

Accepted Manuscript

Accepted Manuscript (Uncorrected Proof)

Title: Detecting Mild Cognitive Impairment to Alzheimer's Disease Progression by fMRI using Convolutional Neural Network and Long-Short Term Memory

Authors: Sima Ghafoori¹, Ahmad Shalbaf^{1,*}

1. *Department of Biomedical Engineering and Medical Physics, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*

***Corresponding Author:** Ahmad Shalbaf, Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
Email: shalbaf@sbmu.ac.ir

To appear in: **Basic and Clinical Neuroscience**

Received date: 2024/11/16

Revised date: 2025/05/12

Accepted date: 2025/06/2

This is a “Just Accepted” manuscript, which has been examined by the peer-review process and has been accepted for publication. A “Just Accepted” manuscript is published online shortly after its acceptance, which is prior to technical editing and formatting and author proofing. *Basic and Clinical Neuroscience* provides “Just Accepted” as an optional and free service which allows authors to make their results available to the research community as soon as possible after acceptance. After a manuscript has been technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as a published article. Please note that technical editing may introduce minor changes to the manuscript text and/or graphics which may affect the content, and all legal disclaimers that apply to the journal pertain.

Please cite this article as:

Ghafoori, S., Shalhaf, A. (In Press). Detecting Mild Cognitive Impairment to Alzheimer's Disease Progression by fMRI using Convolutional Neural Network and Long-Short Term Memory. *Basic and Clinical Neuroscience*. Just Accepted publication Jul. 10, 2025. Doi: <http://dx.doi.org/10.32598/bcn.2025.2034.8>
DOI: <http://dx.doi.org/10.32598/bcn.2025.2034.8>

Abstract

Mild Cognitive Impairment (MCI) is the stage that happens before Alzheimer's Disease (AD) and there is a high risk of progression to AD. However, this progression is not guaranteed and there is a chance of staying at this stage. In this study, we aim to diagnose possible AD progression among MCI subjects from a combination of resting-state functional Magnetic Resonance Imaging (fMRI), clinical assessment, and demographic information for starting treatments in case of progression or reducing medical expenses in case of future stability. For this work, we have used Deep Learning methods called three-dimensional Convolutional Neural Network (CNN) and Long-Short Term Memory (LSTM). The models were developed using 266 samples from 81 MCI subjects over an average of five years between baseline and the last timepoint. Results showed, the best validation scores belonged to the CNN-LSTM model after integrating with clinical attributes based on accuracy of 92.47%. Consequently, our proposed algorithm demonstrated high performance in predicting MCI to AD progression, indicating the potential of deep learning approaches in processing fMRI data and the efficiency of integrating data types.

Index Terms: Alzheimer's Disease, Deep Learning, Convolutional Neural Network, Long-Short Term Memory, Mild Cognitive Impairment, Magnetic Resonance Imaging.

I. Introduction

Alzheimer's disease (AD) is the most customary cause of dementia in old age, which causes irreversible damage to the brain. However, if implemented in the early stages of the disease, proven treatments could be practical and weaken the process by stopping or slowing down the destruction of brain tissue [1], [2]. Accurate diagnosis of various stages of Alzheimer's is based on cerebrospinal fluid (CSF) pathology, an expensive and invasive procedure [3], [4]. Donepezil, memantine, galantamine, rivastigmine, and aducanumab are FDA-approved drugs and are prescribed for AD treatment in the early stages of the disease [5]-[8]. Mild Cognitive Impairment (MCI) is a pre-Alzheimer's stage in which some of the symptoms of Alzheimer's are poorly experienced; therefore, there is a high risk of developing Alzheimer's. However, this progression is not guaranteed, and the individual will likely return to normal or stay at this stage [9], [10]. Given that Alzheimer's effects on the brain start years prior to clinical symptoms, it is critical and attainable to predict the conversion of MCI to AD from brain information [11-12]. In this case, people can prepare for the disease from different aspects and start treatment methods as early as possible to attain the best results.

Researchers and clinicians frequently exploit brain imaging methods like magnetic resonance imaging (MRI), positron emission tomography (PET), and functional MRI (fMRI) for this task. MRI shows the brain structure with high spatial resolution. PET uses ionizing radiation to deliver an insight into the brain's anatomy and functionality with lower spatial resolution than MRI. fMRI provides knowledge about the anatomy and metabolic mechanism of the brain with higher temporal and spatial resolution than PET while being harmless at the same time [13]-[15]. Other techniques, such as magnetoencephalography (MEG), and others, have their advantages and disadvantages [15]. Finally, invasive data such as genetic information and CSF and the cheapest, most accessible, and safest data types like clinical cognitive tests, have been used to predict AD among MCI subjects.

In recent years, many studies have predicted the conversion of MCI to AD by Machine Learning (ML) methods from neuroimaging methods and clinical information with relatively good results. These researchers have been attempting to solve this problem using new ML techniques such as Deep Learning (DL) methods since these methods can discover and learn hidden patterns and give a more accurate answer [16], [17]. There are many ML research in neuroimaging-related studies, including MRI, PET, and fMRI [18]-[32], and researchers often exploit machine learning techniques to analyze manually extracted features. Among the mentioned studies, some of them are as follows: In 2017 [18], Hojjati et al. used the connection matrix and extracted graph properties from resting-state fMRI data. Then a number of these features were selected for evaluation with a Support Vector Machine (SVM). In 2018 [19], he also combined the characteristics of the same fMRI data and MRI of the subjects. In recent years, deep learning strategies have been attempting to solve this problem. In 2019 [20], Abrol extracted time-varying features from fMRI, then deployed a three-dimensional convolutional neural network (3D-CNN) to extract attributes from MRI data and, at last, trained an SVM on the combination of all features. In 2020 [32], Gao Fei et al. developed a 3D-CNN (containing six layers of convolution and three layers of max-pooling) based on MRI images. In 2020 [28], Loris Nanni planned to compare the performance of transfer learning models with a 3D-CNN in processing MRI images. To do this, she designed a simple one-layer convolutional model and fine-tuned several pre-trained models. Finally, an average ensemble model was designed on five retrained models. In 2021 [31], Xiaoxi Pan developed a novel integrated CNN-based model called Multi-view Separable Pyramid Network (MiSePyNet). In this model, slice-wise and spatial-

wise CNNs were performed on three views (axial, coronal, and sagittal) of PET images, and then all outcomes were put together for classification. Fusun Er [29] used a CNN model whose filters were calculated using an auto-encoder to extract prognostic features from each patient's volume. Then these features were delivered to an SVM model for classification. Also, some of these machine learning studies have deployed invasive data such as genetic information and CSF to predict MCI to AD [21]-[23], [30]. In 2020 [30], Lane et al. Developed an SVM model by combining information extracted from PET, MRI, CSF, genetic data, and clinical cognitive tests. Finally, in some machine learning studies, clinical cognitive tests and brain signals have been used to predict Alzheimer's [33]-[35]. In these studies, large samples were needed to obtain reliable results. In 2018 [33], Grassi et al. used clinical and demographical features. After applying the feature selection process, several machine learning algorithms were designed, from which the best model was an SVM with a kernel-based radial base function. Also, in 2019 [34], Grassi used only demographic data and cognitive clinical tests. At first, all the features were used, then three more feature sets were created by three feature selection methods, and then 13 machine learning algorithms were developed according to these four sets resulting in 52 final models. Furthermore, a weighted averaging model was designed by which the results of 52 models were averaged and presented on the test data. Finally, Mengjia Xu [36] used a general form of Graph2Gauss architecture, a DL-based method called multiple graph Gaussian embedding model (MG2G), to classify graph-based features derived from MEG regional time series.

