Smart Media Journal / Vol.12, No.3 / ISSN:2287-1322
https://dx.doi.org/10.30693/SMJ.2023.12.3.30

Diagnosis of Alzheimer's Disease using Wrapper Feature Selection Method

비슈나비 라미네니, 권구락
(Vyshnavi Ramineni , Goo-Rak Kwon)

Abstract

Alzheimer's disease (AD) symptoms are being treated by early diagnosis, where we can only slow the symptoms and research is still undergoing. In consideration, using T1-weighted images several classification models are proposed in Machine learning to identify AD. In this paper, we consider the improvised feature selection, to reduce the complexity by using wrapping techniques and Restricted Boltzmann Machine (RBM). This present work used the subcortical and cortical features of 278 subjects from the ADNI dataset to identify AD and sMRI. Multi-class classification is used for the experiment i.e., AD, EMCI, LMCI, HC. The proposed feature selection consists of Forward feature selection, Backward feature selection, and Combined PCA & RBM. Forward and backward feature selection methods use an iterative method starting being no features in the forward feature selection and backward feature selection with all features included in the technique. PCA is used to reduce the dimensions and RBM is used to select the best feature without interpreting the features. We have compared the three models with PCA to analysis. The following experiment shows that combined PCA &RBM, and backward feature selection give the best accuracy with respective classification model RF i.e., 88.65, 88.56% respectively.

Keywords : Alzheimer's disease | Early Mild Cognitive impairment | Late Mild Cognitive Impairment | Principal Component Analysis | Restricted Boltzmann Machine

I. INTRODUCTION

Alzheimer's disease is the most common form of dementia. Memory loss due to dementia and the inability to recognize the world around us are frightening experiences. Our physical memory stores the entire history of our lives and plays an essential role in defining our character and identity. Over 50 million people are affected and expected to be 10 million by 2050[1,2] as per World Alzheimer's report. Dementia is not a single illness, but a group of symptoms caused by damage to the brain. AD is one of the neurodegenerative disorders which will cause dementia and the main symptom would be memory loss, typically over the course of weeks to months [3-5] Diagnosis of Alzheimer's disease still progresses insufficiently due to the various illness or physical and mental changes shown by the patient.

In this condition, a computational-based analysis tool is an optimistic goal and helps in the diagnosis of Alzheimer's disease at the initial stage. In present years, Machine learning greatly improves health care due to its large-scale integration of data speed. Machine learning algorithms are used for

* This study was supported by research funds from Chosun University, 2022.

Manuscript : 2023.02.17
Revised : 2023.04.07
Confirmation of Publication: 2023.04.10
Corresponding Author : Goo-Rak Kwon e-mail : grkwon@chosun.ac.kr

computational and statistical techniques for training and predicting using the given data. Using these Machine learning techniques[6,7] makes it easy to expose the patterns in the data to distinguish the diagnostics subjects and scenarios. Machine learning techniques have significantly improved in terms of medical application and show success in the [8,9]detection and prediction of different diseases.

Concerning Alzheimer' s disease(AD), they are 4 different stages for identification of the disease from mild cognition impairment (MCI) to severe stage AD. Observing diffusive morphological[10] changes in the brain i.e; Gray and white matter to determine the stage of AD. In other words, the shrinkage of the brain cells. The Alzheimer' s disease patient' s brain has more shrinkage than a healthy brain. Reasons for the cause of still unknown, however, they are reported as aging, genetics, and environmental issues. the purpose of diagnosis of AD is to slow down the progress of AD though they are no permanent cure for AD. so, it' s critical to detect Alzheimer' s disease at an early stage.

In the present paper, sMRI intends to perform AD classification, the structural MRI influences cumulative loss and results in the compression of the neuropil which shows the volume of cortical and subcortical thickness as a biomarker for Alzheimer' s disease. Intending to detect AD, the present paper has used cortical and sub-cortical. While working on different models it was accepted to see time consumption, complexity, and other issues that are during the implementation of the model. The present paper focused on overcoming the accuracy and making the model simple for medical images.

