Detection and Classification of Mild Cognitive Impairment Disease in the Elderly using Deep Learning

Latifah Alamer [1], Ebtesam Shadadi [2]

[1] *Department of Information Technology and Security, College of Computer Science & Information Technology, Jazan University, Saudi Arabia.*
[2] *Department of Computer Science, College of Computer Science & Information Technology, Jazan University, Saudi Arabia.*

laalamer@jazanu.edu.sa[1], ashedadi@jazanu.edu.sa[2]

Abstract *- Elderly people served nation better and public authorities are in a position to secure their tranquility and better living conditions. The future of such people has extended with mechanical progressions and study tells that the elderly populace will turn out to be twofold in the year. The major concern for elder people is that, as they get older diseases related to cognitive impairment (Alzheimer, Vascular Dementia, and Dementia) started to begin and it's quintessential to determine those early stages by healthcare specialists. As the advancements of emerging technology are revolutionizing, the usage of Deep Learning, a class of Machine learning brings a huge potential to these fields. As a result, this research offers the following steps for an effective DL model for rapid recognition of cognitive impairment (CI): a) Data was obtained on 244 subjects from two repositories: According to the Alzheimer's Disease Neuroimaging Initiative (ADNI) website, 123 entries came from ADNI, 121 entries came from AD Repository Without Borders, and 121 entries came from ADNI, b) Preprocessing were done to remove anomalies from the raw data were the selection of instances, selection of clinical scores, imputation of missing values and Data Imbalance stages are taken care, c) Feature extraction was fuzzy logic will be used for extracting certain features for the election procedure, d) Feature Selection Using Recursive Feature Elimination (RFE) and finally e) Convolutional Neural Networks for Classification (CNN). In the research, the CNN-RFE method is superior to other state-of-the-art models (accuracy of 0.96, sensitivity of 0.97, specificity of 0.88, detection rate of 0.95, TPR of 0.95, and FPR of 0.5).*

Keywords *- Alzheimer's disease, Cognitive Impairment, Convolutional Neural Network, Classification, Deep learning, Dementia, Machine learning*

DOI Number: 10.48047/nq.2022.20.19.NQ99119 **NeuroQuantology2022;20(19): 1312-1323**

1. Introduction

Symptoms of dementia include memory loss, caused by neurodegeneration from the death of neurons. In the year 2050, there will be 131 million people living with dementia, up from the current 47 million [1]. MCI is considered a crucial factor in Alzheimer's disease (AD). The Alzheimer's Disease Neuroimaging Initiative (ADNI) database further divided mild cognitive impairment into early and late forms (EMCI and LMCI) [2]. When dementia is caught and treated early, it can be improved through lifestyle changes (such as improved nutrition) and neurocognitive enrichment, as well as pharmaceutical treatments to slow or even stop the progression of symptoms [3].Neuropsychological testing [4, 5] is a critical clinical criterion for identifying MCI and AD. Fluid and picture biomarkers may be employed in specialist healthcare situations to augment clinical testing such as cerebrospinal fluid markers, p-tau, and picture biomarkers [6]. Normal controls (CN), EMCI, LMCI, and AD have all been classified using neurocognitive testing [7, 8]. Several studies have used logistic regression and machine learning techniques to discover the best classifiers [9]. MRI and

genetic data have also been used in conjunction with neurocognitive measures in some studies [10-12].

The number of people affected by MCI is growing exponentially, and by 2040, 640 million people are expected to be affected [2]. According to the World Alzheimer Report 2014 (Figure 1), individuals with hypertension in midlife (individuals aged 40-64 years old) are likely to develop vascular dementia in later life. Dementia is characterized by blocked or decreased blood flow to the brain. Alzheimer's disease is a symptom of diabetes-related brain abnormalities. The power of blood pushing against each other within our veins is excessively high when we have hypertension and it hurts our hearts and veins. This makes the cells work harder, and therefore less effective. This makes the cells work harder, and therefore less effective. In a recent study published in a journal named Neurology Trusted Source, it was revealed that older people have higher blood pressure which is likely to lead to tangles and plaques in the brain. A possible link between SBP and AD, and specifically tangles, has been found.

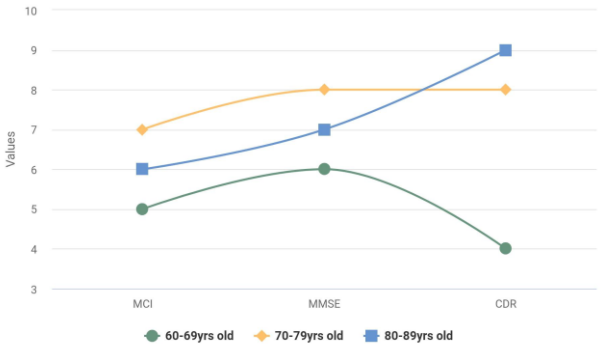

Fig. 1. The overall rate of MCI, MMSE, and CDR over various age periods

Computer vision has received greater interest in the identification of Alzheimer's illness in recent times because of the rapid advancement of artificial intelligence. Deep learning has risen to prominence as the most popular discipline of machine learning since it can tackle problems that traditional methods can't. Deep learning systems have dominated the field of medical imaging in recent years (Nguyen and Chu, 2020; Prakash, et al., 2021; Puttagunta and Ravi, 2021; Shone, Nguyen, Vu Dinh & Shi, 2018) [13-16] and the feature was also extracted from medical data for calculating the occurrence of Alzheimer's disease (Nguyen and Chu, 2020; Prakash et al., 2021.

This paper focuses on bringing an effective model for detecting cognitive impairment which following are the objectives:

- Detection and classification of Cognitive Impairment using deep learning
- Deep Learning model for accurately detecting the CI in elderly people
- Feature extraction and selection using fuzzy and RFE respectively
- Finally classification of CI using Convolutional Neural Network and thereby classify as Alzheimer, Vascular Dementia, and Dementia.

