Multimodal MRI Based Classification and Prediction of Alzheimer's Disease Using Random Forest Ensemble

Thushara A
Dept of Computer Science and Engineering,
TKM College of Engineering, Kollam,
APJ Abdul Kalam Technological University,
Kerala , India.
tusharaa@gmail.com

UshaDevi Amma C
Dept. of Electrical and Electronics Engineering,
TKM College of Engineering, Kollam,
APJ Abdul Kalam Technological University,
Kerala, India
ushadevi.a.c@gmail.com

Ansamma John
Dept of Computer Science and Engineering
TKM College of Engineering, Kollam,
APJ Abdul Kalam Technological University,
Kerala, India
ansamma.john@gmail.com

Reshma Saju
Dept of Computer Science and Engineering
TKM College of Engineering,Kollam,
APJ Abdul Kalam Technological University,
Kerala, India
reshsara1009@gmail.com

Abstract— Alzheimer's disease (AD) is a neurodegenerative disorder that affects millions of people worldwide and it accounts for a significant decrease in the quality of life of patients and their families. Currently, available treatment options for AD is merely palliative and no drugs are available for the inexorable progression of the disorder that is diagnosed during the later stage of the disease. So the early diagnosis of AD is an optimal strategy in formulating the treatment plan. Neuroimaging modalities like Magnetic Resonance Imaging (MRI), resting-state functional Magnetic resonance imaging (rs-fMRI), Diffusion Tensor Imaging (DTI) and Positron emission tomography (PET) are used to diagnose the structural and functional alteration caused by AD. For the past few years, machine learning methods are widely used to analyze the neuroimaging data acquired from MRI imaging modalities for the diagnosis and prediction of neurological disorder. In this work, the random forest classification algorithm is used to classify and predict Alzheimer's disease. The data set that is used in this study is TADPOLE data set, which has been acquired from Alzheimer's neuroimaging Initiative (ADNI). In this work, the multiclass classification that distinguishes the different level of Alzheimer's disease has achieved an accuracy comparable to current research in the prediction of AD.

Keywords— Alzheimer's disease, ADNI, MultiModal MRI, Random Forest.

I. INTRODUCTION

Alzheimer's disease (AD) is an incurable neurodegenerative disease that destroys the neurons and affects many parts of the brain, including the cerebral cortex, entorhinal cortex, hippocampus and the lobes (frontal, occipital, parietal, and temporal) of the brain. The causes of Alzheimer's Disease are still unknown, but risk factors include age, sex, family history, lifestyle, heart health, down syndrome, poor sleep patterns and past head trauma [1]. Diagnosing the disease at an early stage can decrease the rate of the progression of the disease [2]. But it is hard to diagnose at its early stage as memory loss is a part of normal ageing [4]. So by the time the disease has been identified, it might have worsened the condition of the AD patient. The clinical diagnosis of AD can be done based on patient history, Mini-Mental State Examination (MMSE) [5], Alzheimer's Disease Assessment Scale (ADAS) [6] both ADAS11 and ADAS13 and by neuroimaging methods like Magnetic Resonance Imaging (MRI)[7,8]. Based on the level of similarity this disease has been mainly divided into three stages: Alzheimer's disease(AD), Normal Control (NC) and mild cognitive impairment (MCI) [3].

Neuroimaging techniques like structural MRI [9], resting-state functional MRI (rs-fMRI) [10,11], diffusion tensor imaging (DTI) [12] and positron emission tomography (PET) [13] were used for identifying the biomarkers that can effectively diagnose the progression of AD. Structural MRI can measure the structural changes like cortical thinning, atrophy in multiple brain regions and variation in regional tissue density, caused by neurodegeneration. rs-fMRI can measure the functional connectivity between the different region of the brain by quantifying the correlation between the blood oxygen level-dependent signal (BOLD) and neural activity. rs-fMRI studies [14] have shown that alteration in the functional connectivity can capture the level of the impairment caused by the disease. The disruption of connectivity between the different brain region is the biomarker that can be obtained from rs-fMRI. Diffusion tensor imaging can determine the diffusion pattern of the brain and this abnormal diffusion pattern is a biomarker for AD diagnosis [15]. Positron emission tomography can distinguish between AD and healthy individual by characterizing the pattern of glucose metabolism and amyloid deposition [16]. Several biomarkers from different MRI imaging modalities provide complementary information. Indeed, to achieve optimal accuracy for the diagnosis of AD, images from a single modality is not sufficient. As stated in [17], only by combining the features obtained from multimodal MRI, the classification and prediction of the disease could be effectively used to formulate the treatment plan in the clinical realm.