The feature extracted from this process is used for classification in the analysis of AD. The present paper used both subcortical and cortical features for the classification which is usual to see the time consumption and complexity. This paper uses combined Principal Component Analysis (PCA)[11,12] and RBM to overcome this to a certain extent. PCA is used for the reduction of dimensions of the data given to the model, which helps to train and visualize the model faster without facing any complexity. Later that reduced dimensions result is given to the input for RBM to select the features. On the other hand, wrapping methods[13] have their unique importance in selecting the feature sequentially by comparing each accuracy and selecting the sequence of features with good performance.

This experiment signifies the performance of some traditional Machine learning approaches Support vector machine (SVM), [14-16] k-Nearest neighbor (k-NN), and Random Forest (RF) using multi-class classification [17,18] with AD, HC, EMCI, and LMCI. The reason for the proposed method is to get the feature selection in the classification model for Alzheimer' s disease with multiclass classification using prediction class and actual class. Support vector machine (SVM) is taken for classification in considering regression problems and overfitting problems. K-Nearest neighbor (K-NN) is used because the labeled data to obtain and achieve better accuracy with a simple algorithm structure.

The results obtained from the three models forward feature selection, backward feature selection, and PCA & RBM are compared with PCA as PCA has been the most powerful tool in the data analysis and reduces the dimensions within the dataset the same time retaining much information.

II. METHODOLOGY

2.1 sMRI Dataset

For diagnosis, the Alzheimer's disease, Alzheimer's disease neuroimaging initiative database (ADNI) (http://adni.loni.usc.edu/) has been taken. ADNI project created in 2003, made a public-private partnership to examine AD at the early stage. ADNI Consist of a different combination of data like MRI, PET, and other biomarker images with neuro-physical assessment. ADNI data used in this performing this paper is found in ADNI serve data.

2.2 Subjects

Experimental data are taken from the ADNI website, which has thousands of subjects with different age structures. It is authorized by the Institutional review board (IRB). We have taken about 278 subjects from the pre-processed images from different patients which have fulfilled the ADNI protocol.

Table 1. Subject Report

Group	No.of Subjects	Age Range
AD	58	76.65 ± 8.6
LMCI	73	72.80 ± 6.9
EMCI	75	74.83 ± 6.1
HC	72	79.83 ± 5.7

The dataset was divided into 4 categories:

(1) AD with 58 subjects: Age ranging from 76.65 ± 8.6

(2) LMCI with 73 subjects: Age ranging from 72.80 ± 6.9

(3) EMCI with 75 subjects: Age ranging from 74.83 ± 6.1

(4) HC with 72 subjects: Age ranging from 79.83 ± 5.7

Table 1 shows the demographic information about the participants. The data set has been split in (a 75/25) ratio for the evaluation. Which have 75% for training the data and 25% for testing were allocated.

2.3 Feature Extraction

A total, of 930 features have been used during this study which has taken from subcortical and cortical segmentation. Free surfer software has been used to automate workflow which performs the stage of pre-processing step to get the designed result of brain parcellation images of the data of each subject's space.

The extracted data have been normalized from the preprocessing – zero mean and variance using scalar f_x. Normalizing the data help clear out the abnormality in the data, by doing so the analysis can be complicated Normalized matrix $\chi(i,j)$:

