Review Article

Single and Combined Neuroimaging Techniques for Alzheimer's Disease Detection

Morteza Amini [ID],[1] **Mir Mohsen Pedram** [ID],[2,3] **Alireza Moradi** [ID],[4,5] **Mahdieh Jamshidi** [ID],[6] **and Mahshad Ouchani** [ID][7]

[1]Department of Cognitive Modeling, Institute for Cognitive Science Studies, Shahid Beheshti University, Tehran, Iran
[2]Department of Electrical and Computer Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran
[3]Department of Cognitive Modeling, Institute for Cognitive Science Studies, Tehran, Iran
[4]Department of Clinical Psychology, Faculty of Psychology and Educational Science, Kharazmi University, Tehran, Iran
[5]Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran
[6]Department of Mathematical Sciences, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran
[7]Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran

Correspondence should be addressed to Morteza Amini; mor_amini@sbu.ac.ir

Alzheimer's disease (AD) consists of the gradual process of decreasing volume and quality of neuron connection in the brain, which consists of gradual synaptic integrity and loss of cognitive functions. In recent years, there has been significant attention in AD classification and early detection with machine learning algorithms. There are different neuroimaging techniques for capturing data and using it for the classification task. Input data as images will help machine learning models to detect different biomarkers for AD classification. This marker has a more critical role for AD detection than other diseases because beta-amyloid can extract complex structures with some metal ions. Most researchers have focused on using 3D and 4D convolutional neural networks for AD classification due to reasonable amounts of data. Also, combination neuroimaging techniques like functional magnetic resonance imaging and positron emission tomography for AD detection have recently gathered much attention. However, gathering a combination of data can be expensive, complex, and tedious. For time consumption reasons, most patients prefer to throw one of the neuroimaging techniques. So, in this review article, we have surveyed different research studies with various neuroimaging techniques and ML methods to see the effect of using combined data as input. The result has shown that the use of the combination method would increase the accuracy of AD detection. Also, according to the sensitivity metrics from different machine learning methods, MRI and fMRI showed promising results.

1. Introduction

Alzheimer's disease (AD) can be considered a gradually progressive neurodegenerative disease process that involves gradual synaptic integrity and loss of cognitive functions [1]. Early detection of AD will help to prevent catastrophic damage to the brain. One of the most significant signs and biomarkers for AD are beta-amyloid, tau protein, and miRNA [2]. This marker has a more critical role for AD detection than other diseases because beta-amyloid can extract complex structure form [3]. Detecting biomarkers

deposition and brain structure examination with neuroimaging techniques like functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) approaches have been widely exploited nowadays. Amyloid PET was used to determine brain amyloid plaque load scores as a biomarker [4]. Observation of fMRI techniques will help detect dementia and change in neuron connections, determining the change in brain function. On the other hand, the level of amyloid deposition in certain parts of the brain, which can be seen by amyloid PET biomarker, will help to survey AD severity for patients [5]. However, especially in

cases of dementia diagnosis, mentioned biomarkers could not help accurately identify or predict cognitive deterioration due to individual threshold differences in each subject [6]. The use of PET for AD detection will consume more time, and because of isotope injection, it is considered an invasive technique. MRI and its branches like fMRI for AD detection and classification are other areas of research in neuroimaging. For the change in structure detection use of the fMRI technique is more convenient than other neuroimaging techniques.

The better change in brain detection structure using structural magnetic resonance imaging (sMRI) can help. Suppose that, with the survey of sMRI, enough structural features have not been extracted. In that case, resting-state fMRI (rs-fMRI) can provide more valuable and complementary information to distinguish early-stage dementia and AD detection in each patient [7]. Combining fMRI and PET into one unique scanner has allowed the researcher to explore the underlying neurochemistry of brain function in more detail [8]. A combination of PET and fMRI techniques resulted in spatial and quantifiable inconsistencies as active data acquisition. By the combination of PET and fMRI, Wehrl et al. [9] extracted functional connectivity of the subjected rat brain. In a nutshell, the fMRI technique will help extract nine well-known biological neural networks in brain structure.

