

RESEARCH

Open Access



# A systematic review on machine learning and deep learning techniques in the effective diagnosis of Alzheimer's disease

Akhilesh Deep Arya<sup>1</sup>, Sourabh Singh Verma<sup>1</sup>, Prasun Chakarabarti<sup>2</sup>, Tulika Chakarabarti<sup>3</sup>, Ahmed A. Elngar<sup>4\*</sup>, Ali-Mohammad Kamali<sup>5</sup> and Mohammad Nami<sup>6</sup>

## Abstract

Alzheimer's disease (AD) is a brain-related disease in which the condition of the patient gets worse with time. AD is not a curable disease by any medication. It is impossible to halt the death of brain cells, but with the help of medication, the effects of AD can be delayed. As not all MCI patients will suffer from AD, it is required to accurately diagnose whether a mild cognitive impaired (MCI) patient will convert to AD (namely MCI converter MCI-C) or not (namely MCI non-converter MCI-NC), during early diagnosis. There are two modalities, positron emission tomography (PET) and magnetic resonance image (MRI), used by a physician for the diagnosis of Alzheimer's disease. Machine learning and deep learning perform exceptionally well in the field of computer vision where there is a requirement to extract information from high-dimensional data. Researchers use deep learning models in the field of medicine for diagnosis, prognosis, and even to predict the future health of the patient under medication. This study is a systematic review of publications using machine learning and deep learning methods for early classification of normal cognitive (NC) and Alzheimer's disease (AD). This study is an effort to provide the details of the two most commonly used modalities PET and MRI for the identification of AD, and to evaluate the performance of both modalities while working with different classifiers.

**Keywords** Alzheimer's disease, Dementia, MCI, Neurodegenerative, Machine learning, Deep learning

## 1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease in elderly people which causes impairment of cognitive and memory function. A person suffering from

this disease finds it difficult to perform routine tasks. Not able to remember maps and sometimes the name of a person. For the past few years, many research groups are trying and making efforts for early detection of AD using different machine learning models.

There are various risk factors for AD some of them are age, genetics, education, and coexisting health problems [1]. Most cases of this disease are observed in people aging 65 and above. While the percentage of AD patients between the ages of 65 and 74 is 5%, the risk increases by 50% above the age of 85. It is observed that people with higher education are at less risk. Higher education leads to the formation of more synaptic connections in the brain. This creates a synaptic reserve in the brain, enabling patients to compensate for the loss of neurons as the disease progresses [1, 2]. One of the major causes

\*Correspondence:

Ahmed A. Elngar

eIngar\_7@yahoo.co.uk

<sup>1</sup> SCIT, Manipal University Jaipur, Jaipur 303007, India

<sup>2</sup> ITM (SLS), Baroda University, Vadodara, Gujarat 391510, India

<sup>3</sup> Sir Padampat Singhania University, Udaipur, Rajasthan 313601, India

<sup>4</sup> Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef 62511, Egypt

<sup>5</sup> Department of Neuroscience, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>6</sup> Cognitive Neuropsychology Unit, Department of Social Sciences, Canadian University Dubai, Dubai, UAE

of AD is coexisting health problems. It is observed that people suffering from cardiovascular disease, high blood pressure, or type-2 diabetes increase the risk of Alzheimer’s disease [2].

There is no medication available to cure and prevent Alzheimer’s disease. The only way to lower the risk of Alzheimer’s disease is by decreasing the risk of heart disease. It is observed that people at a high risk of heart disease are also at a high risk of Alzheimer’s disease. Major factors include excess weight, high blood pressure, diabetes, and high cholesterol. As no proper treatment is available to cure and prevent Alzheimer’s disease the only way is to maintain the mental function of the patient and with some medication delays the symptoms of Alzheimer’s disease. Early prediction of Alzheimer’s disease is really helpful to a doctor to maintain the mental health of the patient and delays the effect which helps the patient to live a better life [2].

There are five stages associated with AD as shown in Table 1. Figure 1 represents the MRI and PET scans of different stages of AD. In the effective prediction of Mild Cognitive Impairment 25 (MCI) two main factors are to be considered, the first is the qualitative change (i.e., progression of AD) and the second one is a quantitative change which includes the cognitive scores. Qualitative changes can be measured using classification between the MRI images of MCI-C (Mild Cognitive Impairment convertor) and MCI- NC (Mild Cognitive Impairment non-convertor). To calculate quantitative changes, clinical

scores, e.g., Mini-Mental State Examination (MMSE) and Alzheimer’s disease Assessment’s scale- Cognitive Subscale (ADAS- cog.) are measured at different points of time. Disease progression is measured based on the clinical scores at the previous point in time [3].

High diagnostic accuracy can be achieved by the estimation of functional connectivity of the brain using functional Magnetic Resonance Image (fMRI). As fMRIs are high dimensional in nature, this is always a challenge to estimate functional connectivity, but this can be done by using discriminative group sparse representation [4].

Studies have shown that computer vision and deep learning models have shown quite good results in diagnosing viral [5] or non-viral diseases using medical imaging. In AD, an image processing framework can be used to track changes in white matter at different points of time using diffusion tensor imaging (DTI). This helps to find new insights into the progression of AD. A pipeline is created 40 by 2 scans of a subject at two different times and using test–reset methods and bootstrap method high precision in prognosis is achieved [6]

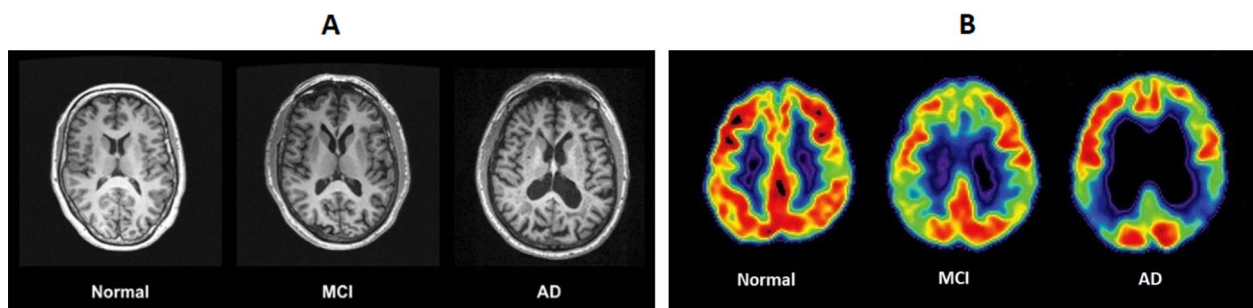
Multiple methods have been attempted to predict prevalence at the outset of AD. The graph-based multiple instances learning method used pre-processed images where patches were used as features for classification. This mi-Graph-based method provides results that are consistently better than the traditional approach to image processing [7].

Taeho Jo. et al. 2019 [8] presented a comparative study of traditional machine learning methods and deep learning methods in the early detection of AD and progression of MCI to AD. They considered 16 studies out of which 4 were using deep learning approaches and traditional machine learning together and 12 studies were using only deep learning approaches. With the use of a deep learning 50 approach with the traditional machine learning approach, the efficiency of 96.0% for feature selection and 84.2% for MCI to AD conversion is achieved. Using convolutional neural networks (CNN) in the deep

**Table 1** Different stages of MCI/AD

There are five stages associated with Alzheimer’s disease:

1	Preclinical Alzheimer’s disease
2	Mild cognitive impairment (MCI)
3	Mild dementia
4	Moderate dementia
5	Severe dementia due to Alzheimer’s disease



**Fig. 1** A MRI scans images of different stages of AD. B PET scan images of different stages of AD

learning approach the accuracy of 96.0% is achieved in feature selection and an accuracy of 84.2% is achieved in MCI to AD conversion prediction. In addition to that, it is observed that performance in the classification can be improved by using multimodal neuroimaging with fluid biomarkers.

Garam Lee et al. [9] use multimodal recurrent neural networks (RNN) to identify the progression of MCI to AD. Cross-sectional neuroimaging, longitudinal CSF analysis, and cognitive performance were used in the proposed framework to improve prediction accuracy. The study found that using a single modality, the accuracy achieved was 75%, whereas using multiple modalities increased the accuracy to 81%.

Zhe Xiao et al. 2017 proposed a framework that extracts multiple features from MRI to train a multi-featured model. This model can effectively identify the subjects suffering from AD. It was noticed that the multi-featured model performs better in comparison to the single feature model [10].