2. Literature Review

According to Gorji and Kaabouch (2019) [17], they intend to apply deep learning to MRI results to distinguish between healthy people and MCI groups. Machine learning's deep learning branch is among the most powerful. Convolutional neural networks (CNNs) to determine the most effective were utilized to classify persons into healthy, EMCI, or LMCI categories. This research used 600 MRIs, with 200 patients being controlled normal (CN), 200 being EMCI, and 200 being LMCI. In this research, we used 70% of the data to train our system and 30% of the data for testing it. The CN and LMCI groups had the best overall classification accuracy with 94.54 percent in the sagittal view, according to the data. Further, the pair EMCI/LMCI achieved 93.96 percent accuracy and the pair CN/EMCI recorded 93.00 percent accuracy, respectively.

According to Li et al. (2021) [18], Deep Learning Genetics (DLG) offers three steps: quality control, DNA polymorphism coding, and classification. The DLG models were created using the ResNet architecture, and the comparison was made to simple convolutional neural network classifications. An equal ratio of 9:1 was randomly assigned to each of the two training/validation and test groups. The model was cross-validated fivefold. Compared to traditional GWAS analysis, they compared DLG model classifications among the three groups.

Hutchins et al. (2021) [19] propose deep learning (DL) method that provides resting-state scalp electroencephalogram (EEG) data to classify persons having Alzheimer's disease (AD), and moderate cognitive impairment (MCI), and healthy aging (HA). Depending on the 10-20 system orientation, the graphs were tiled topography maps. This pipeline was used to process resting-state EEG data from 52 age-matched AD patients (82.4±4.6 years old), 37 MCI subjects (78.5±5.0 years old), and 52 HA subjects (79.5±6.1 years old). As a result, there are 16197 topographical images in the data set.

Kang et al. (2019) [20] assessed the reliability of an algorithm they developed using multicenter NPT data. According to the results of 14,927 neuropsychological exams (Seoul Neuropsychological Screening Battery), 14

927 patients were categorized into three categories based on cognitive abilities: normal cognition (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD) (ADD). The machine learning model was trained using 46 variables in TensorFlow (www.tensorflow.org). They put it to the test on ten datasets that were chosen at random, and the accuracy of their forecasts was examined.

3. Methodology

Figure 2 depicts the overall architecture of the proposed system in which the following are the stages; a) Data Collection: The Alzheimer's illness Repository Without Borders website and the ADNI's website were used to collect data on 244 items, including 123 items and 121 topics from the ADNI website. b) Preprocessing was done to remove anomalies from the raw data where the selection of instances, selection of clinical scores, imputation of missing values, and Data Imbalance stages are taken care of, c) Feature extraction where fuzzy logic will be used for extracting certain features for the selection procedure, d) Feature Selection Using Recursive Feature Elimination (RFE) and finally e) Classification using Convolutional Neural Network (CNN).

3.1 Data Collection

In this study, data were collected from two databases: Alzheimer's illness Repository Without Borders (ARWIBO), and the Alzheimer's illness National Institute (ADNI).

The ARWIBO [22] consists of 29 AD subjects (10 males, 19 females; age ± SD = 71.3 ± 4.15 years, education level = 8.36 ± 9.30, range = 0–18), 35 cognitively normal HC subjects , 36 patients with mAD , and 25 patients with aAD. The ADNI [21] consists of 32 AD subjects (17 males, 15 females; age ± SD = 72.13 ± 4.22 years, education level = 9.42 ± 3.79, range = 0–18), 28 cognitively normal HC subjects , 25 patients with mAD and 36 patients with a AD.

For this study, the demographics are shown in Tables 1 and 2 for the 244 subjects out of 121 from ARWIBO and 123 from ADNI. The Welch independent - sample t was used to examine the clinical features and demographics of the research groups. The alpha value in this study was set at 0.05, which is a common significant criterion. The Ad network had the lowest educational level, as per the examination of the groups. This research used a 70:30, training: testing ratio to obtain unbiased performance estimates. We used to develop these skills code to train. The developed algorithm was utilized on the rest of the dataset. Moreover, the groups' efficiency was rated on the accuracy, precision, etc based on a single test sample [23, 24].

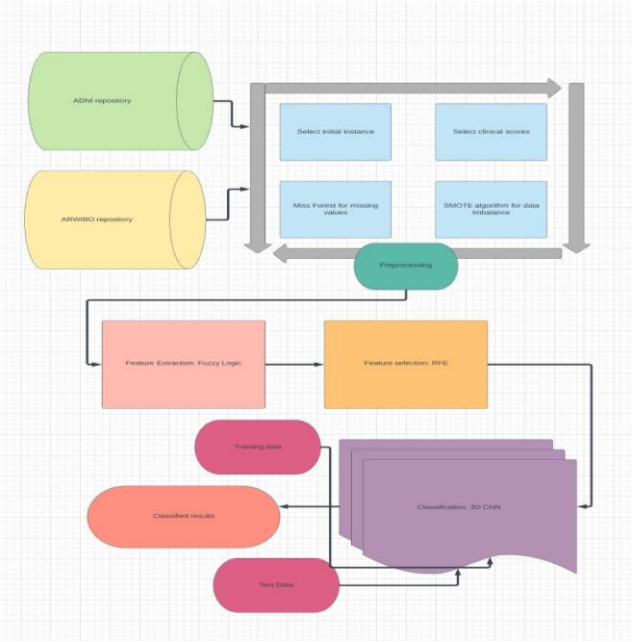

Fig. 2. The overall architecture of the proposed system

Table 1. Demographics of ARWIBO

Group	Subject No.	Age	Male	Female	Education
AD	28	71.23 ± 14.08 [58–81]	11	19	8.36 ± 3.79 [0–19]
EMCI	34	69.8 ± 7.10 [50–83]	15	21	7.66 ± 4.22 [0–19]
LMCI	25	69.44 ± 3.21 [60–80]	11	16	7.98 ± 5.22 [0–19]
HC	33	65.58 ± 9.13 [59–84]	17	18	10.07 ± 3.44 [0–23]