In contrast, the PET technique identified seven glucose metabolism-related biological networks. Different studies have shown comprehensive and complementary information using combination techniques to decode brain function and brain networks further, so that the question of which brain neuroimaging technique will be more helpful and practical for Alzheimer-related disease (ARD) detection and classification arises. In this research, we focus on different branches of AI (artificial intelligence) like machine learning (ML) and deep learning (DL) algorithms for ARD prediction with PET, fMRI, sMRI, and combination methods. Increasing the number of patients with AD-related problems in the future is inevitable. It has been estimated that 1 of 85 individuals in 2050 would suffer from AD-related disease, so, with growth in the number of patients with AD-related problems [10] and with new corona virus pandemics emergence in 2019, different studies have categorized patients with neurodegeneration problems like AD, mild cognitive impairment (MCI), and other ARD at high risk for COVID-19 [11]. Different stages of AD and its complications will be associated with high morbidity and mortality rate. ARD will cause memory capacity loss. Most Alzheimer's patients will forget how to correctly conduct the recommendations from public health authorities or World Health Organization (WHO) to reduce the COVID-19 spread in a high-risk community.

For example, WHO's known recommendations are constant hand washing, covering mouth and nose when coughing or sneezing, monitoring physical and temperature conditions for reporting symptoms of COVID-19, and preserving at least 6 ft physical distance from other people, especially elderly peoples [11]. This specific order is crucial to AD patients, especially when considering these people's age

groups, which usually are more than 65 years. Even patients with MCI or milder dementia may forget to conduct these procedures due to oblivion or depression. Those with more severe dementia cannot correctly comprehend or remember most of these orders due to the strictness of their short-term memory.

Another complication of ARD consists of financial problems. Single Medicare beneficiaries diagnosed with ARD have a higher probability of missing payments on credit accounts. These negative financial consequences continued after the progress of ARD and will cause 10% to 15% of missed payments. These financial complications from AD were common in none of the educational college groups [12]. With this significant complication, proper and fast detection of Alzheimer in the early stages will help patients prevent financial and physical complications. ML and DL algorithms for ARD classification and detection have gained much attention in the past decade. With the growth in computational power and emergence of more sophisticated and supervised algorithms like convolutional neural network (CNN) development of artificial intelligence (AI) application in healthcare has increased rapidly [13, 14]. All artificial intelligence models will use some training data such as pictures from neuroimaging techniques and other electronic healthcare data to extract full features or direct samples to classify, detect, and recognize ARD. Numerous ML applications involve tasks that can be set up as supervised and semisupervised learning [15]. ML algorithms often have reached more than 96% to classify AD [16, 17]. This state-of-the-art result for binary classification of AD is no surprise due to the structure of DL algorithms and the usage of neuroimaging data. Proper preprocessing of input data into the balance group of data mentioned result is no longer a surprise. Nevertheless, the vacancy of specific research on the effect of single and combined neuroimaging techniques for reaching the mentioned results and comparing them is increasing. This paper reviewed AD detection, early recognition, and classification using different machine learning applications with fMRI, PET, sMRI, and combined neuroimaging techniques.

2. Data Set

Brain imaging gathering procedure can be categorized as noninvasive like fMRI and sMRI techniques or invasive techniques like PET. For example, rs-fMRI is a neuroimaging technique commonly used to study the progressive and pathogenic procedure of neurodegeneration diseases like AD. Other techniques we have surveyed in this article impacted the brain's specific marker or specific area. All images presented in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) repository, which can be found at https://www.loni.ucla.edu/ADNI and https://www.adni-info.org. The ADNI database contains 1.5 T and 3.0 T t1w MRI scans for AD, MCI, and cognitively normal conditions in various patients with different ethnic groups. This repository offers data as the image for three categories of patients: NC, AD, and MCI. These three categories will comprise the whole condition of each patient in this research.