$$Xnorm = \frac{x(i,j) - mean\ (x_j)}{std(x_j)} \qquad (1)$$

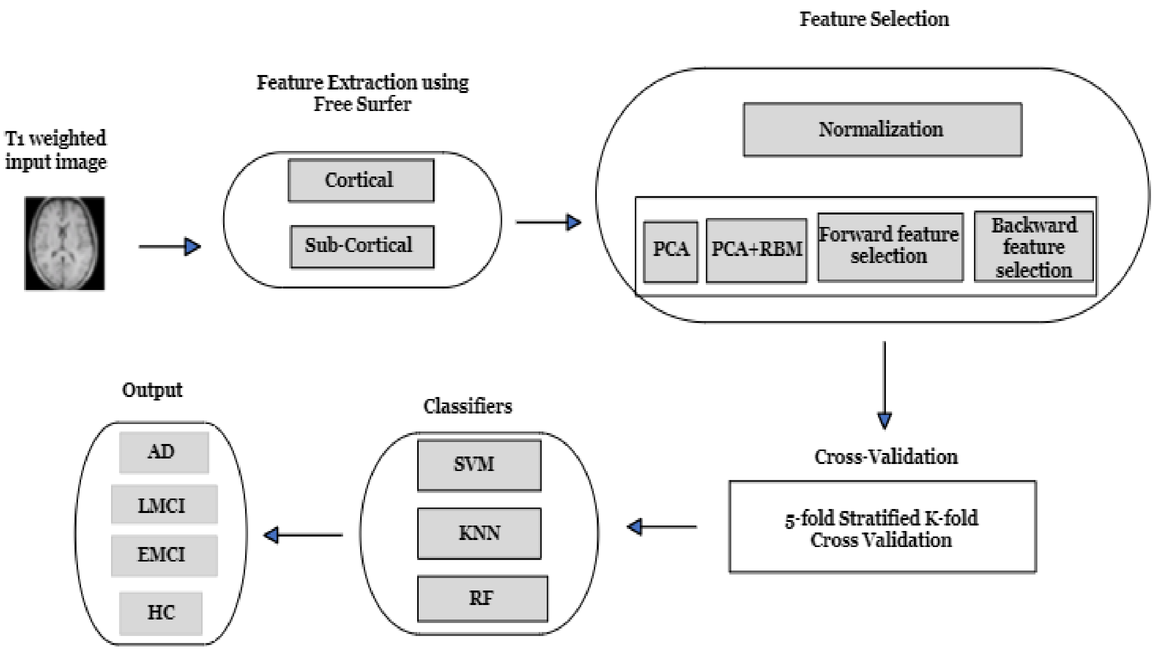


Fig 1. Block diagram of the diagnosis process

2.4 Feature selection

In Combine PCA & RBM, PCA is commonly used to reduce the dimensional sample from higher to lower the sample/features. The dimensional feature can be reduced from 2D to 1D plane using this method. therefore, we have utilized this technique in this case to minimize the feature of both subcortical and cortical features extracted from the free surfer toolbox. PCA builds with starting feature and mapped features of the dataset in d−dimension space in linear combination to a K− dimensional subspace using that k less than d. PC obtains Variables k, except the variation, which is already accounted to all subsequent components, every one PC is addressed to the maximum.

Below is the formula that can be used for computing PC_5(2):

$$Pc_I = a_1x_1 + a_2x_2 + \cdots \qquad (2)$$

PCA can be used extensively for feature selection and just for dimension reduction and later that output is given as the input for RBM used for the feature selection approach. Where RBM is Restricted Boltzmann Machine, it is used to classify the feature without great effort for interpreting the desired feature in this process we do not need to handcraft the relevant features.

Forward and Backward feature selection methods come in the wrapper type of feature selection algorithm. These techniques create the model by referring to a subset of input features and selecting the best ones according to their performance. Forward feature selection considers the subset in a forward manner and Backward feature selection considers the subset in a backward manner in an iterative method. Forward feature selection starts with no features selected in the model whereas the backward feature selection model selects all the features and eliminates the least relevant during the process. Wrapping methods are unique in terms of feature selection. Both models approach by assessing all the possible sets of features for the estimate criterion.

2.5 Classification method

After the feature was extracted from the respective subcortical and cortical regions, which undergoes the stage of normalization before proceeding to the specific feature selection. Later with selected features, can perform the classifications stage, which can discriminate the Alzheimer's disease or not. The flow of the experiment was shown the Figure 1. To perform the classification stage, we have taken three machine-learning models SVM, KNN (k-nearest neighbor), and RF (Random forest). The algorithm was implemented using the Scikit-Learn package in the python programming language.

III. EXPERIMENT RESULT & DISCUSSION

3.1 Performance and evaluation parameters

Table 2. Multiclass confusion matrix

	Prediction classification				
Actual classification	classes	AD	LMCI	EMCI	HC
	AD	**TP**	F_{AL}	F_{AE}	F_{AH}
	LMCI	F_{LA}	**TP**	F_{LE}	F_{LH}
	EMCI	F_{EA}	F_{EL}	**TP**	F_{EH}
	HC	F_{HA}	F_{HL}	F_{HE}	**TP**

Using the confusion matrix from table 2 the multi-class classification is evaluated. Model performance will be estimated using SVM, KNN, and Rf classifiers. Each classifier is responsible for the prediction of the correct number of outputs in the form of a matrix. It can be further divided into truly positive (TP), true negative (TN), false positive (FP), and false negative (FN).