Automated feature extraction, analysis and the classification of brain images acquired from MRI help clinicians to speedily and accurately diagnose the disease. In recent years, there has been tremendous growth in the development of machine learning techniques, including feature selection and classification, which aids clinicians to diagnose various brain disorders at the early stage itself. The majority of neuroimaging studies are based on machine learning techniques like linear discriminant analysis (LDA) [19], support vector machine (SVM) [18], Naive Bayesian [20] or ensemble algorithms [21]. But in the case of multi-modality based feature selection and classification method, ensemble algorithms yielded better results than others.

978-1-7281-6453-3/20/$31.00 ©2020 IEEE

Among the ensemble-based algorithms, Random Forest (RF) algorithm [22] provided better classification accuracies in many neurological diseases, but it has been less applied on data acquired from the Alzheimer's patient's MRI image [23].

This study aims to i) classify data to different labels such as Normal Control(NC), Mild Cognitive impairment (MCI) and Alzheimer's disease(AD), ii) predict whether an MCI patient is non-converted (ncMCI) or converted (cMCI) using RF classification algorithm iii) to compare RF and genetic algorithm in terms of classification and prediction accuracy of AD. The major challenge in applying machine learning algorithm for AD disease classification and prediction is incomplete high dimensional data. To overcome this challenge we proposed a method to improve the classification accuracy based on some appropriate feature selection algorithm using the TADPOLE (The Alzheimer's Disease Prediction of Longitudinal Evolution) [24] data set.

The rest of this paper is organized as follows. In Section II, a brief overview of the previous studies on MRI based feature selection and classification for AD diagnosis. In section III, methodologies and algorithm used in this study have been introduced. In Section IV, results obtained have been analysed and compared. Conclusion and future directions are emphasized in section V.

II. RELATED WORK

Multimodal based studies [33-35] for the diagnosis and prediction of Alzheimer's disease (AD) has achieved significant success when compared to the single modality based classification. In many recent studies, the structural and functional feature like volume, shape, functional connectivity and diffusivity has been extracted from structural MRI, DTI, rs-fMRI and PET [36]. The number of features that have been extracted from these imaging modalities is so large in number. To reduce the number of features, principal component analysis (PCA), Linear discriminant analysis (LDA) and t-test have been widely used. The features extracted from structural MRI along with the demographic information, MMSE score, and genetic data have been combined and studied [37]. For instance, the volumetric information of the grey matter, white matter and cerebral spinal fluid combined with age, sex, and genetic data have been used as the input features for SVM-based AD classification. [26] and [38] combined features extracted from structural MRI, PET and CSF biomarker has been used for SVM-based AD classification. In [39] regional grey matter volume, intensity features from PET, and cerebral spinal fluid are the features, used for classification using SVM. The major pitfall of concatenating the features from different modality is that the classification algorithm considers all features as equal-weighted which decreases the classification accuracy.

The most existing multi-modality data-based classification methods use the filled data set. In Alzheimer's disease Neuroimaging Initiative (ADNI) [25] data set, one-third of the subjects have incomplete MRI, PET, CSF and clinical assessment score data. Because of incomplete multi-modality data, [40,41] explores the possibility of imputing the missing values. To address the incomplete multimodal data, [42] proposed an unsupervised method which maps the incomplete multimodal data into a latent space to obtain a new complete data set. [43] proposed a method to fill the incomplete data by nonnegative matrix factorization (NMF) and sparse regularizer. Filling the incomplete data in the high dimensional space is another factor that improves classification accuracy.

The selection of relevant and most important features helps to reduce the computational time and improves the accuracy in classification. Feature selection methods can be broadly divided into manual feature selection and automated feature selection. For manual feature selection domain expertise is required, whereas, in the automatic feature selection method, no domain expertise is required. The automated methods like t-tests [44], Pearson's correlation coefficient test [45] wrapper based methods [27] and elastic net regression have been used to improve the classification accuracy. In the Pittsburgh Brain Activity Interpretation Competition [23], the best accuracy has been obtained for the teams who applied domain knowledge rather than the automated method for feature selection. So by combining the features obtained from both manual and automated feature selection methods, the best features that can improve the classification accuracy can be selected.

