70590	0.41216	0.86422
0.5	150	83.68746	0.48428	0.87936
0.9	20	74.64103	0.31572	0.84321
0.9	30	78.62295	0.36394	0.85396
0.9	50	80.12446	0.45241	0.86952
0.9	100	85.14953	0.58239	0.89740
0.9	150	85.99633	0.66373	0.91221
1	20	74.37228	0.32788	0.84478
1	30	79.05477	0.38365	0.85793
1	50	80.45211	0.48344	0.87457
1	100	85.07785	0.62558	0.90435
1	150	86.15287	0.68595	0.91618



Receiving Operating Curve Accuracy: 96.16587687161517
 Receiving Operating Curve Accuracy: 96.5823773693516
 Receiving Operating Curve Accuracy: 96.0141773646603
 Receiving Operating Curve Accuracy: 96.67279696025804
 Receiving Operating Curve Accuracy: 95.86224658961727
 Receiving Operating Curve Accuracy: 95.83158385817588
 Receiving Operating Curve Accuracy: 96.7128517189369
 Receiving Operating Curve Accuracy: 96.66772665818672
 Receiving Operating Curve Accuracy: 96.71101941785082
 Receiving Operating Curve Accuracy: 95.79305816277098
 Average Accuracy: 96.30137149714237

Fig. 12. GridsearchCV Results of adaBoost of Table VIII Verified with Hold out Data Results Verification with full Features with Average Accuracy 96.3.



Receiving Operating Curve Accuracy: 92.24488989792022
 Receiving Operating Curve Accuracy: 91.67784243641628
 Receiving Operating Curve Accuracy: 90.8119193588127
 Receiving Operating Curve Accuracy: 91.23084331888616
 Receiving Operating Curve Accuracy: 91.548607052406
 Receiving Operating Curve Accuracy: 92.37635017691973
 Receiving Operating Curve Accuracy: 89.59660719974514
 Receiving Operating Curve Accuracy: 92.29814330924668
 Receiving Operating Curve Accuracy: 92.1426847303852
 Receiving Operating Curve Accuracy: 91.57330098242107
 Average Accuracy: 91.55011884631591

Fig. 13. GridsearchCV Results of Table IX Cross Validated on Hold out Data Average Accuracy 91.55%.

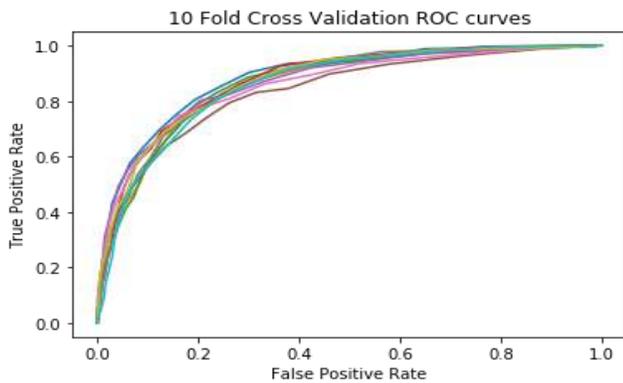
Bagging Classifier Parameter Tuning

TABLE X. BAGGING CLASSIFIER GRIDSEARCHCV RESULTS MAXIMUM ACCURACY 86.86% WITH SPECIFICITY 87.35% AND SENSITIVITY 39.16 USING MAX_SAMPLE FROM BAG 200 AND ESTIMATORS 50 WITH ALL FEATURES