In table 2 we can see the understanding of the calculation mathematically. True negative and true positive indicates the correctly identified controls; False positive and false negative represents the incorrectly identified controls.

Evaluation and accuracy is the parameter in multi-class classifiers that compute the confusion matrix.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$\text{precision} = \frac{TP}{TP+FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{5}$$

However, accuracy may not be accurate because of the unstable class distribution. So, adding on Precision(4), Recall(5), and F1-score. Sensitivity for predicting group accuracy (3) and Recall for the absence of the group's accuracy whereas the F1-score is the harmonic mean of precision and recall.

3.2 Classification results and Discussion

The results were obtained by performing the classification using SVM, K-NN, and RF. The measurements used are cortical and sub-cortical features to get the result. The feature extraction results using PCA & RBM, Forward feature selection, and Backward feature selection are shown in

Table(3,4,5,6) respectively compared with PCA . By seeing the accuracy, we can say that when compared to forward feature selection and backward feature, Backward feature selection has better accuracy and also when compared with PCA in RF classifier with 81.49% in wrapping methods its 84.45% and 88.56% respectively.

Table 3. Classification results using PCA

Classifier's	PCA		
	ACC%	PRE%	RECA%
SVM	79.48	76.66	84.69
K-NN	80.73	78.25	83.67
RF	**81.49**	77.62	89.87

Table 4. Classification results using PCA and RBM

Classifier's	PCA and RBM		
	ACC%	PRE%	RECA%
SVM	85.69	83.89	90.78
K-NN	83.67	80.45	89.32
RF	**88.65**	85.82	93.23

Table 5. Classification results using Forward feature selection

Classifier's	Forward feature selection		
	ACC%	PRE%	RECA%
SVM	80.51	74.68	83.30
K-NN	75.90	73.91	79.89
RF	**84.45**	81.87	88.56

Table 6. Classification results using Backward feature selection

Classifier's	Backward feature selection		
	ACC%	PRE%	RECA%
SVM	83.30	82.31	87.90
K-NN	79.89	76.04	84.78
RF	**88.56**	83.33	89.36

NOTE: ACC: Accuracy; PRE : Precision; RECA: Recall

The features selected from subcortical and cortical using classification and results experiment have been performed in a python environment. In most cases, all techniques will be performed well but the proposed Backward feature selection with random forest classifier shows an accuracy of 88.56% and PCA and RBM with random forest classifier show 88.65% accuracy in the multi−class classification. Similarly, with classification RF, all the models performed better with 81.49%, 88.65%, 84.45%, and 88.56%.

PCA and RBM, and forward and backward feature selection are compared with PCA in the present paper. Although wrapping methods like forward and backward feature and PCA & RBM selection shows good results when compared with PCA in different classifier models. Where PCA with K−NN classifier shows better results than wrapping methods.

Ⅳ. CONCLUSION

In the present paper, where the Backward feature selection and PCA and RBM show better accuracy ishowparison to PCA. Backward feature selection where the process of selecting the most relevant features with all features included in the model. In pca & rbm where principal component analysis is responsible for dimension reduction and RBM is responsible for picks is the best combination of features. We have compared the RBM technique with the wrapping method let alone PCA in identifying the AD from other classes EMCI, LMCI, and HC (early; late; healthy respectively). Taking the dataset from ADNI, In the experiment, used subcortical and cortical features and used the free surfer toolbox. The selected feature then performed the classification stage to distinguish the output. SVM, K-NN, and Random forest have been used for classification. However, the presented work has more improvement in terms of data selection where used only subcortical and cortical, the proposed method has slightly better accuracy leaving room for future study.

ACKNOWLEDGMENT

This study was supported by research funds from Chosun University, 2022. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense, award number W81XWH-12-2-0012). The funding details of ADNI can be found at: http://adni.loni.usc.edu/about/funding/

REFERENCES

[1] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons, "Machine learning in medicine: a practical introduction," *BMC Med. Res. Methodol.*, vol. 19, no. 1, p. 64, Mar. 2019

[2] A. Association, "2015 Alzheimer's disease facts and figures," *Alzheimers Dement.*, vol. 11, no. 3, pp. 332–384, 2015

[3] R. A. Sperling, et al., "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimers Dement.*, vol. 7, no. 3, pp. 280–292, May 2011