Andres Ortiz et al. use a deep learning approach to classify AD, MCI, and non-converting (NC) subjects. They used automated anatomical labeling (AAL) [11] software package to split-brain the area into 3D patches; these patches were used to train deep neural networks. For the prediction part, four voting algorithms were used and the results were compared. This classification model provided an accuracy value of 0.90 in NC/ AD classification [12].

There are certain limitations of the current machine learning techniques in the prediction of AD. One of them is missing data exclusion. Lei Huang et al. [13] proposed a soft split technique to find the missing scores of the subject and use them at all the previous times to predict the scores the next time. The model performs well and the results are better than the previous studies.

Subramoniam et al. [14] proposed a model where they slice the MRI images and use them as input to the residual convolutional neural networks (ResNet-101) for feature extraction and classification. This model classifies and labels images into four classes moderately demented, mild demented, very mild demented, and non-demented. Using 3 layers of CNN followed with the 3 layers of Vanilla-dense neural network (DNN) this model achieves an accuracy of 95.32%. In DNN first 2 dense layers are used with rectified linear unit (ReLU) activation and 1 unit of dense layer with softmax activation.

Edward Challis et al. [15] compared SVM classifiers and Gaussian process logistic regression (GP-LR) and found that the GP-LR method performs well in comparison to SVM. SVM provides a binary classifier, whereas GP-LR provides the probability of class membership. During the study a total of 77 subjects were taken out of which 27

were AD patients, 50 were suffering from MCI, and 39 were controlled subjects. Accuracy of 75% is achieved in the control of MCI conversion whereas 97% in the case of MCI to AD is achieved.

Alzheimer's disease (AD) is a progressive neurodegenerative disease that affects the cognitive and memory function of elderly people and its early detection is crucial for delaying its effects on the patient's mental health. Several risk factors, including age, genetics, education, and coexisting health problems, increase the risk of AD. There is no medication available to cure AD, but maintaining the mental function of the patient and delaying the symptoms is the only way to lower the risk. Computer vision and deep learning models have shown quite good results in diagnosing AD using medical imaging, multiple methods explained above have been attempted to predict prevalence at the outset of AD. The most favored method among researchers for predicting Alzheimer's disease onset using medical imaging is using fMRI/PET scan data with an SVM classifier and CNN [16]. This study presents a statistical analysis of the different machine learning and deep learning models proposed by researchers, along with the various modalities employed.

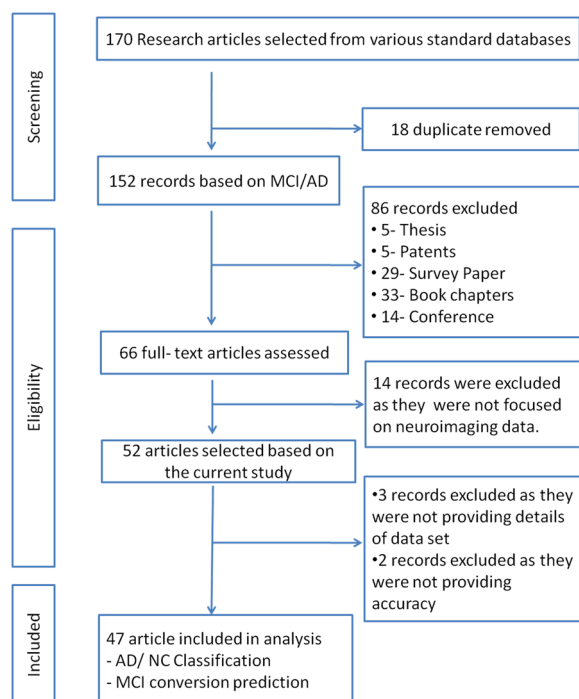
The paper is structured as follows: Sect. 2 outlines the study's material and methods, including the selection process of relevant literature using the Preferred Reporting Item for Systematic Review and Meta-Analysis (PRISMA) flowchart [16], the different datasets used, data pre-processing techniques, various machine learning and deep learning techniques used, and performance evaluation metrics considered for the evaluation. Section 3 presents the experimental details and comparison results. Section 4 discusses the limitations of the study, while Sect. 5 concludes with a future outlook.

## 2 Material and methods used

In this section, selection procedure of the literature for review, various methods used in different studies, and the dataset used by the researchers are discussed.

### 2.1 Literature selection

A systematic review process is performed using the previous publications on the diagnosis and prognosis of Alzheimer's disease and MCI conversion prediction using machine learning and deep learning approaches. During the survey process, the research articles using MRI/PET scans and clinical dataset as modalities are considered in the review. A total of 47 papers were selected and included in the review; out of which 31 papers are used for the comparative analysis and 16 were used for the literature background. The selection procedure is explained using the preferred reporting item for systematic review and meta-analysis (PRISMA) flowchart in Fig. 2.



**Fig. 2** Preferred Reporting Item for Systematic Review and Meta-Analysis (PRISMA) flowchart. The figure explains the selection procedure of literature used for the study

## 2.2 Dataset

There are 3 major sources of the dataset considered in this study:

- ADNI: Alzheimer's disease neuroimaging initiative started in the year 2004 to provide researchers with neuroimages for the effective diagnosis and prognosis of Alzheimer's disease under the leadership of Dr. Michael W. Weiner. 50% of the research articles included in the study are using 1.5 T, 3 T MRI, fMRI images and FDG-PET images as a dataset [9].
- OASIS: Open access series of imaging studies is started with the aim to provide a free neuroimaging dataset to the research community. It provides 3 kinds of data sets:

*a. OASIS-1: contains the study of 416 subjects with the 434 cross-sectional MRI data set of demented and non-demented older adults.*

*b. OASIS-2: contains the study of 150 subjects with the 373 longitudinal MRI datasets of demented and non-demented older adults.*

*c. OASIS-3: contains the study of 1098 subjects with 2168 MRI and 1608 PET images.*

*35% of the research articles included in the study are using the OASIS dataset.*

- Kaggle: offers data set for various research-oriented studies 15% of the research articles use the dataset from Kaggle [13].

Table 3 contains the details such as source, modality, and other numbers of demented and non-demented subjects in the dataset used by different research articles.

## 2.3 Pre-processing of data

To prepare a clean data set various data pre-processing strategies are used in the studies.

Research publications used in this review paper are using two types of data set: one is the clinical data set from OASIS and the second one is image dataset from ADNI and kaggle platform as explained in Sect. 2.2.

There are various data pre-processing methods that can be applied to clinical data sets of Alzheimer's disease to improve the quality and accuracy of the analysis results. Some of the most commonly used pre-processing methods used by the researchers are:

- **Data cleaning:** This involves removing or correcting any incomplete, inconsistent, or erroneous data from the dataset. This is typically done by identifying missing values, outliers, and inconsistencies in the data, and then removing or replacing them [8].
- **Feature selection:** This involves selecting the most relevant features from the dataset that are likely to have a significant impact on the analysis results. This can be done using various statistical techniques such as correlation analysis, principal component analysis, and mutual information [14].
- **Data transformation:** This involves transforming the data into a more suitable format for analysis. This can be done using techniques such as normalization, standardization, and logarithmic transformation [6].
- **Imputation:** This involves filling in missing data with estimated values. This can be done using various imputation techniques such as mean imputation, regression imputation, and multiple imputations.

Image dataset for Alzheimer's consist of MRI ad PET images. Pre-processing methods used in the research articles that were used in this review article are:

- **Image normalization:** This is the process of adjusting the intensity values of images to a common scale. This can be done using techniques such as histogram equalization, contrast stretching, and normalization to improve the visual quality of the images [12].

- **Image cropping and resizing:** This involves resizing images to a uniform size and cropping out any unnecessary parts of the image that do not contain relevant information. This can help reduce the computational complexity of the analysis and improve the efficiency of the algorithms.
- **Image augmentation:** This involves generating additional images from the original dataset by applying various transformations such as flipping, rotating, and scaling. This can help increase the size of the dataset and improve the robustness of the analysis [14].
- **Feature extraction:** This involves extracting relevant features from the images that can be used in the analysis. This can be done using techniques, such as edge detection, texture analysis, and shape analysis [10].
- **Data augmentation:** This involves increasing the size of the dataset by adding variations of the original images, such as changing brightness or contrast, or adding noise [8].