max sam	estimator s	specificity	Sensitivity	accuracy
5	20	61.9225812999993	0.163941299790356	0.777922819236364
5	30	76.6148382202139	0.038993710691824	0.802995223042736
5	50	88.641975308642	0.014255765199162	0.801671267726228
5	100	100	0	0.802664234213609
5	200	100	0	0.802664234213609
10	20	57.3384277875513	0.075471698113208	0.799933507930625
10	30	72.9626890756303	0.049056603773585	0.803988189530117
10	50	84.7058823529412	0.015094339622642	0.803739776801557
10	100	82.1759259259259	0.015513626834382	0.804319144137901
10	200	79.1273054430949	0.035220125786164	0.806636031720447
20	20	54.7680693719426	0.153878406708595	0.804814737266674
20	30	56.1579806137808	0.158909853249476	0.805724751578716
20	50	62.5900805494047	0.121174004192872	0.809365902469857
20	100	61.3947033358798	0.130398322851153	0.810607247464455
20	200	62.4301942004478	0.12746312368973	0.811600453501237
30	20	58.7801119010948	0.20335429769392	0.811764715947454
30	30	64.798210923243	0.183228511530398	0.817226065849393
30	50	68.8224646380291	0.146750524109015	0.817805090972307
30	100	65.4505831704934	0.170649895178197	0.815654211124469
30	200	66.3242335725634	0.161844863731656	0.816647143390507
50	20	64.2439406993273	0.187421383647799	0.819293719391148
50	30	68.5296108891203	0.165199161425577	0.819873839597037
50	50	74.0037009529775	0.167295597484277	0.82376261612585
50	100	71.4577484693443	0.19874213836478	0.825252100078264
50	200	74.6157054563379	0.19832285115304	0.827072470195777
100	20	70.126649647095	0.228092243186583	0.828561577713419
100	30	71.4837866188963	0.238574423480084	0.83063046322352
100	50	75.0823274419283	0.259538784067086	0.836587988377063
100	100	79.7148769872991	0.252830188679245	0.839731184506766
100	200	81.2326043291691	0.258700209643606	0.841800001574182
200	20	85.3025104450034	0.368553459119497	0.862898554593137
200	30	86.8260797310868	0.390775681341719	0.868111286438458
200	50	87.3511737282356	0.391614255765199	0.868608111895578
200	100	87.2558236365198	0.389517819706499	0.868194273195141
200	200	87.8557218134902	0.381551362683438	0.867532158651515

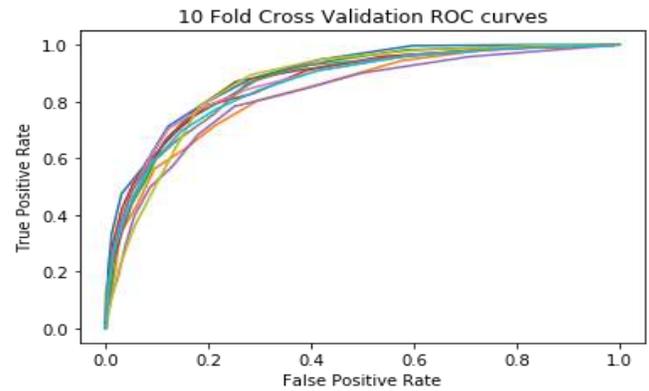
TABLE XI. BAGGING CLASSIFIER GRIDSEARCHCV RESULTS MAXIMUM ACCURACY 86.1% WITH SPECIFICITY 85.90% AND SENSITIVITY 35.72 USING MAX_SAMPLE FROM BAG 200 AND ESTIMATORS 100 WITH FDR FEATURES