Table 2. Demographics of ADNI

Group	Subject No.	Age	Male	Female	Education
AD	32	72.15 ± 4.22 [55–79]	18	16	9.42 ± 3.79 [0–19]
EMCI	25	69.15 ± 8.36 [49–84]	13	14	7.98 ± 4.21 [0–19]
LMCI	38	67.10 ± 5.82 [61–82]	23	17	8.03 ± 7.11 [0–19]
HC	28	64.03 ± 6.46 [64–85]	19	11	11.42 ± 6.57 [0–23]

3.2 Preprocessing

The data collected from the 2 repositories are subjected to the preprocessing module which comprises the selection of instances, clinical scores, Imputation of

missing values, and data balancing.

Selection of instances; The collected data from ADNI and ARWIBO have the instances with their visit range varying from one to twelve. From the clinic diagnosis, we observe that 50% of MCI patients are converted to AD and other types within the time duration of 24 to 36 months. So, we preprocess the data to have the instances with a minimum of 24 to 36 months between baseline [25].

Clinical scores; The paired t-test is used as a statistical procedure to compare the CDRSUM score of the instances at two different visits. We choose CDRSUM in the statistical analysis because it is a measure of cognitive dysfunction and dementia severity. To perform statistical analysis, we assume the null hypothesis and calculate the paired t-test scores between two different visits using Equation 1. Null Hypothesis: The mean difference in CDRSUM score between two visit times is zero and the p-value ≥0.05.

$$t = \frac{m}{s/\sqrt{n}} \qquad (1)$$

Where, t is the t-statistic; m is the mean of the group; s is the standard deviation, and n is the sample/group size. The obtained p-value is 2.3e-16. At a 96% confidence interval, we reject the null hypothesis, concluding that CDRSUM scores are significantly different between the two datasets.

Missing Values; The missing values observed in the obtained data from 2 datasets are computed using the Miss Forest Imputation method [26]. Let the considered data be in the form of a matrix with columns corresponding to clinical scores Pj, where [j=1,2,...m] and rows corresponding to instances Vi, where [i=1,2,..n] of normal and demented subjects. Miss forest begins with the initial calculation of missing values by the mean imputation method. In the next step, the number of missing values for Pj's are determined and sorted based on their increasing order. Where NAPj represents the count of missing values (NA) for clinical score Pj. Based on the number of missing values, rearrange Pj with NAP1 < NAP2 < NAP3 .. < NAPm. Fit a random forest for the observed values of Px and Pj as the target and predictor variables respectively. Then predict one of the missing values of Px based on trained random forest and repeat it for every attribute Pj. This process continues for several iterations until the difference between previous imputed matrix I and newly imputed matrixi+1 increases for the first time.

Algorithm 1; Miss Forest

Input: Original data with N*M matrix with missing values.
Output: Imputed data with N*M matrix without missing values
Step 1: The initial calculation of missing values is performed by using the mean imputation method.
Step 2: Let data be of N * M where N is the no. of records and M is the no. of attributes.
Step 3: Each of the attributes Aj has several missing values where j = 1,2,3 … m.
Step 4: The missing values should be rearranged in ascending order.
Step 5: An analysis of a random forest is conducted using both observed values for the target variable, Ax, and the independent variable, Aj.
Step 6: Based on the trained random forest, the missing value of Ax is predicted.
Step 7: If the difference between the matrix and newly created matrixi+1 meets the stopping criteria, then imputation stops

Data Imbalance: Data Imbalance: SMOTE, or Synthetic Minority Oversampling Technique, is a hybrid sampling approach, in which samples are synthesized from minority classes, while samples from majority classes are removed at random until a balanced ratio is obtained between the two classes. Algorithm 2 describes the synthesis of new minority samples by considering each sample Mi from the minority class and its nearest neighbors. The nearest neighbors are selected based on the minimum Euclidean distance concerning each minority sample considered [27]. After that, the variance between every minority sample's feature vector and one of its randomly chosen neighbors is determined. The estimated variance is then multiplied by a value between 0 and 1 at random. A new sample is created by adding this result to the Mi. It is followed by the under-sampling of samples from the majority class.

Algorithm 2; SMOTE

Input: Data, with M number of Samples from Minority Classes, Amount of synthetic samples N%.
Output: N% of M Synthetic Minority Samples. Initialize: Number of nearest neighbor k, n = N% of M.
Step 1: K nearest neighbors are computed for each instance Mi from Minority Class M.
Step 2: Take one of Minority Sample Mi's K nearest neighbors at random.
Step 3: Diff: Calculate the difference between the nearest neighbor and Mi.
Step 4: Alpha: A random number between 0 and 1.
Step 5: New Sample: Mi + (Diff *Alpha)
Step 6: Repeat steps 2 - step 5 till n new samples are synthesized.

3.3 Feature Extraction

The first step is to do the data fuzzification process. There are different attributes in each dataset entry. This study turns the character into language terms such as VL, L, M, H, and VH in the first phase. If a dataset contains 'P' records, each with 'n' attributes,

$$P_i = [f_{i1}, f_{i2}, \dots \dots f_{in}] \quad (2)$$

The ith record in the dataset is indicated by i and the jth feature of the record is represented by fij. The -type method is used to provide membership values to each feature in the database. Based on the five linguistic terms, the -type assigns fuzzified values. If there are 'n' features per record, the feature vector comprises 5*n fuzzified characteristics.

$$f_{ij} = \left[\mu_{VL}(f_{ij}), \mu_L(f_{ij}), \mu_M(f_{ij}), \mu_H(f_{ij}), \mu_{VH}(f_{ij})\right] \quad (3)$$

Initially, the information is turned into fuzzy membership numbers. The size of the feature space grows dramatically when each characteristic is allocated five membership values, which adds complexity. To solve this difficulty, the current study uses principal component analysis to store all of the important features while discarding the ones that aren't. Furthermore, the dataset has a large number of missing values. The missing data is filled using imputation methods, and all significant characteristics are restored after the fuzzification process [28, 29].