Nevertheless, none of these studies contain enough information alone to decisively predict the prognosis from MCI to AD. Also, some studies have used invasive predictors such as CSF and genetics [21]-[23], [30]. For all we know, the current study is the first to attempt to use deep learning methods called three-dimensional CNN and Long-Short Term Memory (LSTM) in processing information from fMRI data to solve the described problem. These methods automatically extract and classify features, which is their primary upside compared to the conventional machine learning techniques. We also aim to introduce a new approach that combines the features extracted from clinical cognitive tests and neuroimaging data in a neuronal way and then examine its effectiveness in improving the results.

II. Materials and Methods

A. Dataset

In this study, clinical and fMRI data from the patients who had been in the MCI stage were gathered from the Alzheimer's Disease Neuroimaging Initiative (ADNI; available at adni.loni.usc.edu) for predicting MCI conversion to Alzheimer's. ADNI is a multicenter study designed to develop clinical, imaging, genetic, and biochemical markers for the early diagnosis of Alzheimer's disease. Since its inception, which was more than a decade ago, public-private partnerships have played a significant role in Alzheimer's research, enabling data to be shared among researchers worldwide. In ADNI, when a person is admitted according to inclusion and exclusion criteria, an initial diagnosis is made based on cerebrospinal fluid pathology. The clinical cognitive tests are approximately taken every six months, and the rest of the biomarkers are collected at time intervals of one to two years. This process continues as long as the person remains in the study and specialists constantly update the patient's diagnostic status. In this study, we intended to predict the conversion of MCI to Alzheimer's by combining clinical and fMRI data. Thus, related information from the patients who had primarily been in the MCI stage was gathered. Some of these individuals remained moderately in the MCI phase over the years or eventually returned to normal (sMCI group); however, the rest progressed to Alzheimer's after a while (pMCI group). In this paper, we did not include subjects who had been in the study for less than a year. Furthermore, because

we wanted to utilize a combination of imaging and clinical data, given the difference between recording dates, we first examined the imaging dates for each person. Then we considered the test results that were registered around the same time (maximum of one month). Therefore, the corresponding clinical data were collected for each series of imaging data on one date. If each data was not available on a date, the other data was not used. Finally, we had 81 subjects, 28 had progressed from MCI to AD, and 53 either remained in MCI or returned to normal. 266 samples from these individuals were available, of which 78 belonged to the pMCI class, and the remaining 188 samples were sMCI. Details of the data are reported in **Table I** in the two groups. Tesla Phillips 3 Magnetic Resonance Imaging Scanner was used with EPI to record functional and structural images while people were resting. The size of the functional image matrix was 64 by 64, and each 3D image of the brain volume was divided into 48 sections, each section being 3.3 mm, as well as an 80-degree rotation angle, a repetition time of three seconds, and a reflection time of 30 milliseconds. Ultimately, 140 functional images were available for each person. Moreover, we have chosen 17 clinical features from the ADNI organization, including common demographic characteristics and cognitive neurophysiological test results. The ADNIMERGE file, which the ADNI organization officially prepares, contains the essential variables of the types of data collected. Demographic information is age, gender, years of education, and marital status. Clinical cognitive tests are Sum of Boxes score of Clinical Dementia Rate (CDR-SB), Functional Activity Questionnaire (FAQ), Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale (total of 11 activity scores [ADAS11], Alzheimer's Disease Assessment Scale (total of 13 activity scores [ADAS13], Score of task 4 of Alzheimer's Disease Assessment Scale [ADASQ4]), Rey Auditory Verbal Learning Test (RAVLT) scores (immediate [RAVLT-I], Learning [RAVLT-L], forgetting [RAVLT-F], and percent-forgetting [RAVLT-PF]), Trail Making Test version B (TRABSCOR) and the total delayed recall score of the Logic Memory subtest of the Wechsler Memory Scale-Revised (LDELTOTAL) [37]-[43]. We also examined the type of MCI in primary diagnosis (early or late). Dataset Details, reported in **Table II**, shows the names and abbreviations of the 17 clinical features and their statistics in the two groups.

TABLE I : Details of the dataset. 266 samples from 81 subjects were available, from which 28 subjects were progressive MCI (pMCI) and the rest were stable MCI (sMCI). 78 samples belonged to the pMCI group and the rest of 188 samples belonged to the sMCI group.

	sMCI	pMCI	TOTAL
SUBJECTS	53 (65.43 %)	28 (34.56 %)	81
SAMPLES	188 (70.67 %)	78 (29.32 %)	266

TABLE II : 14 continuous and 3 categorical features were selected from ADNIMERGE dataset. Their abbreviation, description and percentage of missing values; mean and standard deviation (std) for continuous features, and percentage for categorical features are reported in each group.

Continuous variables	Description	pMCI			sMCI			Missing values (%)
		Mean	±	std	Mean	±	std	
AGE	Age in years	73.39	±	6.47	70.51	±	7.56	0.0
PTEDUCAT	Years of education	15.54	±	2.5	15.99	±	2.77	0.0
CDRSB	Sum of Boxes score of the Clinical Dementia Ratio	2.47	±	1.04	1.03	±	0.68	1.88
FAQ	Functional Assessment Questionnaire	6.58	±	5.19	1.62	±	2.6	2.25
MMSE	Mini-Mental State Examination	26.65	±	1.69	28.12	±	1.9	1.5
ADAS11	Sum of scores of 11 tasks of the Alzheimer's Disease Assessment Scale	13.51	±	4.05	7.68	±	3.86	1.5
ADAS13	Sum of scores of 13 tasks of the Alzheimer's Disease Assessment Scale	21.29	±	5.6	11.91	±	5.84	1.88
ADASQ4	Score of task 4 of Alzheimer's Disease Assessment Scale	6.88	±	2.13	3.72	±	2.2	1.5
RAVLT-I	Immediate score of the Rey Auditory Verbal Learning Test	30.26	±	8.59	37.32	±	10.77	2.25
RAVLT-L	Learning score of the Rey Auditory Verbal Learning Test	2.95	±	1.98	5.06	±	2.67	2.25
RAVLT-F	Forgetting score of the Rey Auditory Verbal Learning Test	5.9	±	2.11	4.49	±	2.26	2.25
RAVLT-PF	Percent Forgetting score of the Rey Auditory Verbal Learning Test	83.44	±	23.1	52.77	±	30.17	2.25
TRABSCOR	Trial making test, version B	122.28	±	64.46	97.88	±	58.19	4.13
LDELTOTAL	Total delayed recall score of the Logic Memory substest of the Wechsler Memory Scale-Revised	4.17	±	3.12	9	±	3.88	25.9
Categorical variables	Description	pMCI			sMCI			Missing values (%)
		%			%			
PTGENDER	Sex, (male and Female)							0.0
	Male	51.28			52.12			
	Female	48.71			47.78			
PTMARRY	Marital status							0.0
	Married	78.2			68.61			
	Never married	2.56			3.72			
	Widowed	19.23			12.76			
	Divorced	0.0			14.89			
DX_bl	Subtypes of MCI: Early MCI (EMCI) and Late MCI (LMCI)							0.0
	EMCI	44.87			69.14			
	LMCI	55.12			30.85			