To classify the different types of subjects into AD, MCI and NC, machine learning algorithms [28-30] has been used. The brain imaging data obtained from MRI rely on the machine learning algorithms like Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), ensemble algorithms or Naive Bayesian to effectively distinguish between different subject classes. Among these algorithms, ensemble algorithms showed better accuracy for the classification of subjects based on the brain images obtained from multimodal MRI. RF [31] algorithm has better accuracy than other ensemble approaches in the classification of the brain image with the neurological disorder. But very less attention has been paid in applying the RF algorithm for the multiclass classification problem of AD. Noisy data can be handled more efficiently by RF algorithm than the other machine learning algorithm and it also has the capability for parallel processing which reduces the computational time especially in the case of high dimensional data.

For solving high dimensional classification problem genetic algorithm searches in the subset features to find the feature combination that provides the optimal solution for predicting the AD. To build the prediction model, [47] used a genetic algorithm with Logistic Regression (LR). The receiver operating characteristic (ROC) generated by the model has been used to evaluate the fitness of each individual and best individual survives to the next iteration. LR model has been repeatedly applied and the prediction accuracy for neuroimaging data is very high. In [46], the author used fisher criterion to define the objective function which evaluates the fitness of the feature subsets. This method selected the most discriminant features so that the original feature vector is mapped into a low dimensional space. It has been observed that researchers had given more importance to feature extraction phase and not much on feature selection and classification phase of the AD patient data set. Hence, proper integration of feature selection techniques and machine learning model is necessary to improve the classification accuracy of AD.

The limitation of the existing research studies in the classification and prediction of Alzheimer's disease is the inability to select appropriate features that have been obtained from various neuroimaging MRI modalities and clinical assessment. The aim of this study is to i) handle noisy incomplete data ii) to effectively handle the issue of the curse of dimensionality and to find the correlation between the features (iii) achieve better classification accuracy when compared to other methods. The proposed methodology is discussed in the next section.

III. METHODOLOGY

A critical challenge in using the classification framework in the clinical realm is the inability of predictive models to have an accurate diagnosis of new patient data. It is due to the high error rate in the multiclass classification of AD. In our method, the RF-based classification algorithm has been proposed to improve the multiclass classification accuracy of AD. The processing steps followed by the proposed methodology are as follows: i) Data Pre-processing ii) Feature Selection based on domain knowledge iii) Feature Selection by applying RF algorithm iv) Feature selection using Genetic algorithm v) Apply RF classification algorithm on the selected combined features.

A. Data Preprocessing

Data preprocessing has a substantial impact on the classification accuracy of the machine learning algorithm. The data set has been cleaned, by filling the missing data and by removing noisy inconsistent data. One method for handling the missing data is by simply ignoring the tuple. As ignoring the subject data in medical imaging is not appropriate as the acquisition, storage and retrieval of MRI images are time-consuming and the number of available subject data is very limited. So the solution is to retain the data, by applying appropriate algorithms, to improve the classification accuracy. The method adopted in this work to handle the missing data is data imputation. Rather than filling the missing data by a constant or mean or median, a probable value is used to fill the missing data. This is done with the help of an ensemble-based RF algorithm. The strategy adopted is to find the pattern of the missing value by proximity imputation. By using this method, features with the missing value have been ranked from the minimum to maximum based on the number of missing fields. The minimum missing field feature has been initially selected for imputation and filled with rough and inaccurate values. As the next step, the RF imputation algorithm is executed and the proximity matrix is being generated. Based on the proximity matrix the missing values have been recomputed and RF algorithm iterates to improve the result. A Label encoder which converts the categorical data to numerical values has also been used. As our data set contain very large number of the missing field for a particular feature, by applying the mean or median method to fill the missing data has resulted in filling the data field with a large number of zero's, which in turn reduces the classification accuracy. So RF-based imputation has better accuracy, robustness and can perform better on incomplete high dimensional data.

B. Feature Selection and Classification

The Feature selection method reduces the computational complexity of the classifier by picking up the most relevant features for classification and thereby enables the machine learning algorithm to train faster. It also reduces the over-fitting and improves accuracy and robustness. Features can be selected based on domain knowledge or by using the machine learning method. The domain knowledge is very crucial in selecting the most relevant biomarkers that affect disease progression. Domain knowledge can be obtained by analyzing the biomarkers that affect the brain progression to AD, which are identified in clinical studies. This acquired knowledge can be used to manually select the features so as improve the computational efficiency and accuracy of the machine learning algorithm. To find the subset of the manually selected feature, machine learning algorithms are applied to the data set. Two feature selection algorithms used in this study are RF and GA and its classification accuracies are also compared.

RF classifier constructs decision trees using the different bootstrap sample. For the construction of the decision tree, features are randomly selected and Gini index G (T) for a feature is computed as per equation (1)

$$G(T)=1-\sum_{i=1}^{c} p_i^2 \qquad (1)$$

where, c is the number of classes and p_i is the relative frequency of class 'i' in T. The feature which is having the lowest Gini index is selected as the split node. The decision tree generated provides a relative ranking of the features. So the RF selection algorithm selects the most relevant features that reduce the computational time and improves the classification accuracy.