The very first stage consists of two steps: a) data fuzzification and b) extraction of features. The model's neural network features are handled in the second step. A suitable neural network is developed after the extracting features model to execute data classification.

Algorithm 3: Fuzzy Logic for feature extraction

Input: y1, y2,..., yn is a data sequence in which Tamura characteristic xi is associated with the corresponding texture picture, while n is the number of textures. Output: The tamura characteristic has five membership values, each of which is a term placed on the tamura characteristic. 1. Take the first step. Let c0 equal to min(y1, y2,..., yn) and d6 equal to max(y1, y2,..., yn). Calculate d1, d2,..., d5 in the following manner:

$$dj = d0 + \frac{j}{6} * (d6 - d0)$$

Create membership functions using d1, d2,..., d5 as class centres. 2nd step V should be set V = 0 Use one of the following rules to update each element uij for each datum xj: where vij, 1 <=I <=5 and 1 <=j <=n, is the membership value of yj in the ith linguistic term
Rule 1 If yi ≤ d1, vi = 1 and vik≠1 = 0
Rule 2. If cj < yi ≤ dj+1, compute vij = (cj+1 − xi)/(cj+1 − cj)
Rule 3. If yi > d5, vi,k≠5 = 0 and ui,5 = 1

Step 3. Calculate d1, d2,..., d5 using the given equations:

$$ci = \frac{\sum_{j=1}^{n} yijxj}{\sum_{j=1}^{n} yij}$$

The algorithm will terminate if d1, d2,..., d5 remain unchanged; else, go to Step 2.

3.4 Feature Selection

To find the characteristics that allow better discrimination between CI sufferers and the control group, we used a feature selection algorithm based on a wrap-per assessment function called recursive feature elimination [30].

RFE uses a greedy approach to get the overall top subset of characteristics using the backward elimination approach. At every iteration, the RFE technique produces models from the entire collection of accessible features, identifying the worst-performing characteristic. After that, given the leftover qualities, it develops future models until all of the elements have been investigated. In the worst situation, RFE considers N2 subsets if the data contains N characteristics. The feature subset with the greatest performance among those examined is produced by the algorithm as an output. The accuracy, calculated using the K-fold cross-validation method and the xgboost classifier, was utilized as an evaluation function. It's worth noting that the RFE iterative algorithm favors characteristics with low correlations. Let us use an example to demonstrate the method that allows RFE to select a small number of linked characteristics. Let f1 and f2 be two correlated

characteristics in Si, the subset of variables remaining after the i–th iteration, and the procedure will test S0=Si f1 and S00 =Si f2 [31]. Because f1 and f2 are correlated, S0 and S00 have comparable results. As a result, RFE will most likely exclude one of them from the subsetSi+1.

Algorithm 4: RFE- Feature Selection

```
Inputs ;
 Training set S
 Set of p feature G = { G1.....Gp}
 Ranking Method T (S, Q)
Output;
 C is the final rank
Code ;.
 Repeat for i in [ 1 : Q ]
 Rank set F in using T(S, G)
 G* <---- last ranked feature f
 C ( q - i + 1) <------ g*
 G <---- g- g*
```

3.5 Classification

A deep learning network consists of artificial neurons, which are made up of multiple layers. Each 'neuron' is fully connected to all the 'neurons' below it via weighted connections, replicating the way a person's brain works [32]. The model learns, (i) the inputs that represent the information being fed into it, (ii) the output of the network differs from the target value, (iii) The output layer's error signal, and (iv) the error signal from the error signal back to the input layer. Steps 1 through 4 are repeated until the error is as small as possible by modifying the parameters of the deep learning architecture. Furthermore, trained networks allow for the prediction of new observations without having to examine original data.

In Figure 3, you can see that a convolutional layer is alternated with a max pool layer, which is immediately followed by a few fully connected layers. The convolutional layer identifies the major features of a picture, storing them in a 'feature map' that represents their relationship, and then runs the algorithm. In general, nonlinear layers (or activation layers) are placed after every convolutional layer. The model and network's nonlinear properties are improved, but the convolutional layers are left untouched since it just converts all negative activations to 0.

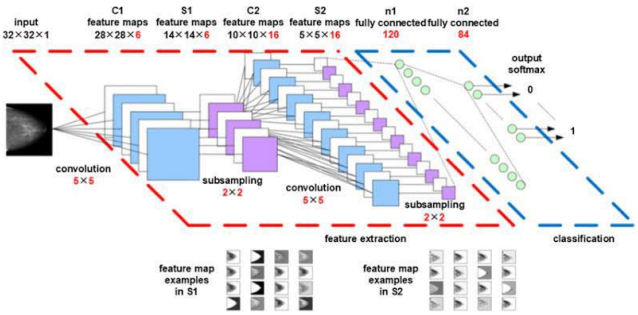

Fig. 3.The regular architecture of CNN

Due to its quick development rate, the Rectified Linear Unit has become the most extensively used activation function. Following that, a pooling (or subsampling) stage that does spatial downsampling. In the last layers of a system, all the neurons from the previous stages are fully connected. A typical deep learning algorithm is more efficient because CNNs reduce the number of parameters that need to be assessed [33].