B. Preprocessing

We first discarded the feature with losing data of more than 5%. Therefore, the LDELTOTAL feature was discarded, while other clinical features were preserved. Then the missing values were replaced for the rest of the features; we substituted the mean for numerical properties and the mode for categorical features (gender, marital status). We then normalized the numerical data and encoded categorical features

to convert them to numbers using the one-hot method. Furthermore, we discarded the first ten measurements among the functional images recorded for each individual. Then, the following steps were performed to preprocess the images: slice-timing correction to the last slice, realignment using a six-parameter rigid-body spatial transformation, EPI normalization, resampling to 3-mm isotropic voxels, detrending, smoothing using a Gaussian filter with FWHM (= 4 mm), band-pass filtering (0.01–0.08 Hz) and elimination of destructive signals such as global mean signal, six head motion parameters, the cerebrospinal fluid (CSF), and the white matter signals. Consequently, in the fMRI dataset, the original data dimensions were (64, 64, 48) with a time dimension of 140. After preprocessing step, the data dimensions were (61, 61, 73) with a time dimension of 130. We used the Python programming language version 3.8 to preprocess clinical features, SPM12 software (<http://www.fil.ion.ucl.ac.uk/spm>), and the DPARSF toolbox (<http://rfmri.org/DPARSF>) to perform imaging data preprocessing. Also, the clinical and imaging data values were normalized to be between zero and one.

C. Convolutional Neural Network

A CNN is a DL model for processing data with a grid pattern, such as images. These networks are designed to learn spatial information and have three primary layers: convolutional, pooling, and fully connected. The convolutional and pooling layers conduct the feature extraction operation, and the fully connected or dense layer is the classification layer. This network can convert and reduce input data into a form that is easy to identify and examine without losing much information. In the convolutional layer, a small matrix of numbers called kernel is applied throughout input by point multiplication to extract feature maps which are then given to a nonlinear activation function. Next, the pooling layer, which has no learnable parameters, down-sizes the results; for instance, the max-pooling layer, the most common type, holds the maximum of patches of each feature map and discards the remaining values. Finally, the last layers' output is converted to a one-dimensional vector by flatten layer and given to a (or more) fully connected layer. Each layer input is connected to the output by a learnable weight. Finally, in the output, we have a probability of belonging to each class for each case [44], [45].

D. Long-Short Term Memory

Long-Short Term Memory (LSTM) network, a type of recurrent neural network, detects temporal dependencies in input data. In an LSTM network, at any time, the node receives its input at its time, output, and hidden state from the last time. Then delivers the output and hidden state in its time. The hidden mode can be considered the node memory, which helps the LSTM remember input from a previous time. Because the node structure is more complex than normal neuronal nodes, we call it a cell. This cell has four input, forget, output, and input modulation gates. Each gate has a weight to control how much information can pass through the gate, and the activation function determines its weight. Fig. 1 exhibits the structure of an LSTM block and how it calculates its output [45, 46].

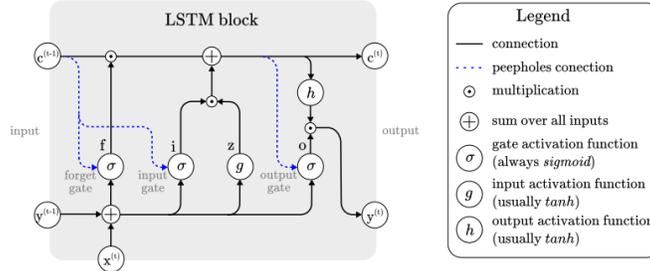


Fig. 1. Architecture of a LSTM block [46].

E. Proposed Model

Functional MRI is a set of three-dimensional images recorded for a limited time. Therefore, CNN and LSTM models could be exploited to process the fMRI spatial and temporal aspects. From what we know, the current study is the first to use DL methods for fMRI processing and combining clinical information to predict the MCI to AD conversion according to the proposed method. First, we wanted to evaluate whether the temporal aspect alone consists of enough discriminative information or not. Therefore, we used the automated anatomical atlas (AAL) to define and divide the brain into 116 regions; thus, the input of the LSTM network was a 130 by 116 matrix (130-time points and 116 features). We designed two LSTM models with 32 and 128 cells to investigate the effect of the number of LSTM layer cells on the diagnostic performance of the model. Since the output size of all models must have been the same and equal to 32, a fully connected layer with 32 neurons was added in continuation of the 128-cell layer. Then the output of both models was entered into a dense layer with two neurons and a softmax activation function for classification. Fig. 2 shows the LSTM models.

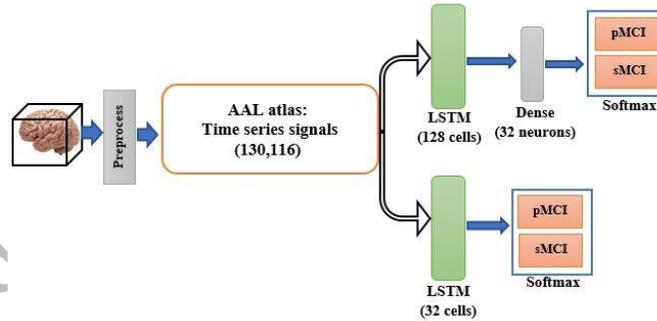


Fig. 2. The LSTM models

Second, if CNN and LSTM models are put together, the concluded model is called CNN-LSTM, composed of a CNN part to extract spatial features, followed by an LSTM layer to extract temporal information. **Fig. 3** illustrates the structure of the CNN part of the CNN-LSTM models. The convolutional part of our proposed model was three-dimensional to prevent initial information loss due to data transformation. It consisted of six convolutional layers, with a filter size of 3x3 and stride of one, and four max-pooling layers with a filter size of 2x2 and stride of two. The input of this network was single functional images (regardless of time dimension). The output of the last layer was 128 extracted features. Since 130 images were available for each sample, we utilized 130 3D-CNN models, each analyzing images of a specific time point, and concatenated their outputs to deliver to an LSTM layer (**Fig. 4**). We designed two CNN-LSTM models since we previously designed two LSTM models.

Moreover, to investigate the effect of combining clinical data on the results, in all of these final models, the output of the imaging section, which was 32 in length, was concatenated to the output of a shallow neural network layer whose input was clinical data. The dense layer had five neurons; thus, the effect of the features extracted from the images was about six times greater than the clinical features. Consequently, a 37-size vector was inserted into the classification layer. Finally, eight networks were developed (two LSTM and two CNN-LSTM models with and without combining with the clinical information). **Fig. 4** shows the CNN-LSTM model with 128 cells in its LSTM layer, combined with the clinical information.

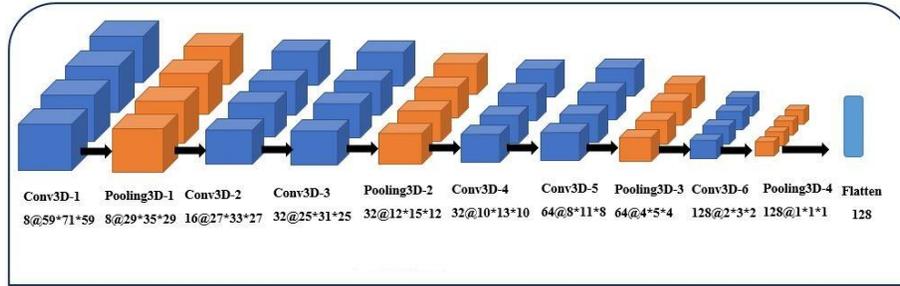


Fig. 3. The structure of the CNN part of the CNN-LSTM models.