max sam	estimators	specificity	Sensitivity	accuracy
5	20	76.9869281045752	0.023480083857442	0.802085790853525
5	30	80.952380952381	0.005031446540881	0.802664405320324
5	50	100	0	0.802664234213609
5	100	100	0	0.802664234213609
5	200	100	0	0.802664234213609
10	20	57.5746807492888	0.108176100628931	0.800346114662716
10	30	52.7995652542967	0.090146750524109	0.802498021150843
10	50	65.9676052770524	0.059538784067086	0.805394447176448
10	100	66.8473163105784	0.092662473794549	0.806221987691952
10	200	68.5488230149631	0.089727463312369	0.808621554039094
20	20	60.6883730203888	0.09601677148847	0.808455101426927
20	30	67.2984063572946	0.09727463312369	0.81019286122253
20	50	75.5615700089384	0.105241090146751	0.813088637042619
20	100	68.6563164386452	0.124947589098533	0.813585325614367
20	200	73.2714047214799	0.116981132075472	0.815074912230811
30	20	67.4581412473286	0.153039832285115	0.815736171240862
30	30	67.1873053086082	0.167295597484277	0.817060674098858
30	50	71.3315605294971	0.148846960167715	0.816729582605702
30	100	70.0808632215366	0.179874213836478	0.81896365453828
30	200	72.0420630600271	0.174842767295598	0.819956826353719
50	20	65.7035032574865	0.183228511530398	0.820700695685681
50	30	71.6890309539053	0.165199161425577	0.823100878016996
50	50	77.848966366415	0.135010482180293	0.821115321477007
50	100	76.7065896952809	0.178197064989518	0.825500512806824
50	200	76.6562019744781	0.184067085953878	0.825417834042228
100	20	73.3607674212325	0.238993710691824	0.832533717433687
100	30	74.5578503120475	0.241090146750524	0.833857604307509
100	50	79.7696255447569	0.212159329140461	0.833608780922833
100	100	80.6639735923208	0.242767295597484	0.838408187387861
100	200	79.3784316065887	0.256603773584906	0.839814376591506
200	20	84.2752589651809	0.341719077568134	0.857438984201032
200	30	83.8258450469914	0.321174004192872	0.853632304673985
200	50	84.6664830924616	0.29643605870021	0.850488081903993
200	100	85.9051302031996	0.357232704402516	0.861492502274864
200	200	84.017160828805	0.358490566037736	0.859755016250004



Receiving Operating Curve Accuracy: 88.6077212947019
 Receiving Operating Curve Accuracy: 86.13601530743381
 Receiving Operating Curve Accuracy: 86.47772069666797
 Receiving Operating Curve Accuracy: 87.59216258055226
 Receiving Operating Curve Accuracy: 86.9632627583638
 Receiving Operating Curve Accuracy: 83.67623048741638
 Receiving Operating Curve Accuracy: 85.90495419479267
 Receiving Operating Curve Accuracy: 86.5789072039072
 Receiving Operating Curve Accuracy: 87.16157031374424
 Receiving Operating Curve Accuracy: 85.93123904332582
 Average accuracy: 86.50297838809061

Fig. 14. Bagging Classifier GridSearchCV Results of Table X Verified using ROC on Holdout Data, Average Accuracy 86.5%.



Receiving Operating Curve Accuracy: 89.2582028251113
 Receiving Operating Curve Accuracy: 83.30865172606707
 Receiving Operating Curve Accuracy: 87.38146156666258
 Receiving Operating Curve Accuracy: 86.89730009557185
 Receiving Operating Curve Accuracy: 82.10881903855447
 Receiving Operating Curve Accuracy: 87.17537108726057
 Receiving Operating Curve Accuracy: 86.77546994821599
 Receiving Operating Curve Accuracy: 86.68428919178224
 Receiving Operating Curve Accuracy: 86.57772635034999
 Receiving Operating Curve Accuracy: 85.90527854724532
 Average accuracy: 86.20725703768213

Fig. 15. Bagging Classifier GridSearchCV Results of Table XI Verified using ROC on Holdout Data, Average Accuracy 86.2%.

First do the prediction for k nearest point, the predict of X point will be 1 if most of k nearest points predict as 1 otherwise -1. The k generally is odd.

1) *KNN classification with full features*: The Gridsearch results of KNN with Full features listed in Table XII maximum accuracy 82.65 % with specificity 60.01% and sensitivity 36.85%, same is verified using hold out data as shown in Fig. 16, with K equals to 5.

2) *KNN Classification with FDR Selected Features*: The accuracy is increased noticeably using FDR, the results are listed in Table XIII showing maximum accuracy 91.5% with specificity 81.54% and sensitivity 74.04% with K equal to 5. The results of Table XIII are verified in Fig. 17 using hold out data using ROC curve.