To account for the volumetric aspect of MR images, a network connection based on digital 3D convolutions was created. These were normalized 3D T1-weighted pictures, and the outcomes were subject groups. Its architecture consists of 12 convolutional layers (2 blocks with 50 kernels of 5 x 5 x 5 with alternate strides 1 and 2 and 10 blocks with 100 to 1600 kernels of 3 x 3 x 3 with alternate strides 1 and 2), a Rectified Linear Unit (activation layer), a fully connected layer, and one output layer (logistic regression), as well as a Rectified Linear Unit (activation layer), a fully connected layer. As a replacement for the max-pooling layers ('all convolutional network' [34]), a regular neural network with a stride of 2 was used, which distinguished it from standard CNNs. The "all convolutional network" is a simple design that yields excellent results without the use of complex activation functions, reply normalization, or max-pooling (Springenberg et al., 2015).

As shown in Figure 4, the 3D CNN (Number 4) has been verified on a large set of subjects and controls with selected binary classifications. The datasets assessed were AD versus HC, stable MCI (s-MCI) versus HC, AD versus c-MCI, AD versus s-MCI, c-MCI versus s-MCI, c-MCI versus s-MCI, c-MCI versus s-MC. The following aspects are included in each classification: (i) training, (ii) validation, and (iii) testing (Fig. 2). Initially, each classification dataset was divided into two groups: a large dataset for training and validation (90 percent of the photos) and a smaller dataset for testing (10 percent of the pictures). On photographs selected for training and validation (rather than testing), data augmentation was performed to create additional false images that can avoid overfitting that could result

from a fully linked layer filling all of the attributes.

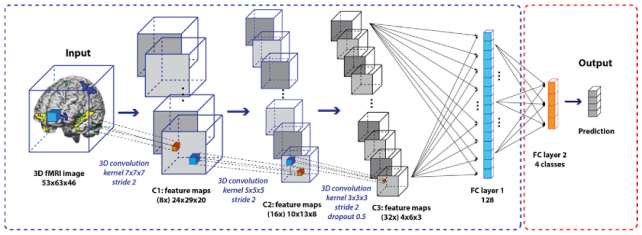

Fig. 4. The architecture of 3D CNN

4. Performance Measures

The model is implemented over hardware specifications like Nvidia RTX 3080 Ti, 12th Gen Intel Core i7, and Windows 10 OS and software specifications like PyTorch, an open-source python library for building the model, and Google Collaboratory, an open-source environment for performing DL models. Evaluation of the proposed model (RFE-3DCNN) is acted over various other models such as VGG16, Googlenet, ANN, Resnet, Alexnet, Densenet, SVM, Regular CNN over various measures like accuracy, sensitivity, specificity, TPR, FPR, AUC score, computation time and memory utilization.

Table 3 depicts the overall analysis of various models regarding the accuracy, sensitivity, and specificity. Figure 5(a,b) shows a graphical representation of various models over 2 datasets in which the proposed model performs well for the ARWIBO dataset when compared to other models. The accuracy, sensitivity, and specificity are 97,98,98 percent respectively and for the ADNI dataset the accuracy, sensitivity, and specificity are 96, 97,98 percent respectively.

Table 3. Overall analysis over accuracy, sensitivity & specificity

Models	Dataset	Accuracy	Sensitivity	Specificity
VGG16		83	86	91
Googlenet		88	90	95
ANN		79	84	90
Resnet	ADNI	80	85	92
Alexnet		85	90	94
Densnet		89	93	96
SVM		87	91	95
Regular CNN		91	95	97
RFE-3DCNN (Ours)		97	98	98
VGG16		81	89	93
Googlenet		85	90	95
ANN		80	86	90
Resnet	ARWIBO	76	83	89
Alexnet		83	89	95
Densnet		89	92	94
SVM		83	87	92
Regular CNN		94	96	97
RFE-3DCNN (Ours)		96	97	98

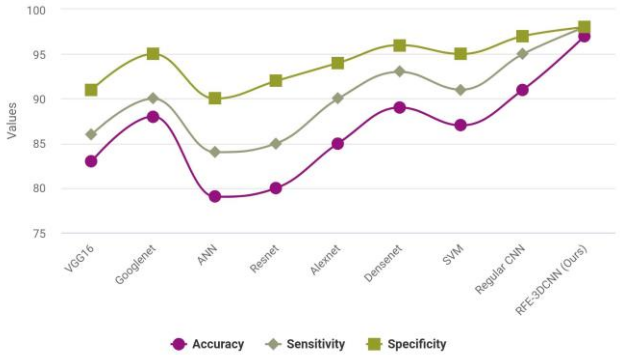

Fig. 5a. Models vs Accuracy, Sensitivity, and Specificity over ADNI dataset

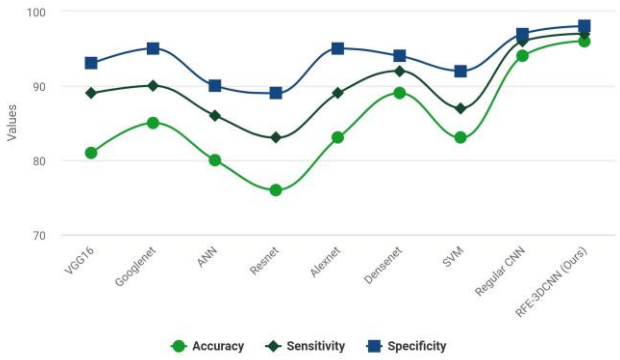

Fig. 5b. Models vs Accuracy, Sensitivity, and Specificity over ARWIBO dataset

In Table 4, the accuracy, recall, and F1-score of various models are evaluated. Figure 6(a,b) shows graphical representation of various models over 2 datasets in which the proposed model outperforms (accuracy; 0.94, recall; 0.86, F1-score; 0.93) for ADNI dataset and (precision; 0.95, recall; 0.88, F1-score; 0.91) for ARWIBO dataset when compared to other models.