## 2.4 Conventional machine learning models used

### A) Logistic regression

Logistic regression is the type of supervised machine learning algorithm. This algorithm is useful in cases where we have to predict the outcome in binary (true or false). The equation of the logistic regression is:

$$p = 1 / (1 + e^{-(b_0 + b_1x + b_2x^2)}) \quad (1)$$

where 'e' is Euler's number which is 2.71828, b<sub>0</sub>, b<sub>1</sub>, and, b<sub>2</sub> are the coefficient of the model, b<sub>0</sub> is the intercept, while 'b<sub>1</sub>' and 'b<sub>2</sub>' are the slope of the liner and quadratic terms of the predictor variable 'x'. These coefficients are estimated from the training data using method called as maximum likelihood estimation.

### B) Support vector machine

SVM is a fast classification algorithm widely used in the area of image classification. SVM is capable of handling multiclass classification problems. SVM classifier creates a hyperplane that separates the possible outputs:

$$W^T X - b = 0, \quad (2)$$

where  $W^T$  is the transpose of the weight vector, which is a vector of coefficients that determines the orientation of the decision boundary in the feature space. The weight vector 'W' is determined by the optimization algorithm used to train the SVM. 'X' is a feature vector, which represents a data point in the feature space. 'b' is the bias term,

which determines the position of the decision boundary in the feature space [17, 18].

### C) Random forest

Random forest is used to predict the outcome based on the target value set from the dataset. It is based on a supervised machine learning algorithm:

$$f = \frac{1}{B} \sum_{b=1 \rightarrow B} fb(x') \quad (3)$$

where 'b' is the number of decision trees in the forest.  $fb(x')$  is the prediction of the bth decision tree on the input feature vector  $x'$  and  $1/B$  is the scaling factor that averages the prediction of all the decision trees [18].

## 2.5 Deep learning models used

### A) Artificial neural network

A neural network is a set of algorithms that are used to extract information and underlying relationships between the data in the MRI dataset. It works in a similar way a human brain works. There are layers of interconnected neurons in the neural network; these neurons are known as perceptrons. These perceptrons are mathematical functions that classify information from the MRI image dataset according to the requirements and specific architecture.

A neural network generally consists of two layers only; this is not a suitable condition for the computation of large networks such as MRI images. So for deep learning, we introduce more layers to the traditional neural network. It ranges from 10 to 100 layers based on the computation network. Each neuron in the layers stores some information and passes that information to the forward neurons. As this data flows from the network, the hidden information is extracted from the MRI images. Lower layer neurons generally collect the raw data.

$$\sum_{i=0 \rightarrow n} Wi * Xi + b. \quad (4)$$

The above equation represents the output of a single neuron in ANN. Where n indicates the inputs to the neuron,  $Wi$  is weight and  $Xi$  is the value associated with the ith input, and 'b' is the bias term. It computes weighted sum of the inputs, adds a bias term, and then applies the activation function to produce the output of the neuron [19, 20].

### B) Convolutional neural networks (CNN)

In a convolutional network, 2 images represented in the matrix form are multiplied to create a new matrix and from that output, features are extracted. There are two processes that a CNN performs first if feature

extraction and second classification. The major advantage of using CNN is the ability to learn and generalize the features from a large data set.

$$fx, y(S) = \max_{a,b=0} S2x + a, 2y + b. \quad (5)$$

The above equation represents a type of pooling operation. In CNN, the input image is convolved with a set of learnable filters to produce a set of feature maps. After convolution operation, a pooling layer is applied to reduce the spatial dimension of the feature maps and extract the most important features. In the equation 'a' and 'b' represents the horizontal and vertical offsets of the subregion from the top-left corner at position (2x, 2y). The value of 'a' and 'b' can either be '0' or '1' which determines the position of the bottom-right corner of the subregion relative to (2x, 2y).

Research articles included in this review paper are using various CNN models such as, ResNet-101, VGG, VoxCNN, DenseNet a brief description of following CNN models are as follows:

- **ResNet-101:** ResNet-101 is one member of a family of ResNet models, including ResNet-50, ResNet-152, and others, with the number indicating the total number of layers in the network. The input to ResNet-101 is an RGB image with a resolution of 224×224 pixels, and the first layer is a convolutional layer with 7×7 filters, followed by a batch normalization layer and a ReLU activation function. ResNet-101 then has a series of residual blocks, each consisting of two or three convolutional layers with batch normalization and ReLU activation functions. The architecture has a total of 101 layers. The final layers of ResNet-101 include a global average pooling layer that takes the average of each feature map, a fully connected layer, and a softmax activation function that produces the output classification probabilities [18–20].
  - **VGG:** VGG is a 2D convolutional neural network (CNN) architecture that is primarily used for image classification. The input to the VGG model is an RGB image with a resolution of 224×224 pixels. The network consists of a series of convolutional layers, each of which has small 3×3 filters. The number of filters in each layer increases with the depth of the network. Following each convolutional layer, VGG applies the rectified linear unit (ReLU) activation function to introduce nonlinearity into the model. After each set of convolutional layers, VGG uses a max pooling layer with a pool size of 2×2 to reduce the spatial size of the feature maps and increase translation invariance. Finally, the output layer of VGG produces the output classification probabilities [18–20].
  - **VoxCNN:** VoxCNN is a powerful 3D convolutional neural network architecture specifically designed to process 3D volumetric data such as medical images and videos. The network includes convolutional layers, activation functions, max pooling layers, dropout layers, and fully connected layers. Each convolutional layer has a small filter size of 3×3×3, and the number of filters increases with the depth of the network. Max pooling layers with a pool size of 2×2×2 are applied after each convolutional layer to capture the spatial and temporal features of the data. Dropout layers are also used to prevent overfitting during training by randomly dropping out some neurons. The output layer uses the softmax activation function to produce the final classification probabilities. Overall, VoxCNN is a highly effective architecture for various applications, including medical image analysis and video processing.
  - **DenseNet:** The main novelty of the DenseNet architecture is the use of dense connections between layers, which enables efficient information flow through the network. The first layer of the network is a standard convolutional layer with a 7×7 filter, followed by batch normalization and ReLU activation function. The network consists of several dense blocks, each containing multiple dense layers where each layer passes the input through batch normalization, ReLU activation function, and a 3×3 convolutional layer with a fixed number of filters. In a dense block, the output of each dense layer is concatenated with the outputs of all previous layers in the same block, allowing each layer to access the previous layer's outputs for efficient information flow. Max pooling layers are used after each set of layers to reduce the spatial size of the feature maps and introduce translation invariance, while dropout layers prevent overfitting during training. DenseNet also includes transition layers between each pair of dense blocks, consisting of a batch normalization layer, a 1×1 convolutional layer to reduce the number of feature maps, and a 2×2 average pooling layer to down sample the spatial size of the feature maps. Finally, the network ends with a global average pooling layer, a fully connected layer, and a softmax activation function for producing output classification probabilities [20, 21].
- C) Recurrent neural network (RNN)
- RNN is used in the effective prognosis of Alzheimer's disease. RNN is used to predict the future health of the patient suffering from AD; also it predicts whether an MCI patient will convert into AD or not based on the

clinical data of a patient at different time stamps fed into it. Nguyen et al. in august 2020 proposed a minimal-RNN and fed it with patent data taken at three different time stamps and predict the state for the next three time stamps. This data was collected from ADNI [22].

The formula for current state:

$$h_t = f(h_{t-1}, x_t). \tag{6}$$

Based on the previous state  $h_{t-1}$  and current input  $x_t$ , a recurrent neural network (RNN) calculates its current state  $h_t$ . An RNN processes a sequence of inputs  $x_1, x_2, \dots, x_T$  and outputs  $y_1, y_2, \dots, y_T$ . The non-linear transformation  $f'$  maps inputs and previous state to current state.

The formula for activation function:

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}X_t). \tag{7}$$

The activation function in a recurrent neural network (RNN) computes the current state  $h_t$  based on the previous state  $h_{t-1}$  and the current input  $x_t$ . The non-linear hyperbolic tangent activation function (tanh) maps its input to a value between -1 and 1. RNNs use the hyperbolic tangent activation function because of its benefits. The output is always finite because it is bounded. Second, it is differentiable, enabling backpropagation network

training. Third, it is symmetric around the origin, reducing the RNN vanishing gradient problem.

The formula for output:

$$h_t = W_{hy}h_t. \tag{8}$$

In the given equation,  $h_t$  is the hidden state vector of the RNN at time step  $t$ ,  $W_{hy}$  is the weight matrix that maps the hidden state to the output vector  $y$ , and  $y_t$  is the output of the RNN at time stamp  $t$ .