F. Gaussian Naive Bayes

It is a probability based classifier that works on Bayes theorem that states the outcome of an event can be measured from the past probability of events. It's a non parametric algorithm. As there are no major parameters to vary so GridsearchCV testing is not done for Naive Bays.

1) *Naive bayes classification with full features*: Naive Bayes results average accuracy 71.23614190687361specificity 85.95% sensitivity 32.78%.The results are cross validated with ROC accuracy on hold out data as shown in Table XIV.

2) *Naive Bayes Classification with FDR Selected Features*: FDR helped to improve average accuracy 74.86 specificity 86% sensitivity 37%. The results are cross validated with ROC accuracy on hold out data as shown in Table XV.

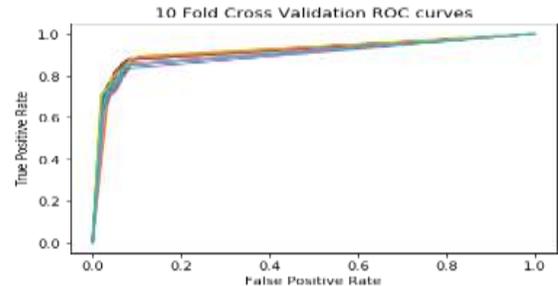
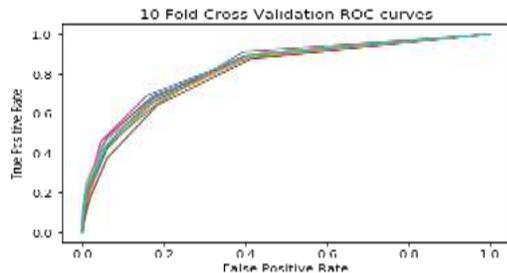
KNN Parameter Tuning

TABLE XII. KNN GRIDSEARCHCV RESULTS WITH ALL FEATURES MAXIMUM ACCURACY 82.65%, SPECIFICITY 60.01% AND SENSITIVITY 36.85%

param_n_neighbors	mean_test_spf	mean_test_recall	mean_test_accuracy
5	60.012901927092	0.368553459119497	0.826576089616115
9	61.358111609077	0.284696016771488	0.822934904503632
21	61.6474825441166	0.148427672955975	0.81317206867676
43	63.6911556294288	0.072536687631027	0.807959610602183
77	66.9705668401321	0.037735849056604	0.805890998862825
89	62.4337623814821	0.025576519916143	0.804401754459811

TABLE XIII. GRIDSEARCHCV RESULTS OF KNN USING FDR SELECTED FEATURES SHOWS GREAT ACCURACY OVER FULL FEATURES, ACHIEVED ACCURACY OF 91.56% WITH SPECIFICITY 81.54% AND SENSITIVITY 74.04%

param_n_neighbors	mean_test_spf	mean_test_recall	mean_test_accuracy
5	81.5485312602163	0.740461215932914	0.915604590177174
9	82.1180030569542	0.706918238993711	0.911715437213589
21	82.7013558282025	0.614675052410902	0.898477013352826
43	83.3771383841857	0.472117400419287	0.877047950603545
77	83.644701871265	0.293920335429769	0.849247010680139
89	85.1447572468791	0.254088050314465	0.843951326299478



Receiving Operating Curve Accuracy: 82.15695827072376
 Receiving Operating Curve Accuracy: 81.93653392513502
 Receiving Operating Curve Accuracy: 82.06034314209442
 Receiving Operating Curve Accuracy: 82.36647671448222
 Receiving Operating Curve Accuracy: 83.22803372846145
 Receiving Operating Curve Accuracy: 80.14621887137308
 Receiving Operating Curve Accuracy: 82.74669279949138
 Receiving Operating Curve Accuracy: 82.62315515141213
 Receiving Operating Curve Accuracy: 81.78181660072175
 Receiving Operating Curve Accuracy: 82.76304217006896
 Average Accuracy: 82.18092713739642