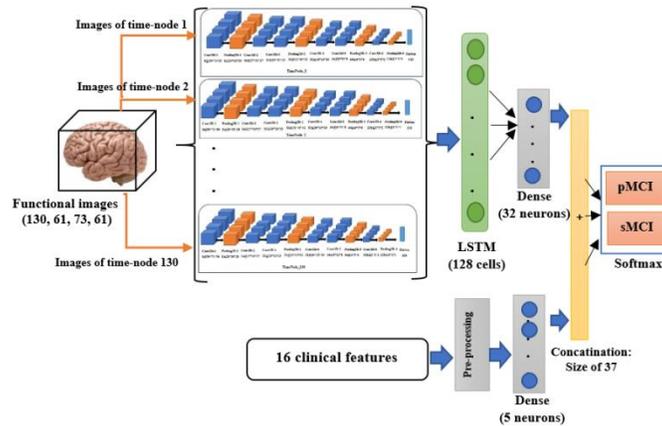


Fig. 4. The 128-cell CNN-LSTM model after combining with clinical features

The models were designed in Google Colaboratory (12 GB RAM) using the Keras framework with Tensorflow as the backend. The optimizer used to train the CNN-LSTM models was Adam, with a learning rate of 0.0001. For controlling the error of overfitting, dropout was used, its value for the outputs of the inner layers (except the first layer) and the LSTM layer (32 cells) was 20%, for the last pooling layer 30%, and it was 40% for the 128-cell LSTM layer. Also, L2 regulation was used for the dense layers (32 and 5- neurons) with a coefficient of 0.001. Due to hardware limitations, the data were delivered to the model in batches, and a maximum of three was possible for the CNN-LSTM models.

F. Training and Evaluation

We desired to train and test our models using 5-fold stratified cross-validation. In this method [47], the dataset is divided into five folds, from which four folds are the training data, and the left-out fold is the

testing data. The folds could be stratified, meaning that the ratio of each class (sMCI, pMCI) is nearly equal in each fold and the whole dataset. We needed our train and test sets to be independent of each other. Therefore, we randomly divided all the data into five parts so that each part contained about 20% of the data and all parts were independent of each other based on the patients; in other words, samples belonging to one subject were assigned to the same set. In our dataset, several samples were available for each person. Also, we attempted to assure that 30 percent of the samples in each set belonged to class one. Consequently, we had 81 subjects, 28 pMCI class, and 53 sMCI class. 266 samples from these individuals were available, of which 78 belonged to the pMCI class, and the remaining 188 samples were sMCI. In 5-fold cross validation, in testing, 16 subjects (5 pMCI and 11 sMCI) with 53 samples (15 pMCI and 38 sMCI) are used in each fold, approximately. Eventually, five training and testing subsets were produced, and all models were designed according to each training set and evaluated based on its corresponding test set. Also, we considered 10% of each training set for evaluation to measure the appropriateness of the models' design during the training process. Thus, in the end, we had five sets of training, testing, and evaluation with ratios of 70, 20, and 10 percent. It is worth noting that the clinical data preprocessing for training and test sets were performed separately to preserve the independence of the two sets. Finally, in this study, to report and compare the performance of the models, the evaluation metrics of accuracy, sensitivity, specificity, AUC, and confusion matrix were calculated in the five sets of testing.

III. Results

The mean of accuracy and loss curves in the 5-fold cross-validation training procedure of the CNN-LSTM models for the train and validation sets in four models (CNN-LSTM models with 32 and 128 cells and before and after combining with the clinical data) is shown in Fig. 5. The models were trained in about 80 iterations. We optimized the model architecture and the hyper-parameters by trial and error after examining the progress of learning based on validation and training metrics values to avoid over-fitting.

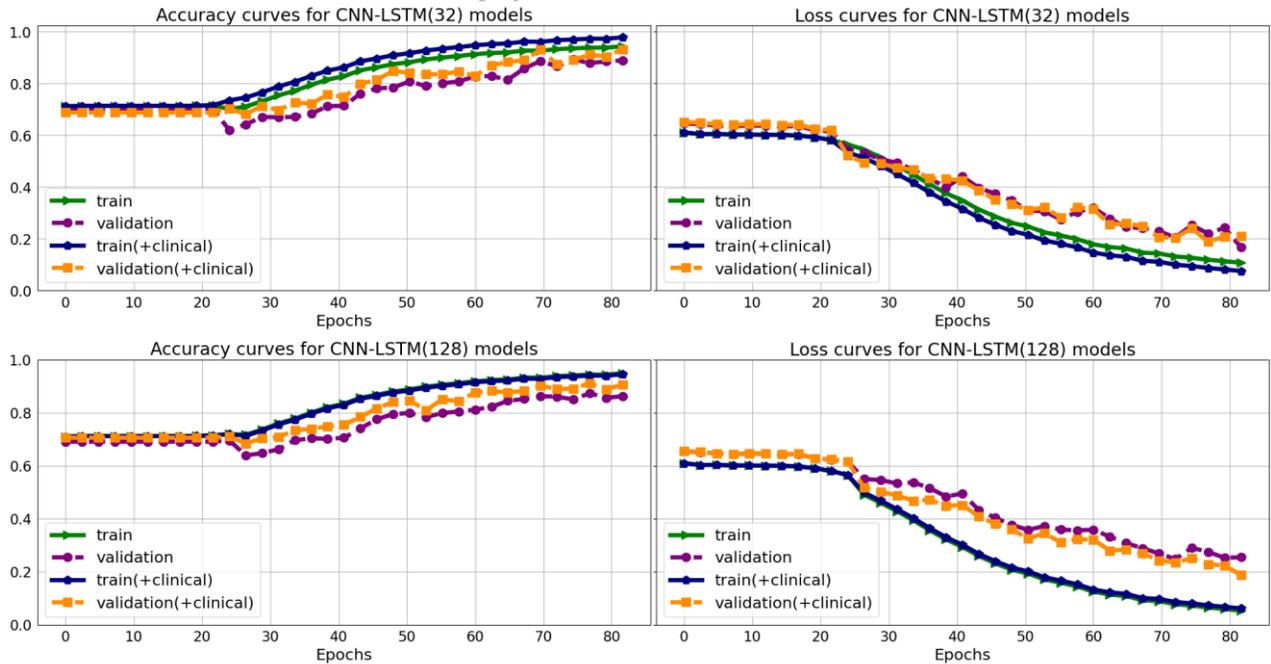


Fig. 5. Mean of Accuracy and loss curves in the cross-validation process for the CNN-LSTM model with 32 and 128 cells before and after combining with clinical data

Mean and standard deviation values of accuracy, sensitivity, specificity, and AUC metrics for the five test subsets are reported for eight models (LSTM models with 32 and 128 cells, CNN-LSTM networks with 32 and 128 cells, before and after combining with clinical data) in Table III.

The evaluation values for the 32-cell LSTM network before combining with clinical features were 71.8 (accuracy), 9.00 (sensitivity), 97.32 (specificity), and 61.51 (AUC) percent. In this model, we observed an increase in accuracy (78.94%), sensitivity (31.66%), specificity (98.4%), and AUC (68.42%) after combination with the clinical information. In the next step, by increasing the number of cells to 128 in the LSTM model before combining with clinical, the accuracy, sensitivity, specificity, and AUC increased to 74.81, 14.16, 100, and 70.07%, respectively. Finally, the 128-cell LSTM model, in combination with the clinical data, had the best performance among the LSTM models based on AUC (75.99%), accuracy (79.60%), and sensitivity (38.5%). As for the CNN-LSTM networks, we also observed that increasing the number of recurrent layer cells improved the results effectively. Validation metrics for the 32-cell CNN-LSTM model, before combination with clinical, were 86.84 (accuracy), 69.33 (sensitivity), 94.15 (specificity), and 90.62 (AUC) percent. Then, by increasing the number of cells to 128, we saw an improvement in results based on the accuracy (89.84%), sensitivity (83.33%), specificity (92.54%), and AUC (92.4%). As with the results of CNN-LSTM models, we also observed an increase in the values of evaluation parameters after adding clinical features, and both models achieved higher sensitivity, accuracy, and AUC afterward. Finally, the 128-cell CNN-LSTM model demonstrated the best results after combining the clinical data. It reached an accuracy of 92.47 %, an AUC of 96.67 %, a sensitivity of 92.33 %, and a specificity of 92.55%.