In machine learning**,** one of the uses of genetic algorithms is to select the right number of features to create a predictive model. The selection of the right subset of variables is a combinatorial optimization problem. The advantage of the genetic algorithm over others in feature selection is that it allows the best solution to emerge from the best of prior solutions, and it improves the selection over time. In this work, GA encodes the features as a binary chromosome where the gene value of one indicates that the feature has been selected and zero otherwise. Using positional indexing the feature with gene value one is selected and chromosomes are ranked based on the fitness evaluation done through logistic regression. The features with the highest rank have been selected and passed to the next generation, and the least significant features in the current population have been passed to crossover and mutation function. This process iteratively continues until an optimal score or the maximum generation has reached.

The Machine learning algorithm for features selection and classification has been applied to the TADPOLE data sets. This data set has been generated from ADNI and it consists of: i) Demographic information ii) Genetic biomarkers iii) Clinically acquired cognitive test result iv) Cerebral Spinal fluid biomarker and v) Multimodal MRI data which includes features extracted from Structural MRI and positron emission tomography. The TADPOLE data set consist of three kinds of data: a training data set, prediction data set and test data set. The training data set contains

measurement that can be used to train the model, whereas the prediction data set contains data that can be used to predict the subject's disease. To evaluate the model test data set can be used.

IV. RESULT

The proposed methods are applied to the TADPOLE data set which contains clinical and neuroimaging biomarkers. Table 1 shows the demographic information of the patient obtained from the data set. As the data set is an incomplete high dimensional, the RF imputation algorithm has been applied to fill the missing value. RF imputation algorithm when applied to fill the missing value improved the classification accuracy as this method filled the high dimensional incomplete data more accurately than filling by the mean or median method. Table 2 and Table 3 shows the same sample data set before and after applying the imputation algorithm. Around 2000 features in the data set have been reduced to 486 features manually based on prior biological knowledge about the AD biomarkers. Among the 486 some of the most relevant biomarkers are listed in Table 4.

The two feature selection algorithms, RF feature selection and the genetic algorithm have been applied independently on this reduced feature set. In the RF based feature selection process, 100 trees were used for forest creation. Features are randomly selected from 486 manually selected features, and the Gini index has been computed for each selected feature. Based on the level of purity, the importance of each feature has been computed and plotted in Figure 1. The final decision on the level of feature importance has been computed by averaging the prediction of each tree. The based on the final value, the features are ranked and the top 21 features were selected using the RF algorithm. The sample feature set after feature selection using RF and GA is as shown in Table 5 and Table 6 respectively.

TABLE 1: SUBJECT DEMOGRAPHIC INFORMATION

CLASS	NO OF VISITS		AGE (YEARS) (MEAN ± SD)		MMSE (MEAN ± SD)	
	MALE	FEMALE	MALE	FEMALE	MALE	FEMALE
AD	989	744	74.35 ± 6.79	73.11 ± 7.14	26.56 ± 3.78	26.62 ± 4.16
MCI	2373	1559	74.34 ± 6.79	73.12 ± 7.14	26.57 ± 3.77	26.62 ± 4.15
cMCI	218	154	74.36 ± 6.79	73.09 ± 7.14	26.56 ± 3.78	26.61 ± 4.16
NC	1317	1351	74.34 ± 6.79	73.11 ± 7.14	26.56 ± 3.77	26.63 ± 4.15

TABLE 2: SAMPLE DATA SET BEFORE IMPUTATION

AGE	GENDER	EDUCATION	APOE4	FDG	ADAS13	MMSE	CLASS
71.7	Male	14	0		10	27	NC
71.7	Male	14	0	1.47245	9	28	NC
67.4	Male	20	0		42	15	AD
77.1	Female	18	1		5	30	NC

TABLE 3: SAMPLE DATA SET AFTER IMPUTATION

AGE	GENDER	EDUCATION	APOE4	FDG	ADAS13	MMSE	CLASS
71.7	Male	14	0	1.25573	10	27	3
71.7	Male	14	0	1.47245	9	28	3
67.4	Male	20	0	1.24932	42	15	0
77.1	Female	18	1	1.26356	5	30	3