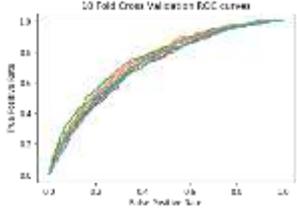
Receiving Operating Curve Accuracy: 91.94355482489823
 Receiving Operating Curve Accuracy: 91.51060955102596
 Receiving Operating Curve Accuracy: 91.51995101107273
 Receiving Operating Curve Accuracy: 91.61473850079078
 Receiving Operating Curve Accuracy: 89.30411280393969
 Receiving Operating Curve Accuracy: 91.84231716559303
 Receiving Operating Curve Accuracy: 90.42768397578847
 Receiving Operating Curve Accuracy: 90.06755508898804
 Receiving Operating Curve Accuracy: 92.16000862063807
 Receiving Operating Curve Accuracy: 90.09826182197293
 Average Accuracy: 91.04887933647078

Fig. 16. Table XII Results Cross Validated on Hold out Data using ROC Curves Average Accuracy 82.18%.

Fig. 17. Table XIII Results are Cross Validated on Hold out Data using ROC Curve Accuracy with Average Accuracy 91.04%.

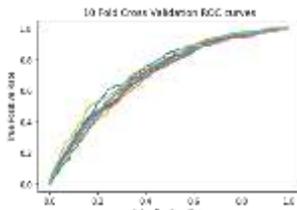
Naïve Bayes Parameter Tuning

TABLE XIV. NAIVE BAYES RESULTS WITH FDR FEATURES



Receiving Operating Curve Area: 73.0420156638747 Receiving Operating Curve Area: 74.59497369959087 Receiving Operating Curve Area: 75.5024765078923 Receiving Operating Curve Area: 71.22811693813246 Receiving Operating Curve Area: 72.60734679369789 Receiving Operating Curve Area: 70.84770490893375 Receiving Operating Curve Area: 73.88755641719222 Receiving Operating Curve Area: 71.54190774670525 Receiving Operating Curve Area: 73.6775428734368 Receiving Operating Curve Area: 71.54451369652202 Average Accuracy : 72.85219263888752	10 fold Accuracy results 74.88913525498891, 75.83148558758315, 76.88470066518846, 72.9490022172949, 74.50110864745011, 73.72505543237251, 76.05321507760532, 74.00221729490022, 75.22172949002217, 74.55654101995566 average accuracy 74.86 specificity 86%
---	---

TABLE XV. NAIVE BAYES RESULTS WITH ALL FEATURES



Receiving Operating Curve Area: 73.06162904924089 Receiving Operating Curve Area: 69.24110542922423 Receiving Operating Curve Area: 69.2456922888916 Receiving Operating Curve Area: 71.07592012214722 Receiving Operating Curve Area: 70.92692772917177 Receiving Operating Curve Area: 71.17445054945055 Receiving Operating Curve Area: 69.77037389976137 Receiving Operating Curve Area: 69.92784514336239 Receiving Operating Curve Area: 73.56140187784698 Receiving Operating Curve Area: 71.41525024323444 Average Accuracy : 70.9400596323315	73.28159645232816, 67.79379157427938, 71.34146341463415, 72.00665188470067, 71.61862527716187, 71.56319290465632, 71.17516629711751, 70.3991130820399, 71.56319290465632, 71.61862527716187 Average accuracy 71.23614190687361 specificity 85.95% sensitivity 32.78%
--	--

VI. RESULTS AND MODEL EVALUATION

The Model is evaluated on the basis of Accuracy, Specificity and Sensitivity and accuracy from Receiving Operating Curve. It's a screening test so more priority is to optimize the Specificity than sensitivity. The formulations of these metrics are: The confusion matrix is defined as

$$\begin{bmatrix} \text{True Positive} & \text{False Positive} \\ \text{False Negatives} & \text{True Negatives} \end{bmatrix}$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negatives}}{\text{True Positives} + \text{False Positive} + \text{False Negatives} + \text{True Negatives}}$$

$$\text{Specificity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{Sensitivity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}}$$

We tried to optimize the accuracy sensitivity and specificity using GridsearchCV method which applied 10 fold Stratified method for a given classifier with a given set of

input parameters. The evaluation results using different classifiers with GridsearchCV method are listed in following tables. The experiments are done twice using feature selection with Fisher Discriminate Ratio method.