Table 4. An overall analysis of Precision, Recall, and F1-score

Models	Dataset	Precision	Recall	F1-score
VGG16		79	72	82
Googlenet		82	79	85

ANN		81	69	80
Resnet	ADNI	73	63	83
Alexnet		84	78	85
Densnet		91	80	88
SVM		86	78	82
Regular CNN		89	81	89
RFE-3DCNN (Ours)		94	86	93
VGG16		80	81	83
Googlenet		87	83	85
ANN		81	75	73
Resnet	ARWIBO	77	70	76
Alexnet		85	84	83
Densnet		90	88	88
SVM		89	81	82
Regular CNN		92	88	84
RFE-3DCNN (Ours)		95	89	90

Fig. 6a. Models vs Precision, Recall, and F1-score under ADNI dataset

Fig. 6b. Models vs Precision, Recall, and F1-score under ARWIBO dataset

Table 5 depicts the overall analysis of various models over TPR, FPR, and Detection rates. Figure 7(a,b) shows graphical representation of various models over 2 datasets in which the proposed model outperforms (TPR; 0.93, FPR; 0.7, Detection rate; 0.92) for ADNI dataset and (TPR; 0.95, FPR; 0.5, Detection rate; 0.94) for ARWIBO dataset when compared to other models.

Table 5. Overall analysis under detection rate, TPR, and FPR

Models	Dataset	TPR	FPR	Detection rate
VGG16		85	15	82
Googlenet		87	13	85
ANN		79	21	78
Resnet	ADNI	76	24	79
Alexnet		83	17	89
Densnet		89	11	84
SVM		86	14	82
Regular CNN		91	9	89
RFE-3DCNN (Ours)		93	7	92
VGG16		87	13	83
Googlenet		84	16	80
ANN		80	20	77
Resnet	ARWIBO	78	22	75
Alexnet		87	13	82
Densnet		90	10	86
SVM		85	15	81
Regular CNN		92	8	91
RFE-3DCNN (Ours)		95	5	94

Fig. 7a. Models vs TPR, FPR, and Detection rate over ADNI dataset

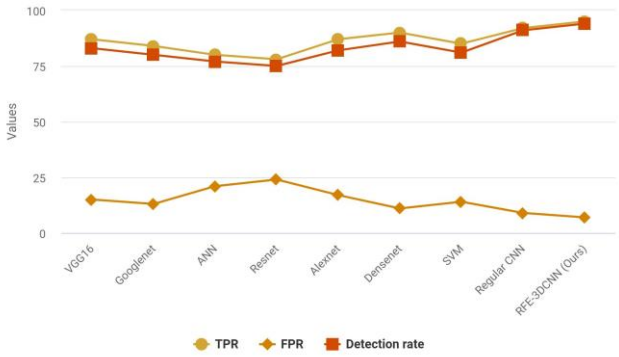

Fig. 7b. Models vs TPR, FPR, and Detection rate over ARWIBO dataset

Table 6 depicts an overall analysis of various models over AUC score, Computation time. Figure 8(a,b,c,d) shows a graphical representation of various models over 2 datasets in which Figure 8a displays the AUC score over various models over ADNI, Figure 8b displays the AUC score of various models over the ARWIBO dataset, Figure 8c displays Computation time of various model and finally figure 8d memory utilization of various models. From all aspects, the proposed model outperforms (AUC: 0.96, Computation time; 4.1, memory utilization; 43%).

Table 6. Overall analysis under AUC score and CT

Models	Dataset	AUC score	Computation time
VGG16		87	9.6
Googlenet		89	13
ANN		80	6.3
Resnet	ADNI	72	11.2
Alexnet		83	8.2
Densnet		91	7.3
SVM		84	5.7
Regular CNN		90	4.9
RFE-3DCNN		93	4.3
(Ours)			
VGG16		85	8.8
Googlenet		87	12
ANN		78	7.7
Resnet	ARWIBO	70	10
Alexnet		81	8.3
Densnet		89	6.2
SVM		82	5.1
Regular CNN		92	4.7
RFE-3DCNN (Ours)		96	4