Figure 3 provides an overview of the modalities, feature selection methods, feature extraction techniques, and classification algorithms employed in the reviewed research articles. The predominant modalities utilized by researchers include fMRI images, PET images, and cerebrospinal fluid analysis. In the feature selection process, physicians rely on brain imaging to identify areas of interest or patches depicting brain tissue deterioration associated with AD. Two biomarkers, namely p-tau and t-tau, play a vital role in the assessment of the presence and progression of Alzheimer's disease (AD). These biomarkers offer valuable insights into the pathological changes that occur in the brain associated with the disease.

Subsequently, features are extracted from the different modalities using methods such as normalization, random tree, and truncated singular value decomposition (SVD). These techniques effectively reduce the dimensionality

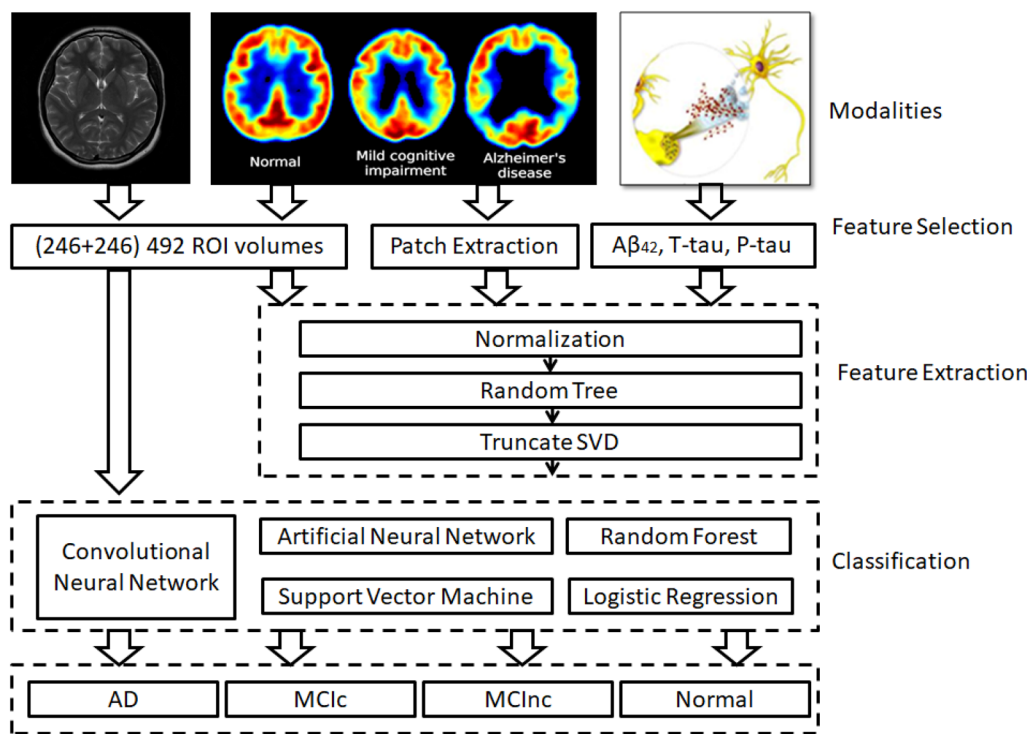


Fig. 3 Description of modalities, feature selection, feature extraction, and classification algorithms used to predict AD and NC cases

of the data while retaining critical information. The extracted features are then employed to train machine learning models for the classification of AD and normal cases. The classification algorithms found in the reviewed articles include convolutional neural networks, artificial neural networks, support vector machines, logistic regression, and random forest.

### 2.6 Performance evaluation measures

To evaluate the performance of the classification model, a confusion matrix is used. Figure 4 shows the confusion matrix. A confusion matrix is an NxN matrix, where N represents the number of classes in the dataset. In our case, as we are doing binary classification the value of N is 2, positive or negative.

Different parameters such as accuracy, specificity, and sensitivity can be calculated for a model using the confusion matrix:

- Accuracy

Classification accuracy shows the correctness of the prediction by a model in the case of AD diagnosis and how many MRIs are correctly labeled out of the total MRI provided as an input to the model [21]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \tag{9}$$

- Sensitivity

Sensitivity is the ratio of correctly positive labeled AD records to all who are suffering from AD in reality [21]:

$$Sensitivity = \frac{TP}{TP + FN}. \tag{10}$$

- Specificity

Specificity is the ratio of correctly labeled negative AD records by the model to all who are cognitively normal in reality [21]:

$$Specificity = \frac{TN}{FP + TN}. \tag{11}$$

- Precision

Precision is the ratio of the correctly positive labeled AD records by the model to the entire positive labeled by the model [15]:

$$Precision = \frac{TP}{TP + FP} * 100. \tag{12}$$

- Area under the curve (AUC)-receiver operating characteristics (ROC) curve

AUC-ROC curve ROC is the probability curve and AUC represents the degree that which good a model separates the cases. In our case, it tells how much the model can distinguish between AD and NC. If the AUC is higher the capability of the model to identify the AD and NC cases will increase [15].

### 3 Experimental

Table 2 represents the various research articles included in the study. It also explains the modality and number of subjects included in the research they conducted. Table 3 explains different methods used by the researchers for the diagnosis of AD and various parameters such as accuracy, sensitivity, specificity, and precision achieved with the different models of machine learning.

Based on the study it is evident that Deep Learning techniques for feature extraction and the traditional machine learning approach of classification using a support vector machine (SVM) classifier are highly efficient in the diagnosis and prediction of Alzheimer’s disease. It has also been observed that multimodality-based diagnosis and prognosis perform better than single modality-based techniques. These can include clinical tests and cognitive tests [41, 42].

A voxel-based feature extraction method for the early diagnosis of AD plays a pivotal role. In this approach, the brain image is divided into 90 regions of interest

		Model Output	
		+	-
Ground Truth	AD	TRUE POSITIVE TP	FALSE NEGATIVE FN
	CN	FALSE POSITIVE FP	TRUE NEGATIVE TN

Fig. 4 Confusion matrix for the prediction of Alzheimer’s disease



**Table 2** Represents the various research articles included in the study. It also explains the modality and number of subjects included in the research they conducted

References	Database	Modality	No. of records				Total
			AD	cMCI/PMCI	ncMCI/SMCI	NC	
Zhang D et al. [3]	ADNI	MRI (T1 weighted), FDG-PET	38	–	–	50	88
Suk H et al. [4]	ADNI	fMRI TR2	12	–	–	25	37
Lu D et al. [6]	ADNI	MRI 1.5 T, FDG-PET	238	217	409	360	1224
Tong T et al. [7]	ADNI	MRI 1.5 T	198	167	238	231	834
Lee G et al. [9]	ADNI	MRI 1.5 T, CFS, Cog. Score	338	307	558	415	1618
Xiao Z et al. [10]	ADNI	MRI 1.5 T	54	–	542	58	654
Ortiz A et al. [12]	ADNI	MRI 1.5 T,FDG-PET	70	26	111	68	275
Subramoniam M et al.(a) [14]	Kaggle	MRI 1.5 T	960	–	2240	3200	6400
Subramoniam M et al.(b) [14]	Kaggle	MRI 1.5 T	960	–	2240	3200	6400
Yue L et al.(a) [23]	ADNI-1	MRI 1.5 T	335	–	542	785	1662
Yue L et al.(b) [23]	ADNI-2	MRI 3 T	366	1304	1583	1106	4359
Goryawala M et al. [24]	ADNI	MRI 1.5 T	55	91	114	125	385
Aderghal [25]	ADNI	MRI 1.5 T	188	–	399	228	815
Korolev (a) [26]	ADNI	MRI 1.5 T	50	43	77	61	231
Korolev (b) [26]	ADNI	MRI 1.5 T	50	43	77	61	231
Cheng and Liu [27]	ADNI	FDG-PET	93	–	146	100	339
Choi H et al. [28]	ADNI	FDG and AV-45 PET	139	79	92	182	492
Cheng and Liu [29]	ADNI	MRI 1.5 T, FDG-PET	93	–	–	100	193
Kumar N et al. [30]	OASIS	MRI 1.5 T	146	–	37	190	373
Ji h et al. [31]	ADNI	MRI 1.5 T	179	–	–	182	361
Battineni G et al. [32]	OASIS	MRI 1.5 T	146	–	37	190	373
Kumari R et al. [33]	OASIS-3	MRI 1.5 T	489	–	–	609	1098
Alickovic E et al. [34]	ADNI	MRI 1.5 T	72	–	–	195	267
Madiwalar S et al. [35]	OASIS	MRI 1.5 T	64	–	–	72	136
Pan D et al. [36]	ADNI	MRI 1.5 T	237	–	–	262	499
Alroobaea R et al. (a) [37]	ADNI	MRI 1.5 T	1731	–	–	2665	4396
Alroobaea R et al. (b) [37]	OASIS	MRI 1.5 T	146	–	37	190	373
Savita et al. [38]	OASIS	MRI 1.5 T	100	–	–	135	235
Li and Yang [39]	ADNI	MRI 1.5 T	260	–	–	300	560
Mehmood A et al. [40]	ADNI	MRI 1.5 T	75	–	–	85	160

(ROI) out of which only informative voxels are selected and stored into a vector for the study corresponding to the baseline. These voxels were provided as input to CNN and variants of CNN such as ensemble system of deep convolutional neural networks and Siamese convolutional neural network [43, 44] for further deep learning. The robustness of the system is then tested against the subset from ADNI. Figure 5 shows that it achieved the accuracy of 97.2% in the case of cognitive normal to AD and 99.4% in the case of MCI-NC to MCI-C. The total subjects included in the study are 1662 which includes 785 cognitively normal, 542 MCI, and 335 AD cases.