2.1. MRI. This data gathering technique will use radio waves and magnetic fields to acquire compelling images with high discriminative features and high-resolution images. This technique can return 2D and 3D images of brain structures with high quality. This technique would not use any harmful radiations from X-rays or other radioactive tracers, so it will be categorized as a noninvasive and nonharmful neuroimaging technique. The most used MRI techniques for AD detection and classification comprise sMRI, which helps technicians evaluate brain volumes in pictures to detect brain loss of tissue, cells, neurons, and other destructive changes. Brain degeneration is an unavoidable component of ARD that causes memory loss and self-unawareness in patients [18]. Figure 1 shows an example of an MRI picture which has been used to detect brain decrease in tissue. In a nutshell, the MRI technique will exploit the nucleus of hydrogen atoms as small magnets as a tracer [19]. Then the vibration of this nucleus of hydrogen atoms can be manipulated to become a tracer and generate a signal that will turn into an image.

2.2. fMRI. fMRI can be categorized as noninvasive technique because this technique focuses on measuring and mapping brain activities without any injected tracer to patients' bodies. With the change in the body's activity, neuron activity in the brain constantly fluctuates [20]. This technique is also using the effect of magnetic fields on brains in order to gather data. The cylindrical tube of an MRI scanner houses a potent electromagnet field, which will be used to gather information whereas patients are conducting some activities. As it has been told before, this strong field will affect the magnetic behavior of atoms. Typically the atomic nuclei orientation is stochastic and does not create any specific pattern. However, under the encouragement of a magnetic field, the pattern of nuclei becomes allied with the direction of the field and makes a specific pattern. As stronger magnetic field becomes significant, the effect will be on the nuclei and degree of alignment. At last, with these powerful aligned signals from these individual nuclei, measurements of these pales will be possible. In fMRI neuroimaging techniques, the behavior of magnetic signal from the core of hydrogen atoms in a water molecule is detected as a tracer. The different neural activity of the body of different parts of the brain will increase; therefore, an improvement in demand for oxygen will be generated (see Figure 2).

With high blood pressure and movement of the capillary blood cells, the local response in a different part of the brain will increase biological neural activity. Red blood cells have hemoglobin. Hemoglobin is diamagnetic when it conveys oxygen in blood and paramagnetic when it conveys carbon dioxide [21], for gathering MRI picture. This difference in magnetic properties will be practical to use. As it has been said before, since the amounts of oxygen that each blood cell will transform will be different according to the different levels of biological neural activity in the brain, these can be used as markers to detect brain activity. A sample of fMRI pictures containing 64 different images of a patient was surveyed in this research. The fMRI picture of the brain is shown in Figure 2. So, with gathering different parts of the brain, brain structure will be revealed, and with different actions in the body like the movement of arms or feet, different parts of the brain will be activated and can be seen in the fMRI scans.

The use of fMRI scans for AD detection was widely used last decade [22, 23]. A different part of brain activity can be discriminated even from a different view with this image. Some pictures from a different region of the brain are shown in Figure 3.

2.3. rs-fMRI. rs-fMRI focuses on low-frequency that is defined as less than 0.1HZ impulsive fluctuations in the level of blood cells that carry oxygen in the brain section. The whole procedure of data gathering rs-fMRI occurs when a patient is at rest situation [24]. The main key characteristic in rs-fMRI was synchronism of low-frequency fluctuations in the tracer signal arising from the right and left primary motor regions of cortices at rest condition of each patient for surveying. With this neuroimaging technique, most researchers have reached a connectivity pattern that is sufficiently close to the activation pattern from a two-sided finger-tapping task [25]. It extracted abnormal patterns in the resting situation of each patient. It disrupted connectivity in various parts of biological neural networks so the brain can help clarify some of the motor and nonmotor shortfalls seen in patients with ARD [26]. A sample of the rs-MRI picture is shown in Figure 4. This picture contains 54 stages of patients in resting condition. The size of samples in this figure is $256 * 64$.

So, with different stages of the brain, even in resting conditions, different patterns can be exploited to classify AD patients.