The ROC and AUC diagrams of the LSTM and CNN-LSTM models with 32 and 128 cells before and after combining with clinical data can be seen in **Fig. 6**. It is shown that the 128-cell CNN-LSTM model, combining fMRI and clinical data, demonstrated the best results. Also, it is shown and concluded that the combination with the clinical predictors made the models more sensitive to the pMCI class since the curves started from points with higher true-positive rates. Finally, confusion metrics of the most performing LSTM and CNN-LSTM models are reported in **Fig. 7** (before adding clinical information) and **Fig. 8** (after adding clinical information). It is illustrated, as well, that combining with clinical data increased the true-positive rates while decreasing false-negative rates, meaning that the models became more sensitive toward the pMCI cases after combining clinical attributes. Lastly, according to the confusion matrices, the 128-cell CNN-LSTM models were the only networks with more false-positive rates than false-negative ones. After combining with clinical data, this model misdiagnosed only three samples from about 50 test set cases. It is eventually worth noting that a false-positive rate being higher than a false-negative rate is desirable, especially in medical issues, since incorrectly diagnosing a patient as unwell is less medically expensive than the other way around.

TABLE III : Mean (%) and standard deviation of cross-validated scores of the LSTM and CNN-LSTM models with 32 and 128 cells before and after combining with clinical data.

Model	Data	AUC	Accuracy	Sensitivity	Specificity
LSTM(32)	fMRI	61.51 (± 0.11)	71.50 (± 1.21)	09.00 (± 3.21)	97.32 (± 1.7)
	fMRI + clinical	68.42 (± 0.53)	78.94 (± 1.33)	31.66 (± 4.04)	98.40 (± 1.3)
LSTM(128)	fMRI	70.07 (± 0.19)	74.81 (± 1.88)	14.16 (± 4.92)	100.0 (± 0.00)
	fMRI + clinical	75.99 (± 0.62)	79.69 (± 1.49)	38.50 (± 1.14)	96.27 (± 3.62)
CNN-LSTM(32)	fMRI	90.62 (± 0.51)	86.84 (± 1.20)	69.33 (± 3.98)	94.15 (± 1.03)
	fMRI + clinical	93.75 (± 0.88)	89.47 (± 1.90)	78.16 (± 3.34)	94.16 (± 1.94)
CNN-LSTM(128)	fMRI	92.40 (± 0.42)	89.84 (± 0.95)	83.33 (± 3.10)	92.54 (± 1.10)
	fMRI + clinical	96.67 (± 0.71)	92.47 (± 1.71)	92.33 (± 2.42)	92.55 (± 1.95)

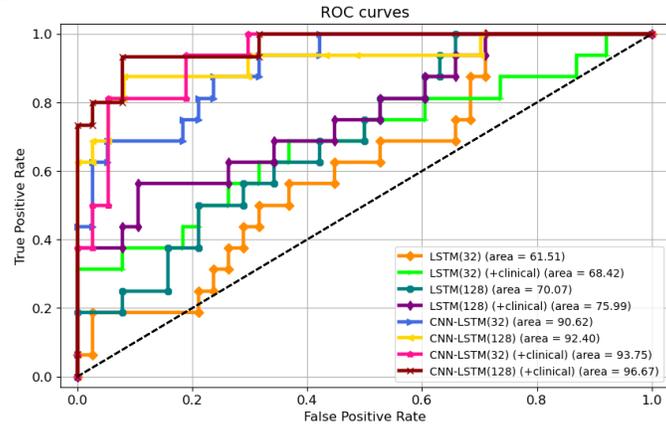


Fig. 6. ROC/AUC curves of the LSTM and CNN-LSTM models with 32 and 128 cells before and after combining with clinical data.

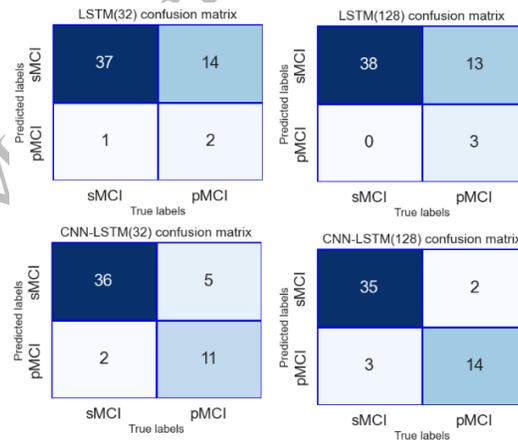


Fig. 7. Confusion matrices of the best performing LSTM and CNN-LSTM models before combining with clinical in 1 fold (54 samples).

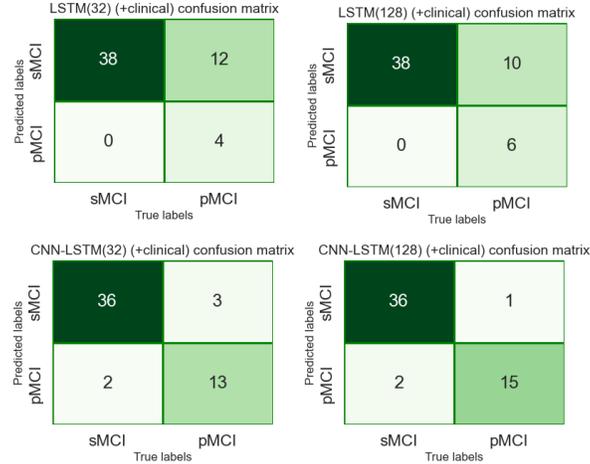


Fig. 8. Confusion matrices of the best performing LSTM and CNN-LSTM models after combining with clinical in 1 fold (54 samples).

IV. discussion

In this study, we evaluated and compared the performance of deep learning techniques in diagnosing the conversion of MCI to Alzheimer's over an average of five years. We have used rs-fMRI, neurophysiological test results, and common demographic characteristics as the noninvasive and clinically available predictors. Our primary focus was to extract appropriate and sufficient features from the imaging data and then improve the results by combining clinical information. Since fMRI data is a series of three-dimensional images recorded over a limited time, it would make sense to use 3D CNN to extract its spatial features, following an LSTM layer to learn its temporal information, which would result in a CNN-LSTM network. Also, to ascertain that it was required to study both spatial and temporal aspects of fMRI to reach good results, LSTM models were separately trained for the comparison. According to the results (Table III), considering both aspects of the information contained in fMRI undoubtedly enhanced the performance. We also observed that the number of cells in the LSTM layers affected the performance of the models. According to AUC and accuracy metrics, we saw an improvement by increasing the number of cells from 32 to 128. Finally, after combination with clinical information, the CNN-LSTM model with 128-cell outperformed other models based on an AUC of 96.67%, an accuracy of 92.47%, a sensitivity of 92.33%, and a specificity of 92.55%.