VII. RESULT COMPARISONS CHARTS

The results of different classification models are compared in Fig. 18, 19 and 20.

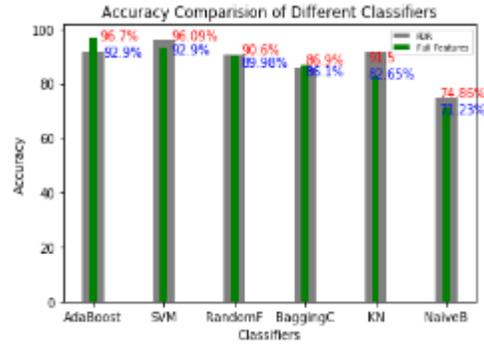
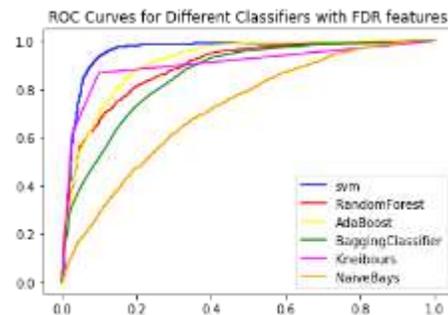
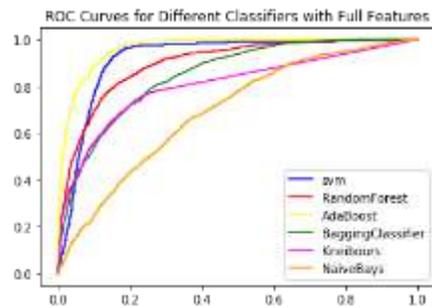


Fig. 18. Accuracy Comparison of different Classifiers with FDR as well as Full Features.



SVM ROC Accuracy 95.88491108807841
Random Forest ROC Accuracy 89.03526077312283
AdaBoost ROC Accuracy 91.3290828085577
Bagging ROC Accuracy 85.61042208468754
KNN ROC Accuracy 90.32602391010393
Naive Bayes ROC Accuracy 71.80806968262688

Fig. 19. ROC Accuracy Comparison of different Classifiers with FDR Features.



SVM ROC Accuracy 92.1662431476976
RandomForest ROC Accuracy 89.92557071323482
AdaBoost ROC Accuracy 96.39966479308009
Bagging ROC Accuracy 84.95671446702912
KNN ROC Accuracy 80.56178731319386
Naive Bayes ROC Accuracy 69.52003889699243

Fig. 20. ROC Accuracy Comparison of different Classifiers with Full Features.

VIII. CONCLUSION

The objective was to design a Decision support system for the Radiologist which help them for fast and correct predictions for the early detection of brain atrophy which can result into Alzheimer in future, we are able to deduce a system where radiologist can input the middle 25 slices from slice_no 110 to 140 of MRI to the system as input and on the basis of data in these slices the system can results the prediction about atrophy of brain. The accuracy of results can be achieved the best with AdaBoost classifier 96.7% and specificity and sensitivity. This study has achieved a better accuracy than the earlier research works because correct registration method and better classifiers that is AdaBoost. It will definitely going to support the radiologist for better decision of brain atrophy. This is a screening test so it's more important to have more specificity than sensitivity. This is an academic research with a purpose to explore machine learning classifiers and their parametric studies. The study also gives a hands out experiences for Image processing, how biomedical texture analysis helpful to extract image signatures which can be used for classification. It's a comparative study on the basis of different classifiers and further how classifiers results can be improved using feature selection criteria, but it also give an insight how some of classifiers are strong classifiers where feature selection criteria does not affect much its performance.