2.4. PET. PET is the most exact and delicate neuroimaging technique for capturing the molecular image, which helps extract communications and pathways within humans' brains [27]. The specificity rises from the variety of positron-emitting radionuclides choice, which can be defined as specific biomarkers for pathway recognition, biochemicals, and pharmaceuticals without disturbing their biological function [28]. Furthermore, the radiation used as a radio-labeled tracer can be sensed when the reflect wave has been reported above the low natural radiation background. The initial provocation for PET use was for human brain studies. Researchers have used PET for gathering anatomical and biological complexity from brain organs. Because of the use of injected tracer in this procedure, this technique will be considered an invasive method [29]. The compassion and chronological resolution of PET scanners provided to researcher scans with kinetical features that constants of neurotransmitter pathways and binding could be extracted from correct input data [30]. The main concern of the PET technique is to gather more high-quality pictures with better discriminative spatial resolution features and increased axial coverage. With the success of this technique, most researchers worldwide have used this picture for regional brain activation detection. PET pictures are sometimes helping to

Figure 1: Example of magnetic resonance imaging (MRI).
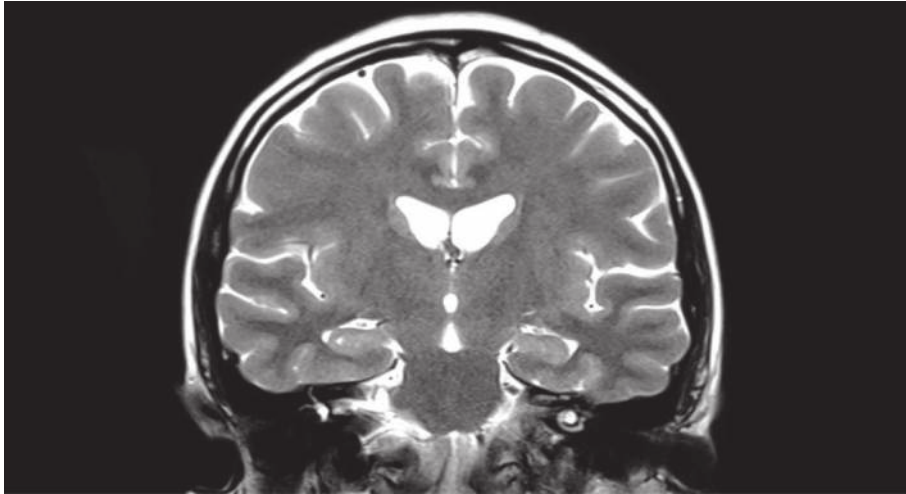

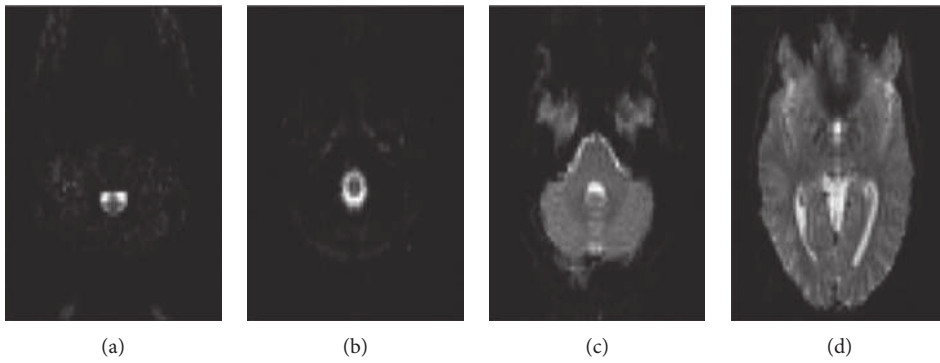(a)   (b)   (c)   (d)