TABLE 4: SAMPLE FEATURE SET AFTER MANUAL SELECTION

Gender, Education, APOE4, FDG, ADAS13, MMSE, Ventricles, Hippocampus, Whole-brain, Entorhinal,Fusiform, Midtemp, Class, Cortical Thickness Average of Right postcentral, Surface Area of Right superoortemporal, Cortical Thickness Average of Leftinsula, Cortical Thickness Average of Left cuneus, Surface Area of Left middle temporal, Surface Area of Left transverse temporal, Surface Area of Left parahippocampal, Surface Area of Left entorhinal, Temporal AV-45 Measure, Hippocampal FDG Measure, 3rd-Ventricle, 5th-Ventricle, Right-Cerebellum-White-Matter, Right-Entorhinal, left-Entorhinal, Superiortemporal, Left-Lateral-Ventricle, Cortical Thickness Standard Deviation of Rightlingual, Cortical Thickness, Average of Right Medial Orbito Frontal, Cortical Thickness Average of Right Isthmus Cingulate, Cortical Thickness Average of Right Inferior Parietal.

TABLE 5: A SAMPLE FEATURE SUBSET AFTER FEATURE SELECTION USING RF

ADAS13	MMSE	Class	Midtemp	AGE	FDG	Cortical Thickness	Volume of left hippocampus	Ventricle	Whole Brain
10	27	3	19500	71.7	1.255731	0.636	3835	21897	1040560
9	28	3	19500	71.7	1.47245	0.624	3702	21897	1040560
42	15	0	19249	67.4	1.24932	0.695	3310	33771	1180520
5	30	3	15838	77.1	1.263558	0.726	3294	27431	901820

TABLE 6: A SAMPLE FEATURE SUBSET AFTER FEATURE SELECTION USING GENETIC ALGORITHM

CLASS	AGE	APOE4	FDG	ADAS13	Hippocampus	Whole Brain	Entorhinal	Fusiform
3	71.7	0	1.255730539	10	7825	1038660	3682	19254
3	71.7	0	1.47245	9	7672	1026760	3925	19019
0	67.4	0	1.249319734	42	7402	1116680	2385	11366
3	77.1	1	1.263558169	5	6095	886668	3464	15256

The RF classification algorithm has been applied to the data set obtained after the feature selection process. Then cross-validation of 10 folds and 20 trees has been done. From the results in Table 7, it has been noticed that classifier achieved an accuracy of 69.33% as well as macro-averaged values of 69.33%, 40.94% and 51.48% for precision, sensitivity and F-score measures respectively. The best performance has been achieved for the class NC (precision 74.17%, Sensitivity 58.09%, F-score 65.15%), while the worst results are attained for the cMCI class (precision 5%, Sensitivity 0.79% and F-score 1.36%). For comparison, RF Classification algorithm has been applied to the data set obtained after the feature selection using Genetic algorithm with a percentage split of 66% and 20 trees. The result is shown in Table 8 and it has been noticed that overall accuracy is 57.54%. NC and cMCI achieved the best and worst classification accuracy respectively. By analyzing the data set, it is understood that the low precision in cMCI is due to less number of subject data in that category. From the results, it has been concluded that RF selection and classification algorithm provided better classification accuracy than the GA based selection and classification algorithm.

The challenge in the diagnosis of AD is the accuracy of multiclass classification. In this work, a four-class classification problem of AD is explored. The proposed method is compared with SVM [48], OVR [49] and SAE-ZEROMASK [48]. The performance comparison is given in Table 9. So the proposed method is an effective choice for multiclass classification of AD.