Among the results acquired from all models (Table III), the LSTM networks performed poorly compared to the CNN-LSTM models, meaning that the LSTM did not extract suitable features from fMRI, which could be due to several factors. First, the structure of LSTM might not have been adequate. We tried to have consistency in our work, so by deleting the CNN parts from the proposed CNN-LSTM models; we would have our LSTM models. As a result, probably, LSTM models with more layers, different numbers of units, and different structural architecture, in general, could bring about better performance. Second, it could be because of the atlas that we used; other atlases divide the brain into different regions and deliver us more and different time series. One last possible factor is the type of fMRI used here, which was resting-state and not task-related. Given that LSTM is mainly concentrated on temporal details rather than spatial, it could be pointed out that resting-state form contains less valuable temporal data than task-related; thus, apparently, not much information was provided here. Therefore, we deduced that the spatial aspect of the fMRI data played a pivotal role in the diagnostic abilities of the models.

We have designed our models based on the characteristics extracted from a combination of imaging data and clinically available noninvasive information. In all of these final models, the output of the imaging section (32 in length) was connected to the output of a 5-neuron dense layer (clinical), delivering a vector containing 37 elements to the classification layer. All four models performed better than their basic form when combined with clinical information. It is because, in general, the models had more information to make decisions; therefore, the final results of the models improved. It was also observed that integrating clinical data with imaging increased the sensitivity of the models to class one (Table 3, Fig. 6, and Fig. 7). Finally, we observed that the proposed method for the combination affected two critical issues: first, the models reached the appropriate results faster (based on the performance of the models on the evaluation data) because, as mentioned, the model had more information to identify. Second, it could have caused an overfitting error if the number of neurons in the clinical layer had increased or the training process had lasted longer.

Many studies have tried to work out the problem of predicting AD among MCI subjects. Even though some studies provided machine learning techniques with a good level of accuracy, the data used by their algorithms was a combination of invasive data, which makes the proposed methods less applicable in real life. Table IV summarizes the articles that predicted the conversion of MCI subjects to AD. As mentioned in this table and indicated in a review of studies that used neuroimaging biomarkers for AD prediction [48], MRI, fMRI, and PET modalities are more frequently used. Regarding the algorithms, CNN from deep learning methods and SVM from machine learning methods were the most common, with CNN-based models achieving higher average performance than other techniques. All articles similar to the present article used cross-validation (CV) for training and evaluation, and the values of accuracy, AUC, sensitivity, and specificity are reported in the table. Also, in reviewing the articles, one should pay attention to the number of folds in the CV method, prediction interval (in years), and the number of pMCI samples. For example, in articles [18] and [19], data from 80 people were used, 18 of them belonged to the pMCI class, and in the CV method, the data were divided into nine parts. Finally, a two-stage AD progression detection framework is proposed. In first stage, a multiclass classification task is used to diagnose a patient in three classes (cognitively normal, MCI, or AD) by LSTM with accuracy of 93.87%. In the second stage, a regression task that predicts the exact conversion time of MCI patients is used. In the regression stage, the LSTM model achieved the best results with mean absolute error of 0.1375.

Our proposed method has some advantages and limitations in comparison to the existing research works. Clinical cognitive tests used in the present study are the most attainable and noninvasive data types that are regularly registered for a basic diagnosis before any further assessments. However, the information provided from different exams has overlaps, and patients could get weary of taking 12 tests altogether. As for the fMRI technique, even though it has advantages over other imaging methods, it is time-consuming and motion-sensitive, which could be inconvenient. Another essential point is that this study's Alzheimer's prediction time interval was an average of five years. While in other studies, it was between three and four years, and we know that the prediction operation is more difficult with increasing prediction time. Also, due to the data imbalance and the fact that most samples belonged to the sMCI class, the models tended to this class by default. As a result, simple models were not responsive and reliable outputs were not delivered, making designing a suitable model challenging. Finally, as far as we can tell, DL methods in processing functional images and combining them with clinical data to predict AD have not been studied. The advantage of using deep neural networks in image analysis such as fMRI

is that the image structure will remain intact. Compared to conventional machine learning methods, deep neural networks are capable of automatic raw data analysis, feature extraction, and classification. Thus, no information is lost at first, and then depending on the architecture of the designed model, the relevant features are provided. In the current research, we tried to design a DL model that converted a volume into 128 single attributes without losing the least amount of information. Therefore, one of the main reasons CNN-LSTM outperforms other machine learning techniques in processing fMRI is that we do not need to manipulate fMRI data to feed it to the model. Thus, the data structure is preserved, while it is not the case for machine learning methods. It should be borne in mind that deep neural networks are complex and computationally expensive and require robust hardware to be trained. Overall, DL methods require more potent hardware, time, experience, and careful precision to be designed and trained without over-fitting or under-fitting. Also, adjusting the rising number of hyper-parameters in deep learning models is challenging since they are affected by one another very closely. Nevertheless, we wanted to explore their utility and performance to our advantage due to their novelty and potential.

Accepted Manuscript (Uncorrected Proof)

TABLE IV : Summary of related works that predicted the conversion of MCI subjects to AD

Study / Year	Modality	Subjects (pMCI)	PI	Method (validation)	Results (%)			
					AUC	Sen	Spe	Acc
[49]/2022	fMRI +clinical	81 (28)	5	3D-CNN (5-CV)	91.72	75.58	92.57	87.59
[50]/2022	Neuroimaging+ clinical	sMCI: 473 pMCI: 140	2.5	LSTM	MAE: 0.137 5	NR	NR	NR
[29]/2021	MRI	294 (125)	2.5	Auto-encoder+ CNN/ SVM (10-CV)	NR	92.4	80.4	87.2
[30]/2020	Multivariate	251 (110)	3	SVM (10-CV, 100 repeat)	94.7	NR	NR	87.1
[32]/2020	MRI	297 (168)	3	3D-CNN (5-CV)	81	77	76	76
[28]/2020	MRI	474 (240)	4	Average 2D-CNN/ transfer learning (CV)	70.6	NR	NR	NR
[34]/2019	clinical	550 (197)	4	Weighted average ensemble /52 ML models (5-CV)	88	77.7	79.9	NR
[20]/2019	fMRI +MRI	134	3	Temporal + convolutional features/ SVM (3-CV, 10 repeat)	78	NR	NR	NR
[33]/2018	clinical	184 (48)	4	SVM (10-CV)	96.2	NR	NR	NR
[19]/2018	fMRI +MRI	80 (18)	4	Graph theory/ SVM (9-CV)	98	94.97	100	97
[18]/2017	fMRI	80 (18)	3	Graph theory/ SVM (9-CV, 1000 repeat)	95	83.83	90.1	91.4
Present	fMRI +clinical	81 (28)	5	3D-CNN-LSTM (5-CV)	96.67	92.33	92.55	92.47

NR: Not Reported; PI: Prediction Interval (in years); CV: Cross Validation; MAE: mean absolute error

C. Future Works

Despite the promising results presented here, several other structures can be designed for the LSTM and the CNN-LSTM models in future works. The changes that can be considered include the number of LSTM layers, the number of cells, and the type of recurrent layer and changing it to gated recurrent units neural network. Furthermore, according to the data from the atlas, other independent structures can be designed for the recurrent model alone. It is also recommended that models be designed and evaluated on more different test and training samples. In addition, it is suggested that transfer learning methods be used in the design of the CNN model and especially the CNN-LSTM model to check the possibility of improving the results with these techniques. Lastly, it is perhaps feasible to use a subset of carefully chosen clinical features to either generate better results or provide more comfort to the patient by reducing the amount of time spent on taking clinical tests.