Figure 2: Example of functional magnetic resonance imaging (fMRI) in different stages. (a) Slice number: 0. (b) Slice number: 10. (c) Slice number: 20. (d) Slice number: 30.
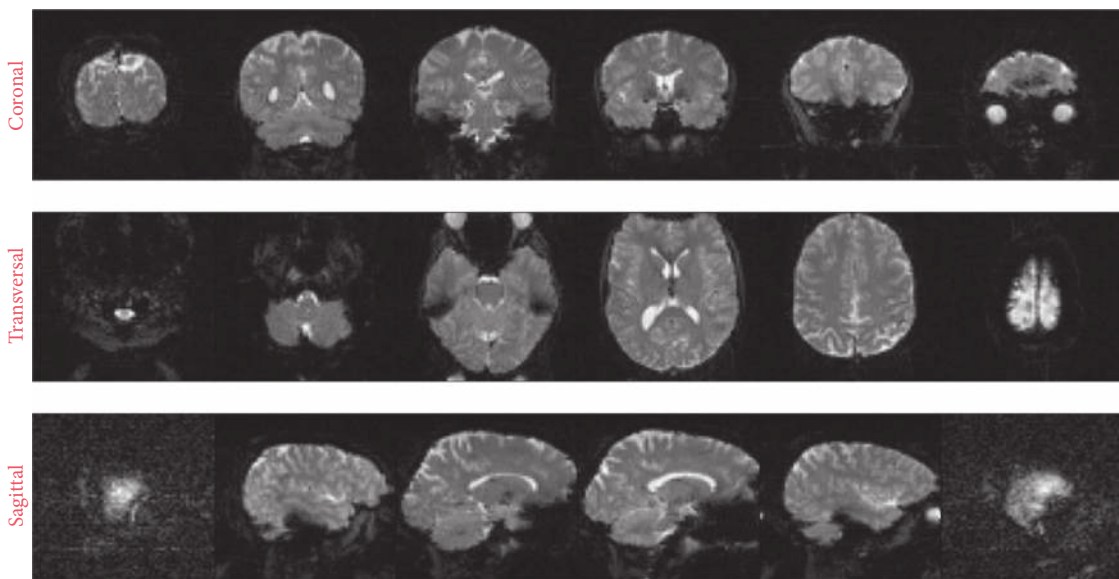

Figure 3: Example of fMRI in different stages.

(a)   (b)   (c)   (d)
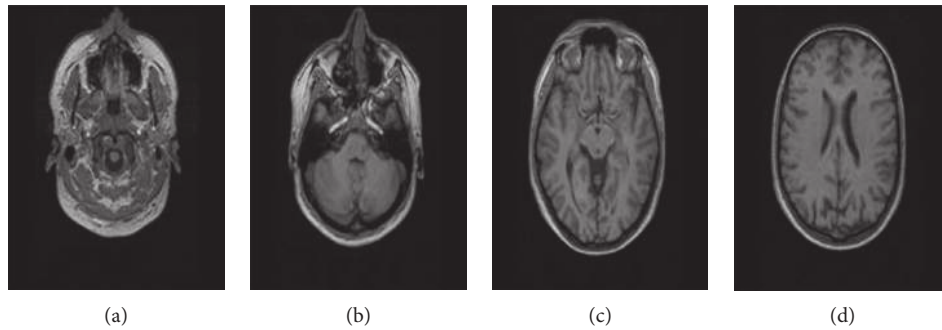
Figure 4: Example of rs-fMRI in different stages. (a) Slice number: 0. (b) Slice number: 10. (c) Slice number: 20. (d) Slice number: 30.

detect diseases like Alzheimer's faster than other imaging procedures [31]. Some samples of PET with $336 * 336$ are shown in Figure 5. As it has been shown in Figure 5, these PET pictures have more different stages than fMRI or rs-fMRI.

Here the number of different pictures has reached 654. This increase in pictures will cause better observation and PET's image will become more conventional. However, as said before, capturing all of these pictures can lead to 4 to 6 hours of tedious procedure. Also, this neuroimaging technique is invasive [32]. The most used case of PET is to survey the chemical activity in the brain tissue of each subject. These neuroimaging techniques will help to determine the different conditions of each patient, which include brain disorders. The pictures from a PET scan provide diverse data which are uncovered by other kinds of neuroimaging techniques, like computerized tomography (CT) or MRI. A PET scan or a combined CT-PET scan enables neuroscientists to identify illness and measure the condition of the brain of each patient more conveniently. PET images nowadays will be recorded by about 1% of the equivalent couples of photons emitted from the brain of each subject and the learned coincidences data from each patient. This data will be stored in list mode as separate events with a time brand or sorted into arrays, sinogram. Then, using this data cluster of tracers 3D distribution of the tracer can be recreated [33] (see Figure 6).