The development of a framework to assess the progression of AD using MRI volumes and neuropsychological scores is also a significant approach to the diagnosis and

prognosis of Alzheimer's disease. In this framework using the features selection method subjects were divided into two groups, early MCI (EMCI) and late MCI (LMCI), effectively with an accuracy rate of 73.6%. This results in a significant amount of improvement and the accuracy achieved with this model is 93.9% in the case of NC/AD classification which can be observed in Fig. 5.

Figure 6 illustrates two primary pieces of information. Firstly, it presents the frequency of usage of machine learning and deep learning algorithms by researchers. Secondly, it showcases the highest accuracy achieved by each specific algorithm. It is evident that there is a prevalent trend of employing deep learning methods, which can be attributed to their notable accuracy, particularly in handling complex three-dimensional data. Deep learning models have demonstrated superior performance in the

**Table 3** Explains different methods used by the researchers for the diagnosis of AD and various parameters such as accuracy, sensitivity, specificity, and precision achieved with the different models of machine learning.

References	Method	AD:NC Acc	SEN recall	SPE	Precision	AUC
Zhang D et al. [3]	Multi-kernel SVM	<b>78.40</b>	<b>79.00</b>	<b>78.00</b>	–	<b>76.80</b>
Suk H et al. [4]	Group sparse representation + SVM	<b>89.19</b>	<b>91.00</b>	<b>88.00</b>	–	<b>95.60</b>
Lu D et al. [6]	DNN	<b>82.93</b>	<b>79.69</b>	<b>83.84</b>	–	–
Tong T et al. [7]	Multiple instance learning + SVM	<b>89.00</b>	–	–	–	–
Lee G et al. [9]	Multi- model deep learning + RNN	<b>81.00</b>	<b>84.00</b>	<b>80.00</b>	–	<b>86.00</b>
Xiao Z et al. [10]	SVM-REF with covariance	<b>85.71</b>	<b>79.63</b>	<b>91.38</b>	–	–
Ortiz A et al. [12]	Deep belief network (SVM)	<b>90.00</b>	<b>86.00</b>	<b>94.00</b>	–	<b>95.00</b>
Subramoniam M et al.(a) [14]	Vanilla DNN	<b>95.31</b>	–	–	–	–
Subramoniam M et al.(b) [14]	CNN	<b>95.32</b>	–	–	–	–
Yue L et al.(a) [23]	CNN	<b>99.40</b>	<b>100</b>	–	<b>98.80</b>	<b>97.20</b>
Yue L et al.(b) [23]	CNN	<b>98.60</b>	<b>97.20</b>	–	<b>100</b>	<b>98.60</b>
Goryawala M et al. [24]	Linear regression model + LDA	<b>93.90</b>	<b>96.30</b>	<b>89.50</b>	<b>93.80</b>	–
Aderghal [25]	2D CNN	<b>91.41</b>	<b>93.75</b>	<b>89.60</b>	–	–
Korolev (a) [26]	3D CNN	<b>79.00</b>	–	–	–	<b>88.00</b>
Korolev (b) [26]	ResNet	<b>80.00</b>	–	–	–	<b>87.00</b>
Cheng and Liu [27]	RNN	<b>91.20</b>	<b>91.40</b>	<b>83.84</b>	<b>95.30</b>	–
Choi H et al. [28]	3D CNN	<b>96.00</b>	–	–	–	<b>91.00</b>
Cheng and Liu [29]	3D CNN	<b>89.60</b>	<b>87.10</b>	<b>92.00</b>	–	<b>94.45</b>
Kumar N et al. [30]	Linear discriminant analysis	<b>98.92</b>	–	–	–	–
Ji h et al. [31]	CNN	<b>98.59</b>	<b>97.22</b>	<b>100</b>	–	–
Battineni G et al. [32]	Linear regression model	<b>98.30</b>	<b>97.40</b>	–	<b>98.60</b>	<b>99.70</b>
Kumari R et al. [33]	CNN	<b>90.25</b>	<b>85.53</b>	–	–	–
Alickovic E et al. [34]	Random forest	<b>85.77</b>	<b>54.17</b>	<b>97.44</b>	–	–
Madiwalar S et al. [35]	Extra tree	<b>93.14</b>	<b>85.00</b>	–	<b>85.00</b>	–
Pan D et al. [36]	CNN	<b>84.00</b>	–	–	–	<b>92.00</b>
Alroobaea R et al. (a) [37]	Random forest	<b>98.89</b>	<b>99.19</b>	–	<b>98.89</b>	–
Alroobaea R et al. (b) [37]	Logistic regression	<b>84.33</b>	<b>84.14</b>	–	<b>84.54</b>	–
Savita et al. [38]	SVM	<b>85.80</b>	–	–	<b>87.83</b>	–
Li and Yang [39]	SVM	<b>90.00</b>	<b>93.90</b>	<b>85.10</b>	–	<b>97.00</b>
Mehmood A et al. [40]	CNN	<b>98.73</b>	<b>98.19</b>	<b>99.09</b>	–	–

domain of Alzheimer's disease research, making them a preferred choice for many researchers.

Figure 7 represents the performance achieved by different research articles in terms of sensitivity (true positive rate) and specificity (false positive rate). The true positive rate indicates that the model predicts the positive cases correctly. And the False positive rate indicates that the model predicts negative to the positive class. Figure 8 is the AUROC curve between true positive rate (TPR) and false positive rates (FPR) achieved by CNN, SVM, and other methods used in the research articles included in the study. Figures 8 and 9 represent the stacked area chart of the TPR and FPR.

Figures 9 and 10 are the stacked representation of the sensitivity and specificity derived from different methods used by research articles referred to in this review paper. Application of 3d CNN with PET images achieved an accuracy of 96% for AD/NC classification and accuracy of 84.2% in the case of MCI-C/ MCI-NC which is much better in comparison to results achieved by SVM classifiers. CNN for the classification achieved an accuracy of 91.41% while RNN on PET images achieved an accuracy of 91.2% in NC/ AD classification. If the landmark detection method with CNN on MRI images was applied then it achieved an accuracy of 91.09%.