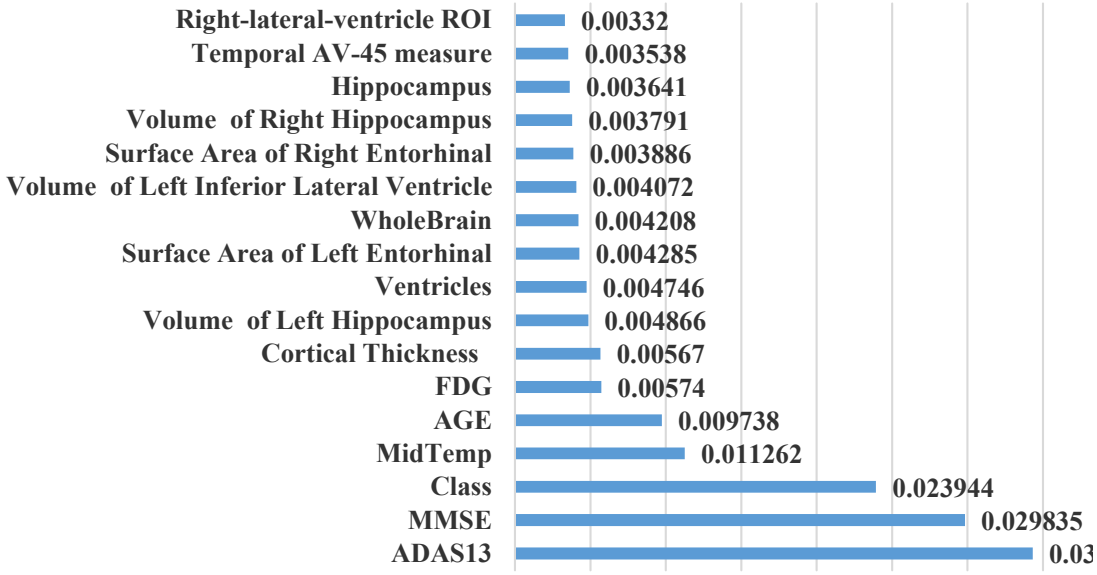

Figure 1: Feature Importance obtained from RF Algorithm

TABLE 7: PERFORMANCE METRIC OF RANDOM FOREST ALGORITHM

Class	Precision	Sensitivity	F-Score
AD	65.54%	28.28%	39.51%
MCI	69.43%	92.89%	79.47%
cMCI	5%	0.79%	1.36%
NC	74.17%	58.09%	65.15%
Macro Average	69.33%	40.94%	51.48%
Accuracy	**69.33%**		

TABLE 8: PERFORMANCE METRIC OF GENETIC ALGORITHM

Class	Precision	Sensitivity	F-Score
AD	75.00%	0.81%	1.60%
MCI	56.65%	99.37%	72.16%
cMCI	0%	0.00%	0.00%
NC	87.91%	10.15%	18.20%
Macro Average	57.54%	36.52%	44.68%
Accuracy	**57.54%**		

TABLE 9: PERFORMANCE (%) COMPARISON OF MULTICLASS AD CLASSIFICATION

Method	AD	MCI	cMCI	NC	Accuracy
SVM	53.74	44.58	46.45	49.74	47.74
OVR	56.33	52.53	50.68	58.99	56.25
SAE-Zero Mask	59.07	52.21	40.17	59.07	64.07
Proposed Method	**77.36**	**75.61**	**47.19**	**79.06**	**69.33**

V. CONCLUSION

Recent studies on the prediction of AD shows that machine learning algorithms are reliable in its early diagnosis. These algorithms use various biomarkers like cognitive tests, demographic information, PET measures and MRI measures for predicting AD. From our study, RF provided better accuracy with a minimum subset of features and it is clear that features selected using RF showed superiority over features selected using Genetic Algorithm. The feature selection algorithm selects features, which includes some demographic measures like age, cognitive test results such as ADAS13 and MMSE, MRI measures like Hippocampus and mid temporal lobe and PET measures. At least one of the features from each biomarker category is included in the resultant data set, so complimentary information is obtained which improved the accuracy. The result of classification showed that the cMCI is either classified as cMCI or AD, which does not have that much effect in the clinical realm. The limitation of this work is the low accuracy in cMCI class, and it is due to the less number of subject data available in the data set. In future work, we will incorporate more data in the prediction class and instead of directly applying a single RF feature selection model, different RF models with different parameters can be modelled and a combined result of each model could be fussed to be fed into the classifier to improve the classification accuracy. In the case of Genetic Algorithm, a variable-length chromosome (VLC) could be used for feature subset selection to minimize feature subset size. Also, different fitness functions and other parameters can be analyzed to identify which one provides better accuracy.

VI. REFERENCES

[1] "2019 Alzheimer's disease facts and figures," *Alzheimer's & Dementia,* vol. 15, no. 3, pp. 321-387, 2019.

[2] J. Rasmussen and H. Langerman, "Alzheimer's Disease – Why We Need Early Diagnosis," *Degenerative Neurological and Neuromuscular Disease,* vol. Volume 9, pp. 123-130, 2019.

[3] Islam, Jyoti, and Yanqing Zhang. "GAN-based synthetic brain PET image generation." *Brain Informatics* 7 (2020): 1-12.

[4] G. Karas et al., "Precuneus atrophy in early-onset Alzheimer's disease: A morphometric structural MRI study," *Neuroradiology,* vol. 49, no. 12, pp. 967-976, 2007.

[5] I. McDowell, B. Kristjansson, G. B. Hill and R. Hébert, "Community screening for dementia: The Mini Mental State Exam (MMSE) and modified Mini-mental State Exam (3MS) compared," *Journal of Clinical Epidemiology,* vol. 50, no. 4, pp. 377-383, 1997.