IX. FUTURE WORK

The Support system lacks the front end, in the future work we can design an automated system which automatically extract middle slices with proper frontend system where radiologist can feed the DICOM image slices and the system should give a report about the slices. Many other texture features can be explored to improve the performance. Many other feature extraction methods as well as classification techniques can be explored for better results. The study consumed much time in preprocessing of data, a fast and error data preprocessing steps can be explored in future work.

ACKNOWLEDGMENT

As the study is a practical study under the domain knowledge of Dr Ritesh Garg, Sr. Radiologist, who is owning MRI Diagnostic Center. The results had been verified under the supervision of radiologist. Our sincere thanks and gratitude to Dr Ritesh Garg for his unconditional support while analysing the data as without his help at every point of analysis, this study would have not completed.

REFERENCES

- [1] Braak and E. Braak, "Neuropathological staging of Alzheimer-related changes," *Acta Neuropathol.*, vol. 82, no. 4, pp. 239–259, Sep. 1991, doi: 10.1007/BF00308809.
- [2] G. B. Karas *et al.*, "A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry," *Neuroimage*, vol. 18, no. 4, pp. 895–907, 2003, doi: 10.1016/S1053-8119(03)00041-7.
- [3] Z. Lao, D. Shen, Z. Xue, B. Karacali, S. M. Resnick, and C. Davatzikos, "Morphological classification of brains via high-dimensional shape transformations and machine learning methods," *Neuroimage*, vol. 21, no. 1, pp. 46–57, 2004, doi: 10.1016/j.neuroimage.2003.09.027.
- [4] K. Specht, M. Minnerop, J. Müller-Hübenthal, and T. Klockgether, "Voxel-based analysis of multiple-system atrophy of cerebellar type: Complementary results by combining voxel-based morphometry and voxel-based relaxometry," *Neuroimage*, vol. 25, no. 1, pp. 287–293, 2005, doi: 10.1016/j.neuroimage.2004.11.022.
- [5] W. Wen, P. S. Sachdev, X. Chen, and K. Anstey, "Gray matter reduction is correlated with white matter hyperintensity volume: A voxel-based morphometric study in a large epidemiological sample," *Neuroimage*, vol. 29, no. 4, pp. 1031–1039, 2006, doi: 10.1016/j.neuroimage.2005.08.057.
- [6] S. J. Teipel *et al.*, "Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment," *Neuroimage*, vol. 38, no. 1, pp. 13–24, 2007, doi: 10.1016/j.neuroimage.2007.07.008.
- [7] V. A. Cardenas, C. Studholme, S. Gazdzinski, T. C. Durazzo, and D. J. Meyerhoff, "Deformation-based morphometry of brain changes in alcohol dependence and abstinence," *Neuroimage*, vol. 34, no. 3, pp. 879–887, 2007, doi: 10.1016/j.neuroimage.2006.10.015.
- [8] S. J. Teipel, T. Meindl, L. Grinberg, H. Heinsen, and H. Hampel, "Novel MRI techniques in the assessment of dementia," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 35, no. SUPPL. 1, pp. 58–69, 2008, doi: 10.1007/s00259-007-0703-z.
- [9] V. A. Cardenas *et al.*, "Deformation-based morphometry reveals brain atrophy in frontotemporal dementia," *Arch. Neurol.*, vol. 64, no. 6, pp. 873–877, 2007, doi: 10.1001/archneur.64.6.873.
- [10] R. Stahl, O. Dietrich, S. J. Teipel, H. Hampel, M. F. Reiser, and S. O. Schoenberg, "White matter damage in Alzheimer disease and mild cognitive impairment: Assessment with diffusion-tensor MR imaging and parallel imaging techniques," *Radiology*, vol. 243, no. 2, pp. 483–492, 2007, doi: 10.1148/radiol.2432051714.