V. Conclusion

In the present study, a 3D-CNN-LSTM algorithm was developed that was very sensitive to correctly identifying MCI patients who may develop Alzheimer's between six months and an average of 5 years.

We have shown that combining clinical data with fMRI based on the proposed approach leads to better results. We have also proved that processing both spatial and temporal aspects of the rs-fMRI data is crucial for reaching decent outcomes. Finally, our model outperformed other ML methods based on AUC (96.67%), accuracy (92.47%), sensitivity (92.33%), and specificity (92.55%) and can be used for initial clinical diagnosis.

Ethics declarations

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Acknowledgment:

This research is financially supported by " Shahid Beheshti University of Medical Sciences" (Grant No 43014246).

Accepted Manuscript (Unconnected Proof)

References

- [1] J. Weller and A. Budson, "Current understanding of Alzheimer's Disease diagnosis and treatment," (in eng), *F1000Res*, vol. 7, pp. 1-9, Oct. 2018. doi:10.12688/f1000research.14506.1.
- [2] K. G. Yiannopoulou and S. G. Papageorgiou, "Current and future treatments in Alzheimer Disease: an update," *J Cent Nerv Syst Dis*, vol. 12, Feb. 2020. doi:10.1177/1179573520907397.
- [3] S. Engelborghs *et al.*, "Consensus guidelines for lumbar puncture in patients with neurological diseases," *Alzheimers Dement (Amst)*, vol. 8, pp. 111-126, 2017. doi:10.1016/j.dadm.2017.04.007.
- [4] A. Anoop, P. K. Singh, R. S. Jacob, and S. K. Maji, "CSF biomarkers for Alzheimer's Disease diagnosis," *Int J Alzheimers Dis*, vol. 2010, pp. 1-12, Jun. 2010. doi:10.4061/2010/606802.
- [5] Q. Li, S. He, Y. Chen, F. Feng, W. Qu, and H. Sun, "Donepezil-based multi-functional cholinesterase inhibitors for treatment of Alzheimer's Disease," *Eur J Med Chem*, vol. 158, pp. 463-477, Oct. 2018. doi:10.1016/j.ejmech.2018.09.031.
- [6] M. M. Koola, "Galantamine-Memantine combination in the treatment of Alzheimer's Disease and beyond," *Psychiatry Res*, vol. 293, p. 113409, Nov. 2020. doi:10.1016/j.psychres.2020.113409.
- [7] B. Ray *et al.*, "Rivastigmine modifies the alpha-secretase pathway and potentially early Alzheimer's Disease," *Transl Psychiatry*, vol. 10, no. 1, p. 47, Feb. 2020. doi:10.1038/s41398-020-0709-x.
- [8] I. S. Padda and M. Parmar, "Aducanumab," *StatPearls [Internet]*, 2021.
- [9] J. Neugroschl and S. Wang, "Alzheimer's Disease: diagnosis and treatment across the spectrum of disease severity," (in eng), *Mt Sinai J Med*, vol. 78, no. 4, pp. 596-612, Aug. 2011. doi:10.1002/msj.20279.
- [10] A. M. Sanford, "Mild Cognitive Impairment," *Clin Geriatr Med*, vol. 33, no. 3, pp. 325-337, Aug. 2017. doi:10.1016/j.cger.2017.02.005.
- [11] D. P. Devanand *et al.*, "Combining early markers strongly predicts conversion from Mild Cognitive Impairment to Alzheimer's Disease," (in eng), *Biological psychiatry*, vol. 64, no. 10, pp. 871-879, Nov. 2008. doi:10.1016/j.biopsych.2008.06.020.
- [12] V. Mantzavinos and A. Alexiou, "Biomarkers for Alzheimer's Disease diagnosis," *Curr Alzheimer Res*, vol. 14, no. 11, pp. 1149-1154, 2017. doi:10.2174/1567205014666170203125942.
- [13] K. Buttenschoen, M. Kornmann, D. Berger, G. Leder, H. G. Beger, and C. Vasilescu, "Endotoxemia and endotoxin tolerance in patients with ARDS," *Langenbecks Arch Surg*, vol. 393, no. 4, pp. 473-478, Jul. 2008. doi:10.1007/s00423-008-0317-3.
- [14] M. S. Judenhofer and S. R. Cherry, "Applications for preclinical PET/MRI," in *Seminars in nuclear medicine*, 2013, vol. 43, no. 1, pp. 19-29; Elsevier.
- [15] T. Varghese, R. Sheelakumari, J. S. James, and P. Mathuranath, "A review of neuroimaging biomarkers of Alzheimer's Disease," (in eng), *Neurol Asia*, vol. 18, no. 3, pp. 239-248, 2013. doi:https://doi.org/10.1186/s13024-019-0325-5.
- [16] M. A. DeTure and D. W. Dickson, "The neuropathological diagnosis of Alzheimer's Disease," (in eng), *Mol Neurodegener*, vol. 14, no. 1, p. 32, Aug. 2019. doi:10.1186/s13024-019-0333-5.
- [17] M. Tible *et al.*, "Dissection of synaptic pathways through the CSF biomarkers for predicting Alzheimer Disease," *Neurology*, vol. 95, no. 8, pp. e953-e961, Aug. 2020. doi:10.1212/WNL.0000000000010131.
- [18] S. H. Hojjati, A. Ebrahimzadeh, A. Khazae, A. Babajani-Feremi, and I. Alzheimer's Disease Neuroimaging, "Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM," (in eng), *J Neurosci Methods*, vol. 282, pp. 69-80, Apr. 2017. doi:10.1016/j.jneumeth.2017.03.006.
- [19] S. H. Hojjati, A. Ebrahimzadeh, A. Khazae, A. Babajani-Feremi, and I. Alzheimer's Disease Neuroimaging, "Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI," (in eng), *Computers in biology and medicine*, vol. 102, pp. 30-39, Nov. 2018. doi:10.1016/j.compbiomed.2018.09.004.
- [20] A. Abrol, Z. Fu, Y. Du, and V. D. Calhoun, "Multimodal data fusion of Deep Learning and dynamic functional connectivity features to predict Alzheimer's Disease progression," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 4409-4413; IEEE.
- [21] Y. Cui *et al.*, "Identification of conversion from Mild Cognitive Impairment to Alzheimer's Disease using multivariate predictors," (in eng), *PloS one*, vol. 6, no. 7, p. e21896, 2011. doi:10.1371/journal.pone.0021896.