[6] G. Weyer, H. Erzigkeit, S. Kanowski, R. Ihl and D. Hadler, "Alzheimer's Disease Assessment Scale: Reliability and validity in a multicenter clinical trial," *International Psychogeriatrics,* vol. 9, no. 2, pp. 123-138, 1997.

[7] C. R. Jack et al., "Magnetic resonance imaging in Alzheimer's Disease Neuroimaging Initiative 2," *Alzheimer's and Dementia,* vol. 11, no. 7, pp. 740-756, 2015.

[8] P. L. McGeer, "Brain Imaging in Alzheimer's Disease," *British Medical Bulletin,* vol. 42, no. 1, pp. 24-28, 1986.

[9] G. Frisoni, N. Fox, C. Jack, P. Scheltens and P. Thompson, "The clinical use of structural MRI in Alzheimer disease", *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67-77, 2010.

[10] E. L. Dennis and P. M. Thompson, "Functional brain connectivity using fMRI in aging and Alzheimer's disease", *Neuropsychol. Rev*., vol. 24, no. 1, pp. 49-62, Mar. 2014.

[11] P VemuriP. Vemuri, D. T. Jones and C. R. Jack, "Resting state functional mri in alzheimer's disease", *Alzheimer's research& therapy*, vol. 4, no. 1, pp. 1-9, 2012.

[12] T. M. Nir et al., "Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging," *NeuroImage: Clinical,* vol. 3, pp. 180-195, 1 1 2013.

[13] A. Nordberg, J. O. Rinne, A. Kadir and B. Langstrom, "The use of PET in Alzheimer disease", *Nature Rev. Neurol.*, vol. 6, no. 2, pp. 78-87, 2010.

[14] M. A. Binnewijzend et al., "Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment," *Neurobiology of Aging,* vol. 33, no. 9, pp. 2018-2028, 1 9 2012.