Accuracy

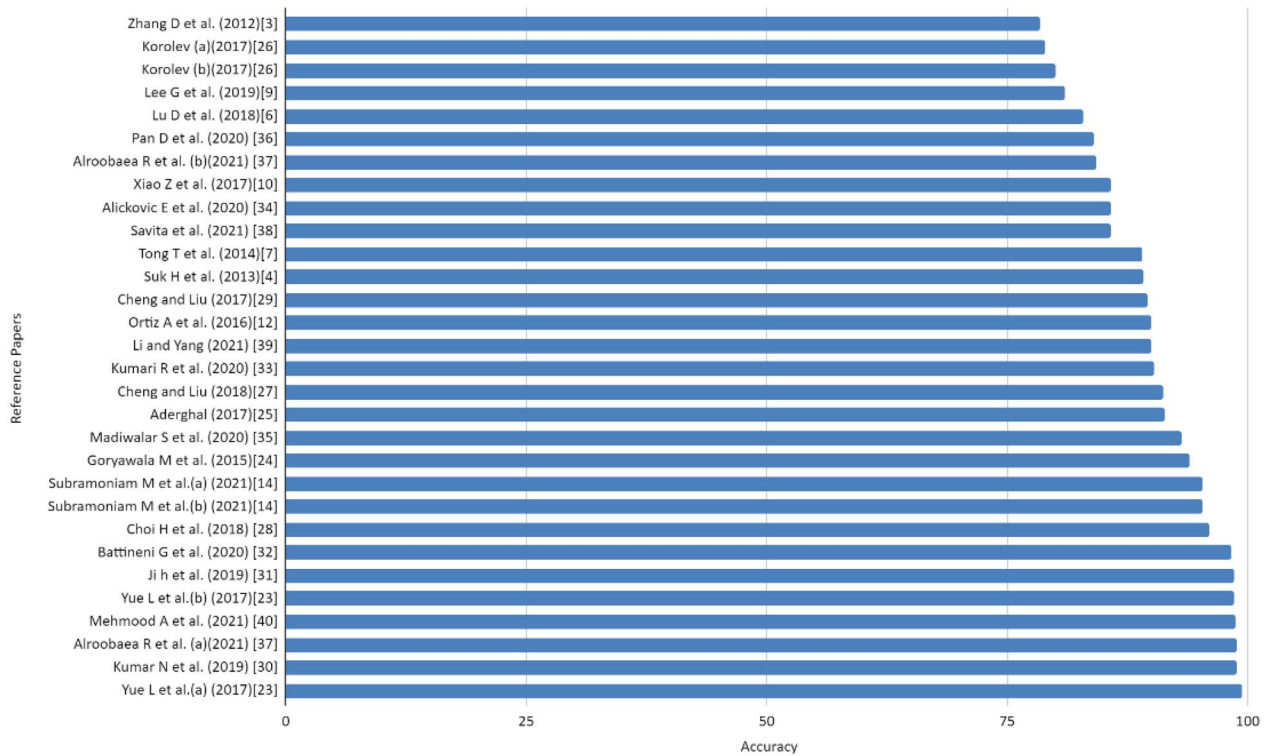


Fig. 5 Comparison chart of accuracy derived from different articles

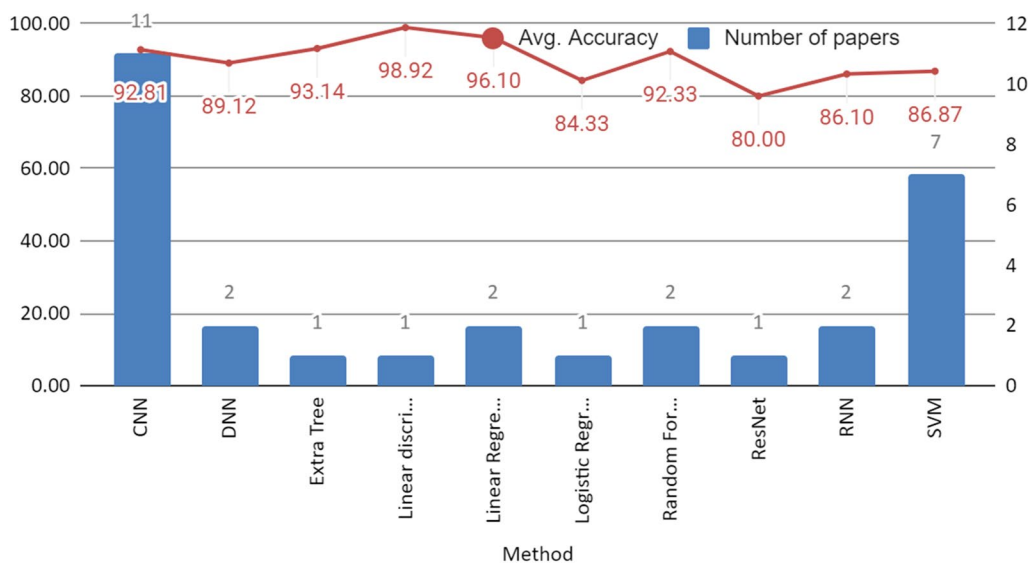
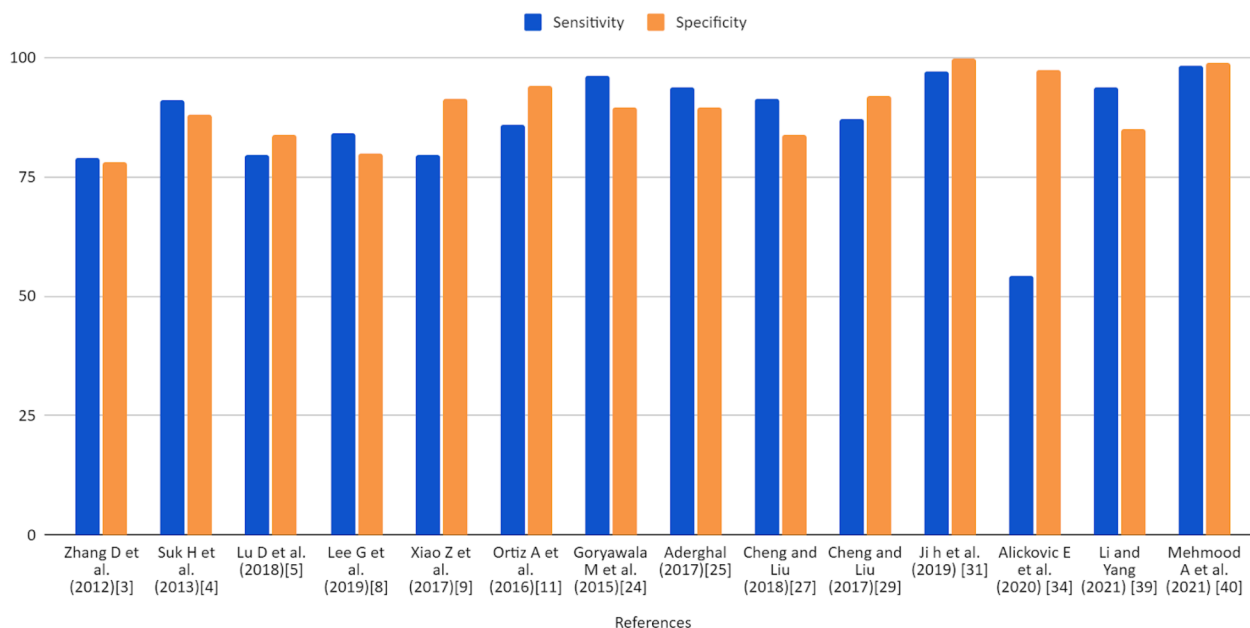


Fig. 6 Frequency of methods used in different articles taken in the study, and average accuracy using different classifiers

4 Limitations

It is crucial to acknowledge the limitations of this study. While there are various modalities available for the diagnosis of Alzheimer’s disease, including medical history,

blood tests, spinal taps, and genetic testing, this study only examines research articles that utilize two specific modalities: MRI/PET scans and cognitive assessments obtained from clinical data sets.



**Fig. 7** Comparison chart of performance in terms of sensitivity and specificity

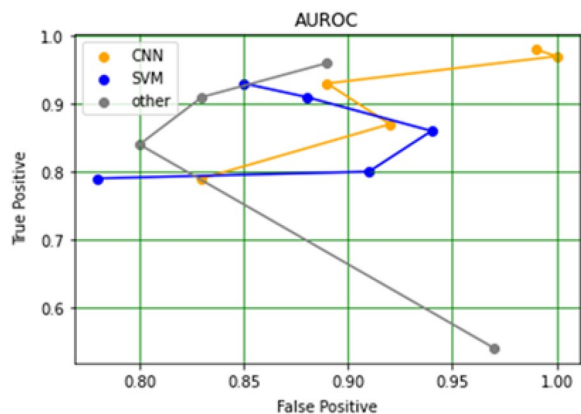
AD patients often exhibit alterations in brain electrical activity, which can be reflected in magnetic field measurements. Magnetometer and gradiometer data, researchers have the opportunity to capture and analyze magnetic field variations, allowing for the identification of potential patterns or signatures that differentiate AD patients from healthy individuals [45].

Additionally, it is important to note that the literature reviewed in this study heavily relies on data from the Alzheimer’s disease Neuroimaging Initiative (ADNI) database, resulting in a substantial overlap in sample populations. This dependency exists due to the high expense of fMRI data acquisition and limited access to other databases.