[4] B. C. Dickerson, et al., "The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals," *Cereb. Cortex*, vol. 19, no. 3, pp. 497–510, Mar. 2009

[5] G. Mirzaei and H. Adeli, "Machine learning techniques for diagnosis of alzheimer disease, mild cognitive disorder, and other types of dementia," *Biomed. Signal Process. Control*, vol. 72, p. 103293, Feb. 2022

[6] T. M. Schouten, et al., "Individual classification of Alzheimer's disease with diffusion magnetic resonance imaging," *NeuroImage*, vol. 152, pp. 476–481, May 2017

[7] A. H. Syaifullah, A. Shiino, H. Kitahara, R. Ito, M. Ishida and K. Tanigaki, "Machine Learning for Diagnosis of AD and Prediction of MCI Progression From Brain MRI Using Brain Anatomical Analysis Using Diffeomorphic Deformation," *Front. Neurol.*, vol. 11, 2021

[8] H. Kaur and V. Kumari, "Predictive modelling and analytics for diabetes using a machine learning approach," *Appl.*

Comput. Inform., vol. 18, no. 1/2, pp. 90–100, Jan. 2020

[9] R. Prajapati, U. Khatri and G. R. Kwon, "An Efficient Deep Neural Network Binary Classifier for Alzheimer's Disease Classification," *in 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pp. 231–234, Apr. 2021

[10] S. Basheera and M. Satya Sai Ram, "A novel CNN based Alzheimer's disease classification using hybrid enhanced ICA segmented gray matter of MRI," *Comput. Med. Imaging Graph.*, vol. 81, p. 101713, Apr. 2020

[11] R. K. Lama, J. Gwak, J.–S. Park and S.–W. Lee, "Diagnosis of Alzheimer's Disease Based on Structural MRI Images Using a Regularized Extreme Learning Machine and PCA Features," *J. Healthc. Eng.*, vol. 2017, p. e5485080, Jun. 2017

[12] F. U. R. Faisal, U. Khatri and G.–R. Kwon, "Diagnosis of Alzheimer's Disease using Combined Feature Selection Method," *J. Korea Multimed. Soc.*, vol. 24, no. 5, pp. 667–675, 2021

[13] H. Das, B. Naik and H. S. Behera, "A Jaya algorithm based wrapper method for optimal feature selection in supervised classification," *J. King Saud Univ. – Comput. Inf. Sci.*, vol. 34, no. 6, Part B, pp. 3851–3863, Jun. 2022

[14] R. K. Lama and G.–R. Kwon, "Diagnosis of Alzheimer's Disease Using Brain Network," *Front. Neurosci.*, vol. 15, 2021

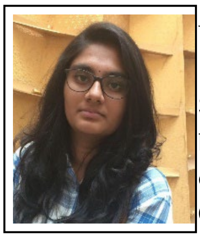
[15] Md. K. Hasan, Md. A. Alam, D. Das, E. Hossain and M. Hasan, "Diabetes Prediction Using Ensembling of Different Machine Learning Classifiers," *IEEE Access*, vol. 8, pp. 76516–76531, 2020

[16] S. Alam, M. Kang, J.–Y. Pyun and G.–R. Kwon, "Performance of classification based on PCA, linear SVM, and Multi–kernel SVM," *in 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 987–989, 2016

[17] M. Liu, D. Zhang, S. Chen and H. Xue, "Joint Binary Classifier Learning for ECOC–Based Multi–Class Classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2335–2341, Nov. 2016

[18] M. Grandini, E. Bagli and G. Visani, "Metrics for Multi–Class Classification: an Overview," *arXiv*, Aug. 13, 2020

Authors



Vyshnavi Ramineni

She has completed her bachelor's degree at the department of Electronics and Communication Engineering in NEC, India, in 2019. Currently, she is pursuing for the Master degree at the department of Information and Communication Engineering in Chosun University, Gwangju, Republic of Korea. Her research interests include Digital Signal processing, Machine learning using Medical Imaging.

Goo–Rak Kwon

He received a Ph.D at the Department of Mechatronic Engineering, Korea University, in 2007. He has been a professor with Chosun University, Since 2017. Currently he has been a director in Basic Research Laboratory of NRF, since 2019. His research interest includes medical image analysis, A/V signal processing, video communication, and applications. He is a senior member of the IEEE.