[15] P. Selnes et al., "Diffusion tensor imaging surpasses cerebrospinal fluid as predictor of cognitive decline and medial temporal lobe atrophy in subjective cognitive impairment and mild cognitive impairment," *Journal of Alzheimer's Disease,* vol. 33, no. 3, pp. 723-736, 2013.
[16] R. La Joie et al., "Region-specific hierarchy between atrophy, hypometabolism, and 2-amyloid (Aβ) load in Alzheimer's disease dementia," *Journal of Neuroscience,* vol. 32, no. 46, pp. 16265-16273, 2012.
[17] S. Rathore, M. Habes, M. A. Iftikhar, A. Shacklett and C. Davatzikos, "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages", *NeuroImage*, vol. 155, pp. 530-548, 2017.
[18] B. Magnin et al., "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI", *Neuroradiology*, vol. 51, pp. 73-83, 2009.
[19] B. M. French, M. R. Dawson and A. R. Dobbs, "Classification and staging of dementia of the Alzheimer type: A comparison between neural networks and linear discriminant analysis," *Archives of Neurology,* vol. 54, no. 8, pp. 1001-1009, 1997.
[20] S. R. Bhagya Shree and H. S. Sheshadri, "Diagnosis of Alzheimer's disease using Naive Bayesian Classifier," *Neural Computing and Applications,* vol. 29, no. 1, pp. 123-132, 2018.
[21 S. Farhan, M. A. Fahiem and H. Tauseef, "An Ensemble-of-Classifiers Based Approach for Early Diagnosis of Alzheimer's Disease: Classification Using Structural Features of Brain Images", *Comput. Math. Methods Med.*, vol. 2014, Sep. 2014.
[22] A. Liaw and M. Wiener, "Classification and regression by randomforest", *R News*, vol. 2, no. 3, pp. 18-22, 2002.
[23] A. Sarica, A. Cerasa and A. Quattrone, "Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review", *Frontiers in Aging Neuroscience*, vol. 9, pp. 329, 2017.
[24] R. V. Marinescu, N. P. Oxtoby, A. L. Young, E. E. Bron, A. W. Toga, M. W. Weiner, et al., "TADPOLE challenge: Prediction of longitudinal evolution in Alzheimer's disease", *arXiv:1805.03909*, 2018.
[25] C. R. Jacket et al., "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging,* vol. 27, no. 4, pp. 685-691, 1 4 2008.
[26] K. Blennow, H. Hampel, M. Weiner and H. Zetterberg, "Cerebrospinal fluid and plasma biomarkers in Alzheimer disease", *Nature Rev. Neurol.,* vol. 6, no. 3, pp. 131-144, 2010.
[27] C.-Y. Wee, P.-T. Yap and D. Shen, "Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns," *Human Brain Mapping,* vol. 34, no. 12, pp. 3411-3425, 2013.
[28] D. Zhang and D. Shen, "Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease," *NeuroImage,* vol. 59, no. 2, pp. 895-907, 2012.
[29] F. Falahati, E. Westman and A. Simmons, "Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging", *J. Alzheimers Dis.*, vol. 41, no. 3, pp. 685-708, 2014.
[30] P. T. Trzepacz, P. Yu, J. Sun, K. Schuh, M. Case, M. M. Witte, H. Hochstetler and A. Hake, "Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia," *Neurobiology of Aging,* vol. 35, no. 1, pp. 143-151, 2014.
[31] F. Aydin and Z. Aslan, "Classification of Neurodegenerative Diseases using Machine Learning Methods," *International Journal of Intelligent Systems and Applications in Engineering,* pp. 1-9,2017.
[32] S. Dimitriadis, D. Liparas and ADNI, "How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database", *Neural Regeneration Research*, vol. 13, no. 6, pp. 962-970, 2018.
[33] D. Zhang, Y. Wang, L. Zhou, H. Yuan and D. Shen, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," *NeuroImage,* vol. 55, no. 3, pp. 856-867, 2011.
[34] D. Wang et al., "Application of multimodal MR imaging on studying Alzheimer's disease: a survey," *Current Alzheimer Research*, vol. 10, no. 8, pp. 877–892, 2013.
[35] L. Mesrob et al., "DTI and structural MRI classification in Alzheimers disease DTI and Structural MRI Classification in Alzheimer's Disease," *Advances in Molecular Imaging,* vol. 2, pp. 12-20, 2012.
[36] X. Tang, Y. Qin, J. Wu, M. Zhang, W. Zhu and M. I. Miller, "Shape and diffusion tensor imaging-based integrative analysis of the hippocampus and the amygdala in Alzheimer's disease," *Magnetic Resonance Imaging,* vol. 34, no. 8, pp. 1087-1099, 2016.
[37] P. Vemuri et al., "Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies," *NeuroImage,* vol. 39, no. 3, pp. 1186-1197, 2008.
[38] J. Dukart, K. Mueller, A. Villringer, F. Kherif, B. Draganski, R. Frackowiak and M. L. Schroeter, "Relationship between imaging biomarkers, age, progression and symptom severity in Alzheimer's disease," *NeuroImage: Clinical,* vol. 3, pp. 84-94, 2013.
[39] X. Zhu, H. I. Suk and D. Shen, "A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis," *NeuroImage,* vol. 100, pp. 91-105, 2014.
[40] T. Schneider, "Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values", *Journal of Climate*, vol. 14, no. 5, pp. 853-871, 2001.
[41] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging,* vol. 21, no. 11, pp. 1421-1439, 11 2002.
[42] H. Zhao, H. Liu and Y. Fu, "Incomplete multi-modal visual data grouping", *Proc. Int. Joint Conf. Artif. Intell.*, pp. 2392-2398, 2016.
[43] W. Deng, D. Liu and Y. Dong, "Feature Selection and Classification for High-Dimensional Incomplete Multimodal Data", *Mathematical Problems in Engineering*, vol. 2018, pp. 1-9, 2018.
[44] I. Beheshti and H. Demirel, "Feature-ranking-based Alzheimer's disease classification from structural MRI", *Magnetic Resonance Imaging*, vol. 34, no. 3, pp. 252-263, 2016.

[45] Y Fan, D Shen and C. Davatzikos, "Classification of structural images via high-dimensional image warping robust feature extraction and SVM" in 8th International Conference on International Conference on Medical Image Computing and Computer-Assisted Intervention, Berlin Heidelberg: MICCAI, pp. 1-8, 2005.
[46] I. Beheshti, H. Demirel and H. Matsuda, "Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm", *Comput. Biol. Med.*, vol. 83, pp. 109-119, Apr. 2017.
[47] P. Johnson, L. Vandewater, W. Wilson et al., "Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease", *BMC Bioinformat*., vol. 15, no. 16, 2014.
[48] S. Q. Liu et al., "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease", *IEEE Trans. Biomed*. Eng., vol. 62, no. 4, pp. 1132-1140, Apr. 2015.
[49] J. Ramrez et al., "Ensemble of random forests one vs. rest classifiers for MCI and AD prediction using ANOVA cortical and subcortical feature selection and partial least squares", J. Neurosci. Methods, vol. 302, pp. 47-57, 2018.