### 5 Conclusion

Early and effective diagnosis of Alzheimer’s disease is very critical in the treatment of the patient. With the early detection of AD, it is possible to minimize the effects of AD on neuron degeneration. It is evident that machine learning and deep learning techniques contribute significantly toward the early detection of Alzheimer’s disease. In traditional ML techniques using the support vector machine (SVM) classifiers, an accuracy of 85.71% can be achieved in the classification of CN/AD. In deep learning using CNN classifier accuracy of 98.6% and with RNN accuracy of 91.2% can be achieved. Our study suggests that decisions made on the predictions based on the multimodalities are more reliable and accurate in comparison of the single modality. This article will be of great assistance in gaining an understanding of the various research approaches that have been utilized in recent studies on Alzheimer’s disease.

Recent trends show that use of the deep learning techniques in the analysis of medical images has increased for faster analysis and better accuracy than a human practitioner. This study will facilitate better accuracy in the future by identifying various combinations of existing models. Ensemble learning techniques can enhance the predictive accuracy of better-performing models by combining them. This involves training multiple models on the same dataset and aggregating their predictions to generate a final output. Ensemble learning can provide benefits, such as improving accuracy, mitigating



**Fig. 8** AUROC curve between the true positive rate and false positive rate derived from CNN, SVM, and other methods

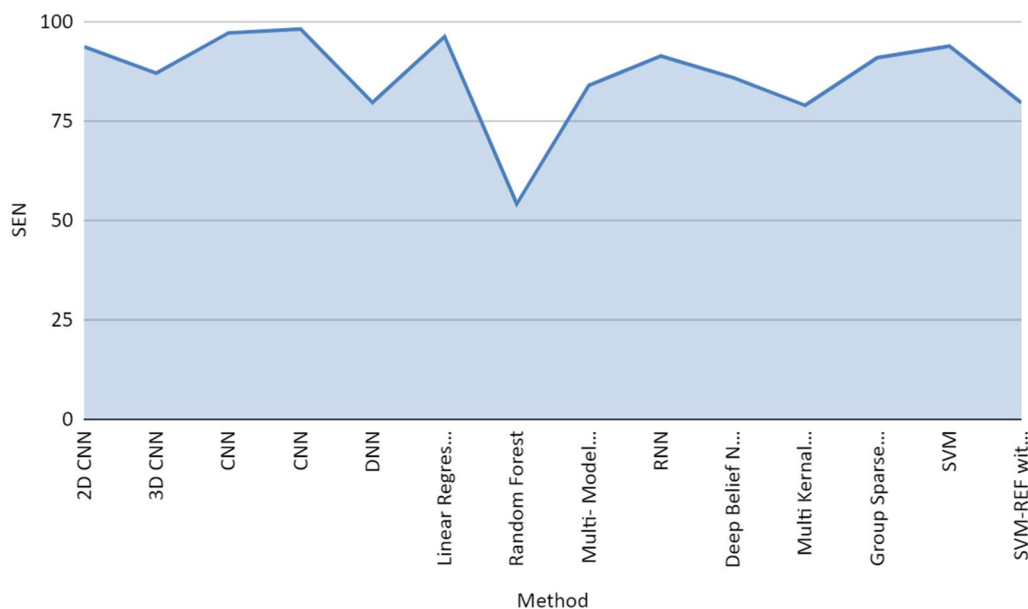


Fig. 9 Stacked area chart of the sensitivity drive from different methods used in the study

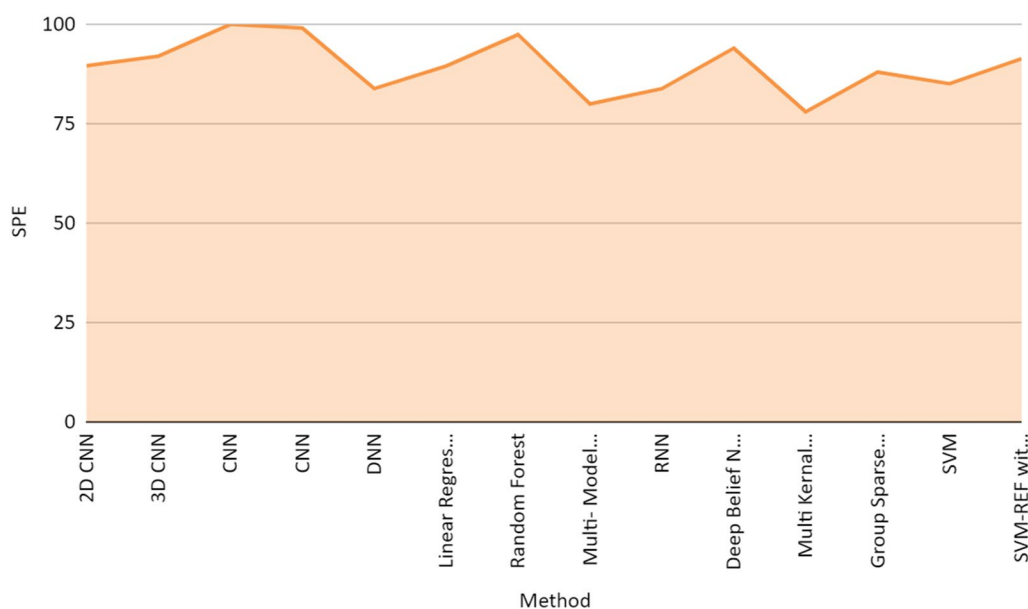


Fig. 10 Stacked area chart of the specificity drive from different methods used in the study

overfitting, and increasing robustness. By leveraging the strengths of multiple models, ensemble learning can help produce more accurate and reliable results.

Future research can expand on the current literature survey and the limitations mentioned in this study. It is recommended that future research must explore

additional modalities not covered in the current study. Additionally, leveraging ensemble learning techniques to combine results from multiple models may be beneficial in investigating the early stages of Alzheimer’s disease [46]. Further Explainable AI (XAI) can be used for the explanation on the reliability and stability of the model

used. XAI techniques aims to provide transparent and interpretable insights into the decision-making processes of machine learning models, particularly in complex domains such as AD diagnosis and classification [47]. To overcome the challenge of limited sample size, multiple neuroimaging databases or transfer learning could be utilized. Overall, future studies should prioritize achieving robust and accurate outcomes.

#### Acknowledgements

The corresponding author is thankful to the co-authors for their support to complete the manuscript. We are also thankful to the anonymous reviewers and Editors for their perceptive comments on the manuscript.

#### Human and animal ethics

Not applicable.

#### Author contributions

All the authors contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript. Conceptualization; ADA. Data curation; SSV. Formal analysis; PC. Investigation; TC. Methodology; AAE. Project administration; MN. Resources; AMK. Software; ADA. Supervision; SSV. Validation; PC. Visualization; TC. Roles/writing—original draft; MN. Writing—review and editing; AAE.

#### Funding

No funding was received to assist with the preparation of this manuscript.

#### Availability of data and materials

Open-source databases accelerate research work in the field of diagnosis and prognosis of Alzheimer's diseases by providing neuroimages. Some of the most common and reliable databases are Alzheimer's disease neuroimaging initiative (ADNI) and OASIS.

#### Declarations

##### Ethics approval consent to participate

The manuscript is original and has not been published, accepted for publication or under editorial review for publication elsewhere.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 27 December 2022 Accepted: 2 June 2023