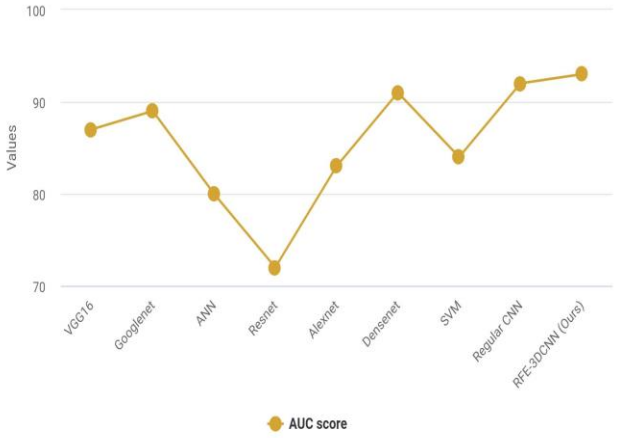

Fig. 8a. Models vs AUC score over ADNI dataset

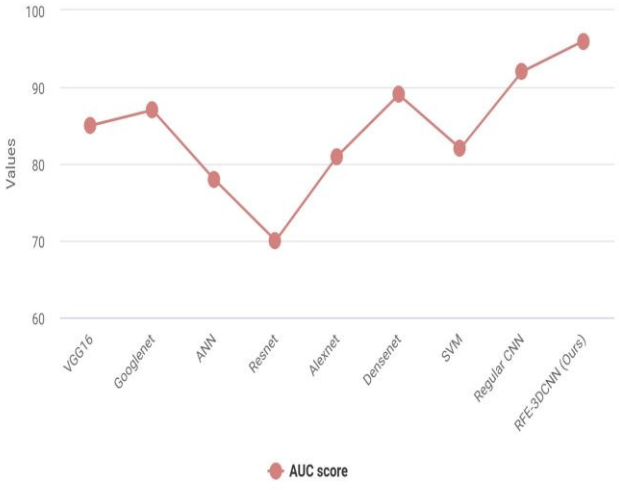

Fig. 8b. Models vs AUC score over ARWIBO dataset

eISSN 1303-5150

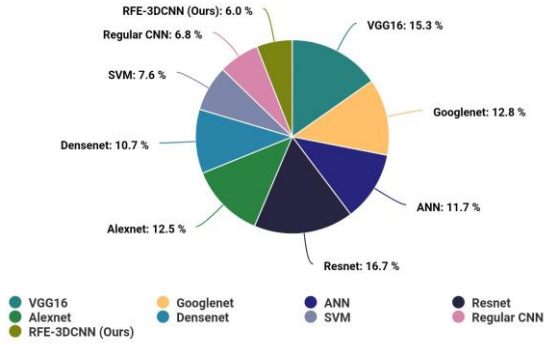

Fig. 8c. Models vs Computation time overall analysis

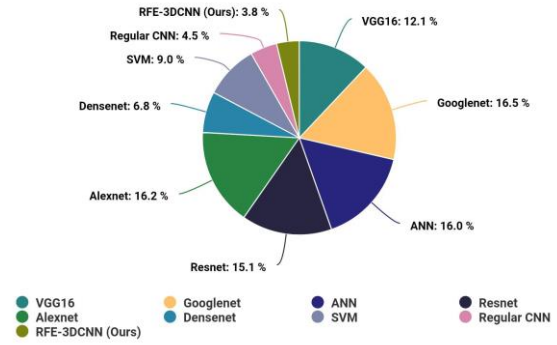

Fig. 8d. Models vs Memory utilization overall analysis

5. Conclusion

This paper brings an effective way of analyzing, detecting, and classifying various Mild Cognitive Impairments in a much more accurate and effective way in elderly people. As when the study progresses, it includes processes like preprocessing, extraction of features, selections, and categorization, as well as the application of successful strategies that produce effective results. Tests have shown that the proposed combination strategy beats different state-art approaches. As DL is getting revolutionized in these medical areas so much, various deep learning approaches can be performed. Finally, this paper will be much helpful for healthcare specialists (psychiatrists or psychologists) and also helpful for research specialists to dig deep and get an understanding and also brings even more advanced integrated models for effective detection of MCI in elderly people as well in a better way.

References

[1] Arvanitakis Z, Shah RC, Bennett DA (2019) Diagnosis and management of dementia: review. JAMA 322, 1589-1599.

[2] Edmonds EC, McDonald CR, Marshall A, Thomas KR,Eppig J, Weigand AJ, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW, Alzheimer's Disease Neuroimaging Initiative (2019) Early versus late MCI: Improved MCI staging using a neuropsychological approach. Alzheimer's Dement 15, 699-708.

[3] Mossello E, Ballini E (2012) Management of patients with Alzheimer's disease: Pharmacological treatment and quality of life. Ther Adv Chronic Dis 3, 183-193.

[4] Tierney MC, Yao C, Kiss A, McDowell I (2005) Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. Neurology 64, 1853-1859

[5] Battista P, Salvatore C, Berlingeri M, Cerasa A, Castiglioni I (2020) Artificial intelligence and neuropsychological measures: The case of Alzheimer's disease. Neurosci Biobehav Rev 114, 211-228.

[6] Molinuevo JL, Ayton S, Batrla R, Bednar MM, Bittner T,Cummings J, Fagan AM, Hampel H, Mielke MM, Mikulkis A, O'Bryant S, Schltens, Sevigny J, Shaw LM, Sores HD, Tong G, Trojanowski JQ, Zeterberg H, Blenow K (2018) Current state of Alzheimer's fluid biomarkers. Acta Neuropathol 136, 821-853.

[7] Hemmy LS, Linskens EJ, Silverman PC, Miller MA, Talley KMC, Taylor BC, Ouellette JM, Greer NL, Wilt TJ, Butler M, Fink HA (2020) Brief cognitive tests for distinguishing clinical Alzheimer-type dementia from mild cognitive impairment or normal cognition in older adults with suspected cognitive impairment. Ann Intern Med 172, 678-687.

[8] Grober E, Hall C, McGinn M, Nicholls T, Stanford S, Ehrlich A, Jacobs LG, Kennedy G, Sanders A, Lipton RB (2008) Neuropsychological strategies for detecting early dementia. J Int Neuropsychol Soc 14, 130-142

[9] Zhu F, Li X, Tang H, He Z, Zhang C, Hung G-U, Chiu P-Y, Zhou W (2020) Machine learning for the preliminary diagnosis of dementia. Sci Program 2020, 1-10.

[10] Kim JW, Byun MS, Sohn BK, Yi D, Seo EH, Choe YM, Kim SG, Choi HJ, Lee JH, Chee IS, Woo JI, Lee DY (2017) Clinical dementia rating orientation score as an excellent predictor of the progression to Alzheimer's disease in mild cognitive impairment. Psychiatry Investig 14, 420-426.

[11] Abd Razak MA, Ahmad NA, Chan YY, Mohamad Kasim N, Yusof M, Abdul Ghani MKA, Omar M, Abd Aziz FA, Jamaluddin R (2019) Validity of screening tools for dementia and mild cognitive impairment among the elderly in primary health care: A systematic review. Public Health 169, 84-92

[12] Gillis, C., Mirzaei, F., Potashman, M., Ikram, M. A., & Maserejian, N. (2019). The incidence of mild cognitive impairment: A systematic review and data synthesis. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 11, 248-256.