- [11] D. Sydykova *et al.*, "Fiber connections between the cerebral cortex and the corpus callosum in Alzheimer's disease: A diffusion tensor imaging and voxel-based morphometry study," *Cereb. Cortex*, vol. 17, no. 10, pp. 2276–2282, 2007, doi: 10.1093/cercor/bhl136.
- [12] P. M. THOMPSON *et al.*, "Tracking Alzheimer's Disease," *Ann. N. Y. Acad. Sci.*, vol. 1097, no. 1, pp. 183–214, Feb. 2007, doi: 10.1196/annals.1379.017.
- [13] X. Hua *et al.*, "Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects," *Neuroimage*, vol. 43, no. 3, pp. 458–469, 2008, doi: 10.1016/j.neuroimage.2008.07.013.
- [14] C. Plant *et al.*, "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease," *Neuroimage*, vol. 50, no. 1, pp. 162–174, 2010, doi: 10.1016/j.neuroimage.2009.11.046.
- [15] R. Casanova *et al.*, "Using high-dimensional machine learning methods to estimate an anatomical risk factor for Alzheimer's disease across imaging databases," *Neuroimage*, vol. 183, pp. 401–411, 2018, doi: 10.1016/j.neuroimage.2018.08.040.
- [16] R. Casanova, F. C. Hsu, and M. A. Espeland, "Classification of Structural MRI Images in Alzheimer's Disease from the Perspective of Ill-Posed Problems," *PLoS One*, vol. 7, no. 10, 2012, doi: 10.1371/journal.pone.0044877.
- [17] S. Adaszewski, J. Dukart, F. Kherif, R. Frackowiak, and B. Draganski, "How early can we predict Alzheimer's disease using computational anatomy?," *Neurobiol. Aging*, vol. 34, no. 12, pp. 2815–2826, 2013, doi: 10.1016/j.neurobiolaging.2013.06.015.
- [18] O. Ben Ahmed, J. Benois-Pineau, M. Allard, C. Ben Amar, and G. Catheline, "Classification of Alzheimer's disease subjects from MRI using hippocampal visual features," *Multimed. Tools Appl.*, vol. 74, no. 4, pp. 1249–1266, 2014, doi: 10.1007/s11042-014-2123-y.
- [19] A. E.R., S. Varma, and V. Paul, "Classification of Brain MR Images using Texture Feature Extraction," *Int. J. Comput. Sci. Eng. Commun.*, vol. 5, no. 5, pp. 1722–1729, 2017.
- [20] T. Altaf, S. M. Anwar, N. Gul, M. N. Majeed, and M. Majid, "Multi-class Alzheimer's disease classification using image and clinical features," *Biomed. Signal Process. Control*, vol. 43, pp. 64–74, 2018, doi: 10.1016/j.bspc.2018.02.019.
- [21] L. Nanni, S. Brahnam, C. Salvatore, and I. Castiglioni, "Texture descriptors and voxels for the early diagnosis of Alzheimer's disease," *Artif. Intell. Med.*, vol. 97, no. May, pp. 19–26, 2019, doi: 10.1016/j.artmed.2019.05.003.

- [22] S. Lee, H. Lee, and K. W. Kim, "Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume," *J. Psychiatry Neurosci.*, vol. 45, no. 1, pp. 7–14, 2020, doi: 10.1503/jpn.180171.
- [23] J. H. Cai *et al.*, "Magnetic Resonance Texture Analysis in Alzheimer's disease," *Acad. Radiol.*, 2020, doi: 10.1016/j.acra.2020.01.006.
- [24] V. P. S. Rallabandi, K. Tulpule, and M. Gattu, "Automatic classification of cognitively normal, mild cognitive impairment and Alzheimer's disease using structural MRI analysis," *Informatics Med. Unlocked*, vol. 18, 2020, doi: 10.1016/j.imu.2020.100305.
- [25] Z. Yaniv, B. C. Lowekamp, H. J. Johnson, and R. Beare, "SimpleITK Image-Analysis Notebooks: a Collaborative Environment for Education and Reproducible Research," *J. Digit. Imaging*, vol. 31, no. 3, pp. 290–303, 2018, doi: 10.1007/s10278-017-0037-8.
- [26] B. Fischl, "FreeSurfer Authos Manuscript," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012, doi: 10.1016/j.neuroimage.2012.01.021.FreeSurfer.