3. Combination of PET and MRI Techniques

MRI and PET techniques have opposite natures for gathering data from the brain. By combining these neuroimaging techniques, more accurate AD diagnosis or classification tasks can be conducted [34]. In Figure 6, a combination of PET and rs-fMRI has been shown. New researches show that a mixture of one or more biomarkers may deliver complementary material for ARD diagnosis; also, combination data can help to increase the classification accuracy. This combination of biomarkers can be presented as fluorodeoxyglucose positron emission tomography (FDG-PET), sMRI, cerebrospinal fluid (CSF) protein levels, and apolipoprotein-E (APOE) genotype [35]. Although published approaches based on the combination of techniques have applied dissimilar biomarkers to develop a new neuroimaging biomarker for AD, this usage of combination

methods may be limited [36]. The usage of combination techniques has led to early detection in some approaches [36]. Based on the performance of ensemble learning models and multikernel learning success on combination neuroimaging techniques in the last decade, these techniques are popular for AD detection and classification [37, 38]. The combining method showed promising results and can be considered the future of AD detection and classification input data, especially in early detection cases [39, 40].

3.1. Combination Method Preprocessing. After gathering different data set use of various preprocessing procedures is necessary for better prediction of results. As it has been said before use of a combination of neuroimaging techniques has resulted in better classification and detection of ARD. Methods like Dartel are considered preprocessing procedures for combination input data. Dartel is a proper tool for increasing intersubject recording or three-dimensional normalization of functional and structural scans, providing less flattening and improving MRI-PET combination data [41]. Another approach has worked on a framework based on an early union procedure that uses different combination rules to combine opposite data from different biomarker modalities into a single feature vector [42]. In another research, they focused on region of interest (ROI) for gathering complementary information of each neuroimaging method. Many researchers used data from the ADNI database and divided brains based on two atlases: LONI Probabilistic Brain Atlas and Automated Anatomical Labeling. Then baseline images of sMRI and 18F-fluorodeoxyglucose PET were used to calculate average gray-matter density and average relative cerebral metabolic rate for glucose in each region [43]. In 2008 compatible PET detector tools for gathering synchronized PET/MR images of the human brain were conducted. With these new tools and new studies, researchers successfully achieved brain glucose consumption images in two subjected patients using 18F-FDG-PET, MR imaging, and MR spectroscopy [44]. With the combination of data, they established that PET/MR imaging combination is possible in humans' brains. With this combination of data for the first time, a field of new possibilities in molecular imaging areas and ARD detection have been unlocked. The essential step of combination
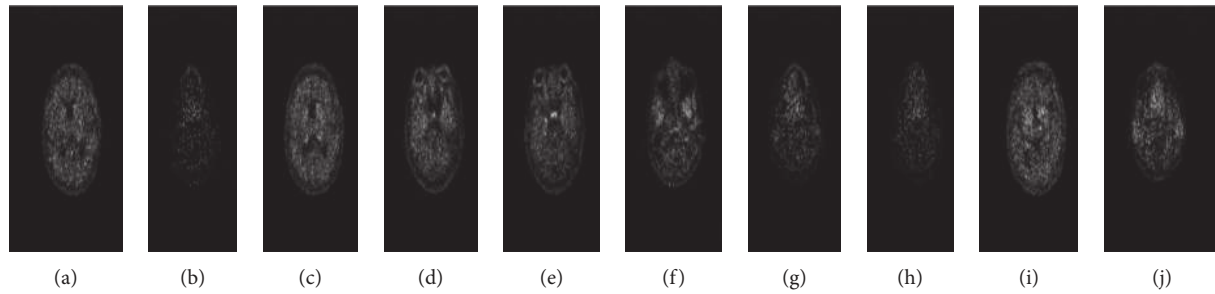
Figure 5: Example of PET in different stages. (a) Slice number: 0. (b) Slice number: 20. (c) Slice number: 40. (d) Slice number: 60. (e) Slice number: 80. (f) Slice number: 100. (g) Slice number: 120. (h) Slice number: 140. (i) Slice number: 160. (j) Slice number: 180.

techniques and preprocessing methods is shown in Figure 7. A typical pipeline for the combination of PET with rs-fMRI is shown in Figure 7.