- [22] D. Zhang, D. Shen, and I. Alzheimer's Disease Neuroimaging, "Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers," (in eng), *PLoS one*, vol. 7, no. 3, p. e33182, 2012. doi:10.1371/journal.pone.0033182.
- [23] G. Gavidia-Bovadilla, S. Kanaan-Izquierdo, M. Mataro-Serrat, A. Perera-Lluna, and I. Alzheimer's Disease Neuroimaging, "Early prediction of Alzheimer's Disease using null longitudinal model-based classifiers," (in eng), *PLoS one*, vol. 12, no. 1, p. e0168011, 2017. doi:10.1371/journal.pone.0168011.
- [24] A. Zhao *et al.*, "Increased prediction value of biomarker combinations for the conversion of Mild Cognitive Impairment to Alzheimer's dementia," *Transl Neurodegener*, vol. 9, no. 1, p. 30, Aug. 2020. doi:10.1186/s40035-020-00210-5.
- [25] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, J. Tohka, and I. Alzheimer's Disease Neuroimaging, "Machine Learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *Neuroimage*, vol. 104, pp. 398-412, Jan. 2015. doi:10.1016/j.neuroimage.2014.10.002.
- [26] C. Cabral, P. M. Morgado, D. Campos Costa, M. Silveira, and I. Alzheimer's Disease Neuroimaging, "Predicting conversion from MCI to AD with FDG-PET brain images at different prodromal stages," (in eng), *Computers in biology and medicine*, vol. 58, pp. 101-109, Mar. 2015. doi:10.1016/j.compbiomed.2015.01.003.
- [27] H. Li *et al.*, "A Deep Learning model for early prediction of Alzheimer's Disease dementia based on hippocampal Magnetic Resonance Imaging data," (in eng), *Alzheimer's & dementia : the journal of the Alzheimer's Association*, vol. 15, no. 8, pp. 1059-1070, Aug. 2019. doi:10.1016/j.jalz.2019.02.007.
- [28] L. Nanni *et al.*, "Comparison of Transfer Learning and conventional Machine Learning applied to structural brain MRI for the early diagnosis and prognosis of Alzheimer's Disease," (in English), *Front Neurol*, Original Research vol. 11, no. 1345, p. 576194, Nov. 2020. doi:10.3389/fneur.2020.576194.
- [29] F. Er and D. Goularas, "Predicting the prognosis of MCI patients using longitudinal MRI data," *IEEE/ACM Trans Comput Biol Bioinform*, vol. 18, no. 3, pp. 1164-1173, May. 2021. doi:10.1109/TCBB.2020.3017872.
- [30] W. Lin *et al.*, "Predicting Alzheimer's Disease conversion from Mild Cognitive Impairment using an extreme learning machine-based grading method with multimodal data," (in eng), *Frontiers in aging neuroscience*, vol. 12, p. 77, May. 2020. doi:10.3389/fnagi.2020.00077.
- [31] X. Pan *et al.*, "Multi-View Separable Pyramid Network for AD prediction at MCI stage by (18)F-FDG brain PET imaging," *IEEE Trans Med Imaging*, vol. 40, no. 1, pp. 81-92, Jan. 2021. doi:10.1109/TMI.2020.3022591.
- [32] F. Gao *et al.*, "AD-NET: age-adjust neural network for improved MCI to AD conversion prediction," *Neuroimage Clin*, vol. 27, p. 102290, Jan. 2020. doi:10.1016/j.nicl.2020.102290.
- [33] M. Grassi, G. Perna, D. Caldirola, K. Schruers, R. Duara, and D. A. Loewenstein, "A clinically-translatable Machine Learning algorithm for the prediction of Alzheimer's Disease conversion in individuals with mild and premild cognitive impairment," (in eng), *Journal of Alzheimer's disease : JAD*, vol. 61, no. 4, pp. 1555-1573, 2018. doi:10.3233/JAD-170547.
- [34] M. Grassi *et al.*, "A novel ensemble-Based Machine Learning algorithm to predict the conversion from Mild Cognitive Impairment to Alzheimer's Disease using socio-demographic characteristics, clinical information, and neuropsychological measures," *Front Neurol*, vol. 10, p. 756, Jul. 2019. doi:10.3389/fneur.2019.00756.
- [35] S. S. Poil, W. de Haan, W. M. van der Flier, H. D. Mansvelder, P. Scheltens, and K. Linkenkaer-Hansen, "Integrative EEG biomarkers predict progression to Alzheimer's Disease at the MCI stage," (in eng), *Frontiers in aging neuroscience*, vol. 5, p. 58, 2013. doi:10.3389/fnagi.2013.00058.
- [36] M. Xu, D. L. Sanz, P. Garces, F. Maestu, Q. Li, and D. Pantazis, "A Graph Gaussian Embedding method for predicting Alzheimer's Disease progression with MEG brain networks," *IEEE Trans Biomed Eng*, vol. 68, no. 5, pp. 1579-1588, May. 2021. doi:10.1109/TBME.2021.3049199.
- [37] C. A. Lynch *et al.*, "The clinical dementia rating sum of box score in mild dementia," *Dement Geriatr Cogn Disord*, vol. 21, no. 1, pp. 40-43, 2006. doi:10.1159/000089218.
- [38] C. Chlebowski, "Wechsler Memory Scale All Versions," in *Encyclopedia of Clinical Neuropsychology*, J. S. Kreutzer, J. DeLuca, and B. Caplan, Eds. New York, NY: Springer New York, 2011, pp. 2688-2690.
- [39] T. A. Salthouse, "What cognitive abilities are involved in trail-making performance?," (in eng), *Intelligence*, vol. 39, no. 4, pp. 222-232, Jul. 2011. doi:10.1016/j.intell.2011.03.001.
- [40] I. Arevalo-Rodriguez *et al.*, "Mini-Mental State Examination (MMSE) for the detection of Alzheimer's Disease and other dementias in people with Mild Cognitive Impairment (MCI)," (in eng), *The Cochrane*

- database of systematic reviews*, vol. 2015, no. 3, p. CD010783, Mar. 2015. doi:10.1002/14651858.CD010783.pub2.
- [41] G. A. Marshall *et al.*, "Functional Activities Questionnaire Items that best discriminate and predict progression from clinically normal to Mild Cognitive Impairment," (in eng), *Curr Alzheimer Res*, vol. 12, no. 5, pp. 493-502, 2015. doi:10.2174/156720501205150526115003.
- [42] E. Khosravi Fard, L. K. J, A. Akbarzadeh Bagheban, and W. K. R, "Comparison of the Rey Auditory Verbal Learning Test (RAVLT) and Digit Test among typically achieving and gifted students," (in eng), *Iranian journal of child neurology*, vol. 10, no. 2, pp. 26-37, Jul. 2016.
- [43] J. K. Kueper, M. Speechley, and M. Montero-Odasso, "The Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog): modifications and responsiveness in pre-Dementia populations. a narrative review," (in eng), *Journal of Alzheimer's disease : JAD*, vol. 63, no. 2, pp. 423-444, Jun. 2018. doi:10.3233/JAD-170991.
- [44] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional Neural Networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611-629, Aug. 2018. doi:10.1007/s13244-018-0639-9.
- [45] X. Zhang, L. Yao, X. Wang, J. Monaghan, and D. McAlpine, *A survey on Deep Learning based Brain Computer Interface: recent advances and new frontiers*. 2019.
- [46] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929-5955, Dec. 2020. doi:10.1007/s10462-020-09838-1.
- [47] C. Schaffer, "Selecting a classification method by cross-validation," *Machine Learning*, vol. 13, no. 1, pp. 135-143, 1993.
- [48] G Valizadeh, R Elahi, Z Hasankhani, HS Rad, A Shalhaf. Deep Learning Approaches for Early Prediction of Conversion from MCI to AD using MRI and Clinical Data: A Systematic Review. Archives of Computational Methods in Engineering, 1-70. 2024
- [49] S. Ghafoori and A. Shalhaf, "Predicting conversion from MCI to AD by integration of rs-fMRI and clinical information using 3D-Convolutional Neural Network," *Int J Comput Assist Radiol Surg*, Apr. 2022. doi:10.1007/s11548-022-02620-4.
- [50] El-Sappagh, S., Saleh, H., Ali, F. et al. Two-stage deep learning model for Alzheimer's disease detection and prediction of the mild cognitive impairment time. Neural Comput & Applic (2022). <https://doi.org/10.1007/s00521-022-07263-9>