Published online: 14 July 2023

#### References

- Alzheimer's Association (2019) Alzheimer's Disease Facts and Figures. Alzheimer's Association Report, 01 March 2019 15:321. <https://doi.org/10.1016/j.jalz.2019.01.010>
- Bhushan I, Kour M, Kour G, et al. Alzheimer's disease: Causes and treatment – A review. *Ann Biotechnol.* 2018; 1(1): 1002.
- Zhang D (2012) Predicting future clinical changes of MCI Patients using longitudinal and multimodal biomarkers. *PLoS ONE* 7:1–15
- Wee C-Y, Suk H-II (2013) Discriminative Group Sparse Representation for Mild Cognitive Impairment Classification. Springer International Switzerland, Cham, pp 131–138
- Verma SS, Prasad A, Kumar A (2022) CovXmlc: High performance COVID19 detection on X-ray images using multi-model classification. *Biomed Signal Processing Control* 71:103272
- Popuri K, Donghuan Lu (2018) Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG- PET images. *Sci Rep* 8:1–13
- Wolz R, Tong T (2014) Multiple instance learning for classification of dementia in brain MRI. *Med Image Anal* 18:808–818
- Jo T, Nho K, Saykin AJ (2019) Deep Learning in Alzheimer's Disease: Diagnostic classification and prognostic prediction using neuroimaging data. *Front Aging Neurosci* Aug 20;11:220. <https://doi.org/10.3389/fnagi.2019.00220>. PMID: 31481890; PMCID: PMC6710444.
- Lee, G., Nho, K., Kang, B. et al. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Sci Rep* 9, 1952 (2019). <https://doi.org/10.1038/s41598-018-37769-z>.
- Ding Yi, Xiao Z et al (2017) Brain MR image classification for Alzheimer's disease diagnosis based on multifeature fusion. *Comput Math Methods Med.* 2017 <https://doi.org/10.1155/2017/1952373>
- Rolls ET, Huang CC, Lin CP, Feng J, Joliot M. Automated anatomical labeling atlas 3. *Neuroimage.* 2020 Feb 1;206:116189. <https://doi.org/10.1016/j.neuroimage.2019.116189>. Epub 2019 Sep 12. PMID: 31521825.
- Munilla J, Ortiz A et al (2016) Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *Int J Neural Syst.* 2016 Nov;26(7):1650025 <https://doi.org/10.1142/S0129065716500258>. Epub 2016 Apr 4. PMID: 27478060.
- Jin Y, Huang L (2016) Longitudinal clinical score prediction in Alzheimer's disease with soft-split sparse regression based random forest. *Neurobiol Aging* 46:180–191
- Aparna T R, M Subramoniam (2021) Deep learning based prediction of Alzheimer's disease from magnetic resonance images. *EESS.*
- Hurley P, Challis E (2015) Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. *Neuroimage* 112:232–243. <https://doi.org/10.1016/j.neuroimage.2015.02.037>. ISSN 1053-8119
- Bhadra S et al (2022) An insight into diagnosis of depression using machine learning techniques: a systematic review. *Curr Med Res Opin* 38(5):749–771. <https://doi.org/10.1080/03007995.2022.2038487>
- Kumar C et al (2022) Autism spectrum disorder diagnosis and machine learning: a review. *Int J Med Eng Inf* 14(6):512–527. <https://doi.org/10.1504/IJMEI.2022.126522>
- Sharma M et al (2021) Early diagnosis of rice plant disease using machine learning techniques. *Arch Phytopathol Plant Prot* 55(3):259–283. <https://doi.org/10.1080/03235408.2021.2015866>
- Kumar CJ et al (2021) The diagnosis of ASD using multiple machine learning techniques. *Int J Dev Disabil Jun* 10;68(6):973–983 <https://doi.org/10.1080/20473869.2021.1933730>. PMID: 36568623; PMCID: PMC9788716.
- Sharma M et al (2022) Ensemble averaging of transfer learning models for identification of nutritional deficiency in rice plant. *Electronics* 11:148. <https://doi.org/10.3390/electronics11010148>
- Martin MH, Cenek. (2018) Survey of Image Processing techniques for brain pathology diagnosis: challenges and opportunity. *Front Robot Al.* <https://doi.org/10.3389/frobt.2018.00120>
- He T, Nguyen M et al (2020) Alzheimer's Disease Neuroimaging Initiative. Predicting Alzheimer's disease progression using deep recurrent neural networks. *Neuroimage.* 2020 Nov 15;222:117203. <https://doi.org/10.1016/j.neuroimage.2020.117203>. Epub 2020 Aug 4. PMID: 32763427; PMCID: PMC7797176.
- Yue XGL (2019) Hierarchical feature extraction for early Alzheimer's disease diagnosis. *IEEE Access* 7:93752–93760
- Goryawala M, Zhou Q, Barker W, Loewenstein DA, Duara R, Adjouadi M. Inclusion of neuropsychological scores in atrophy models improves diagnostic classification of Alzheimer's disease and mild cognitive impairment. *Comput Intell Neurosci.* 2015;2015:865265. <https://doi.org/10.1155/2015/865265>. Epub 2015 May 25. PMID: 26101520; PMCID: PMC4458535.
- Jenny Benois-Pineau, Karim Aderghal (2018) "FuseMe: Classification of sMRI images by fusion of Deep CNNs in 2D+ $\epsilon$  projections," in Research Gate.
- Safuillin A, Korolev S (2017) Residual and plain convolutional neural networks for 3D brain MRI classification. In *IEEE Xplore, Melbourne*

27. Cheng D, Liu M (2018) Alzheimer's Disease Neuroimaging Initiative. Classification of Alzheimer's Disease by Combination of Convolutional and Recurrent Neural Networks Using FDG-PET Images. *Frontiers in Neuroinformatics*. 2018;12:35. <https://doi.org/10.3389/fninf.2018.00035>. PMID: 29970996; PMCID: PMC6018166.
28. Choi H (2018) Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Sci Direct* 344:103–109
29. D Cheng, M Liu (2017). CNNs Based Multi-Modality Classification for AD Diagnosis. In 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. Shanghai: IEEE Xplore
30. Manhas N, Sharma D, Kumar V (2019) Comparative study to measure the performance of commonly used machine learning algorithms in diagnosis of Alzheimer's disease. *J Multimed Inf Syst* 6:75–80
31. Liu H, Yan Z, Klette WQ, Ji R (2019) Early diagnosis of Alzheimer's disease using deep learning. in *Association for Computing Machinery, Jeju*
32. Chintalapudi G, Amenta N, Battineni F (2020) Comparative Machine Learning Approach in Dementia Patient Classification using Principal Component Analysis. In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART*; ISBN 978-989-758-395-7; ISSN 2184-433X, SciTePress, pages 780-784. <https://doi.org/10.5220/0009096907800784>
33. Nigam A, Pushkar S, Kumari R (2020) Machine learning technique for early detection of Alzheimer's disease. *Microsyst Technol* 26:3935;1-10
34. Subasi E, Alickovic A (2020) "Automatic detection of Alzheimer disease based on histogram and random forest. Springer International Publishing, Cham, pp 91–96
35. Patil S, Shashidhar H, Parameshachari BD, Madiwalar S (2020) Classification and investigation of Alzheimer disease using machine learning algorithms. *Biosci Biotechnol Res Commun* 13:15–20, October 2020. *Biosci.Biotech.Res.Comm. Special Issue Vol 13 No 13 (2020) Pp-15-20*, <https://doi.org/10.21786/bbrc/13.13/3>
36. Zeng D, Jia A, Huang L, Frizzell Y, Song T, Pan X (2020) Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning. *Front Neurosci* 14:259
37. Seifeddine Mechti, Roobaea Alroobaea. (2021).Alzheimer's Disease Early Detection Using Machine Learning Techniques. Research Square
38. Vijayalakshmi S, Savita MS (2021) Alzheimer's disease detection through machine learning. *Ann Romanian Soc Cell Biol* 25(3):2782–2792. EID: 2-s2.0-85104586144
39. Yang MQ, Li Q (2021) "Comparison of machine learning approaches for enhancing Alzheimer's disease classification." *PeerJ*
40. Yang S, Feng Z, Wang M, Ahmad AS, Khan R, Maqsood M, Yaqub M, Mehmood A (2021) A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. *Neuroscience* 460:43–52
41. Noor MBT, Zenia NZ, Kaiser MS et al (2020) Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain Inf* 7(11):1–21. <https://doi.org/10.1186/s40708-020-00112-2>
42. Mukherji D, Mukherji M, Mukherji N et al (2022) Early detection of Alzheimer's disease using neuropsychological tests: a predict-diagnose approach using neural networks. *Brain Inf* 9(23):1–14. <https://doi.org/10.1186/s40708-022-00169-1>
43. Islam J, Zhang Y (2018) Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain Inf* 5(2):1–9. <https://doi.org/10.1186/s40708-018-0080-3>
44. Hajamohideen F, Shaffi N, Mahmud M et al (2023) Four-way classification of Alzheimer's disease using deep siamese convolutional neural network with triplet-loss function. *Brain Inf* 10(5):1–12. <https://doi.org/10.1186/s40708-023-00184-w>
45. Yang S, Bornot JMS, Fernandez RB et al (2021) Integrated space-frequency-time domain feature extraction for MEG-based Alzheimer's disease classification. *Brain Inf* 8(24):1–11. <https://doi.org/10.1186/s40708-021-00145-1>
46. Merone M, D'Addario SL, Mirino P et al (2022) A multi-expert ensemble system for predicting Alzheimer transition using clinical features. *Brain Inf* 9(20):1–11. <https://doi.org/10.1186/s40708-022-00168-2>
47. Lombardi A, Diacono D, Amoroso N et al (2022) A robust framework to investigate the reliability and stability of explainable artificial intelligence

markers of mild cognitive impairment and Alzheimer's disease. *Brain Inf* 9(17):1–13. <https://doi.org/10.1186/s40708-022-00165-5>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---