[13] H. Nguyen, N.N. Chu An introduction to deep learning research for Alzheimer's disease, IEEE Consumer Electronics Magazine, 10 (3) (2020), pp. 72-75

[14] D. Prakash, N. Madusanka, S. Bhattacharjee, C.H. Kim, H.G. Park, H.K. Choi Diagnosing Alzheimer's Disease based on Multiclass MRI Scan using Transfer Learning Techniques Current medical imaging (2021)

[15] M. Puttagunta, S. Ravi, Medical image analysis based on deep learning approach, Multimedia Tools and Applications, 80 (16) (2021), pp. 24365-24398

[16] N. Shone, N.Tran Nguyen, P. Vu Dinh, Q. Shi A deep learning approach to network intrusion detection, IEEE Transactions on Emerging Topics in Computational Intelligence, 2 (1) (2018), pp. 41-50

[17] Taheri Gorji, H., & Kaabouch, N. (2019). A deep learning approach for diagnosis of mild cognitive impairment based on MRI images. Brain sciences, 9(9), 217.

[18] Li, L., Yang, Y., Zhang, Q., Wang, J., Jiang, J., & Neuroimaging Initiative. (2021). Use of deep-learning genomics to discriminate healthy individuals from those with Alzheimer's disease or mild cognitive impairment. Behavioural Neurology, 2021.

[19] Huggins, C. J., Escudero, J., Parra, M. A., Scally, B., Anghinah, R., De Araújo, A. V. L., ... & Abasolo, D. (2021). Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. Journal of Neural Engineering, 18(4), 046087.

[20] Kang, M. J., Kim, S. Y., Na, D. L., Kim, B. C., Yang, D. W., Kim, E. J., ... & Youn, Y. C. (2019). Prediction of cognitive impairment via deep learning trained with multi-center neuropsychological test data. BMC medical informatics and decision making, 19(1), 1-9.

[21] Online source: https://www.google.com/url?sa=t&source=web&rct=j&url=https://adni.loni.usc.edu/data-samples/access-data/&ved=2ahUKEwiE5POg8Mz2AhX8S2wGHacQCVQQFnoECDYQAQ&usg=AOvVaw3pYpdGy4LKM9KuWU2PmawL

[22] Online source; https://www.google.com/url?sa=t&source=web&rct=j&url=https://www.gaaindata.org/partner/ARWIBO%23:~:text%3DARWIBO%2520is%2520a%2520cross%252Dsectional,with%2520Alzheimer%27s%2520disease%2520(AD).&ved=2ahUKEwiOieb48Mz2AhUJT2wGHfaaBm8QFnoECAQQBQ&usg=AOvVaw04NuphlJpwnurMdHy4j2EC

[23] Huckvale, E. D., Hodgman, M. W., Greenwood, B. B., Stucki, D. O., Ward, K. M., Ebbert, M. T., ... & Miller, J. B. (2021). Pairwise Correlation Analysis of the Alzheimer's Disease Neuroimaging Initiative (ADNI) Dataset Reveals Significant Feature Correlation. Genes, 12(11), 1661.

[24] Pinaya, W. H., Scarpazza, C., Garcia-Dias, R., Vieira, S., Baecker, L., F da Costa, P., ... & Mechelli, A. (2021). Using normative modeling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-

eISSN 1303-5150

sectional multi-cohort study. Scientific reports, 11(1), 1-13.

[25] Sarraf, S., Desouza, D. D., Anderson, J. A., & Saverino, C. (2019). MCADNNet: recognizing stages of cognitive impairment through efficient convolutional fMRI and MRI neural network topology models. IEEE Access, 7, 155584-155600.

[26] Iddi, S., Li, D., Aisen, P. S., Rafii, M. S., Thompson, W. K., & Donohue, M. C. (2019). Predicting the course of Alzheimer's progression. Brain informatics, 6(1), 1-18.

[27] Yan, Y., Liu, R., Ding, Z., Du, X., Chen, J., & Zhang, Y. (2019). A parameter-free cleaning method for SMOTE in imbalanced classification. IEEE Access, 7, 23537-23548.

[28] Haouas, I., Moussa, M., & Douik, A. (2021, June). Classification and Identification of Alzheimer's Disease With Fuzzy Logic Method. In 2021 IEEE International Conference on Design & Test of Integrated Micro & Nano-Systems (DTS) (pp. 1-6). IEEE.

[29] Mallika, R. M., UshaRani, K., & Hemalatha, K. (2019). A Fuzzy-Based Expert System to Diagnose Alzheimer's Disease. In the Internet of Things and Personalized Healthcare Systems (pp. 65-74). Springer, Singapore.

[30] Ota, K., Oishi, N., Ito, K., Fukuyama, H., Sead-J Study Group, & Alzheimer's Disease Neuroimaging Initiative. (2015). Effects of imaging modalities, brain atlases, and feature selection on prediction of Alzheimer's disease. Journal of neuroscience methods, 256, 168-183.

[31] Kim, J., & Lee, B. (2017, July). Automated discrimination of dementia spectrum disorders using extreme learning machine and structural t1 MRI features. In 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 1990-1993). IEEE.

[32] Wu, C., Guo, S., Hong, Y., Xiao, B., Wu, Y., Zhang, Q., & Alzheimer's Disease Neuroimaging Initiative. (2018). Discrimination and conversion prediction of mild cognitive impairment using convolutional neural networks. Quantitative imaging in medicine and surgery, 8(10), 992.

[33] Wen, D., Wei, Z., Zhou, Y., Li, G., Zhang, X., & Han, W. (2018). Deep learning methods to process fmri data and their application in the diagnosis of cognitive impairment: a brief overview and our opinion. Frontiers in neuroinformatics, 12, 23.

[34] Khvostikov, A., Aderghal, K., Benois-Pineau, J., Krylov, A., & Catheline, G. (2018). 3D CNN-based classification using sMRI and MD-DTI images for Alzheimer's disease studies. arXiv preprint arXiv:1801.05968.

[35] Islam, J., Zhang, Y., & Alzheimer's Disease Neuroimaging Initiative. (2018, December). Deep convolutional neural networks for automated diagnosis of Alzheimer's disease and mild cognitive impairment using 3D brain MRI. In International Conference on Brain Informatics (pp. 359-369). Springer, Cham.

eISSN 1303-5150