In the first step, frontal commissure and subsequent commissure correction for all subject images can be extracted with this combination of data. After data gathering use of N4 bias field correction using ANT's toolbox will help correct the intensity of nonhomogeneity for each patient's image [45]. In some cases, elimination of skull has been conducted, which was unnecessary if images were already preprocessed. For the MRI images, aligning them to the MNI152 T1-weighted standard image using a standard procedure will be done next. In normalization steps, the extracted features will throw a standard scalar function, which transforms the array of input matrix data sets into a standard distribution with minimum and maximum of each column vector, which can help to reduce the redundancy and dependency of the data [46]. For structural and functional segmentation of brain into the anatomical area and enumerating these extracted features from each specific ROI from each sMRI image, toolbox with a conventional procedure like NiftyReg with 2 mm Brainnetome Atlas template has been used [47]. After gathering ROI from labeled sMRI images, computed volume of gray-matter tissues in that ROI will be used as an input feature for the detection task.

3.2. Single Method Preprocessing. For processing fMRI pictures into using robust data for early MCI detecting, a handful of researchers have used Data Processing Assistant for Resting-State (or in brief DPARSF) [48–51]. To process fMRI pictures in this platform, users need to arrange their DICOM files and specify their intended parameters. DPARSF then will deliver all the preprocessed data as different variety of desirable data for classification. This desirable data consists of slice timing images, normalized images, smooth data, functional connectivity with specific data, ReHo, ALFF/fALFF, degree centrality, and voxel-mirrored homotopic connectivity (VMHC) results [52]. For PET image processing, using a standard CL pipeline is a conventional method [53]. PET images were intensity normalized using the whole cerebellum as a reference region. In simple image preprocessing, the fusion parameters of combination methods have been eliminated.
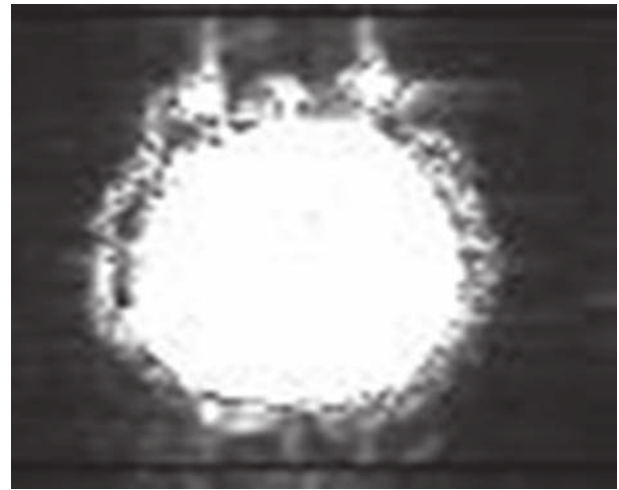

Figure 6: Combination of PET with rs-fMRI pictures.

However, the use of feature selection and normalization part of the procedure remains the same. All of the mentioned algorithms help decrease noise in the picture and use the whole part of the brain, which help model to classify, detect, or recognize ARD. After the acquisition of using complete information, some models will convert pictures into 2D features. With the use of these procedures, the number of features increases significantly. Also, even small shape images like $64 * 64$ or $128 * 128$ have about 4096 or 16384 features. With this amount of information, even power full computational GPU would not conduct the code and render results [54]. So, dimension reduction is a preferable procedure for dealing with image processing tasks. One of the main steps of dimensionality reduction is using correct data and eliminating undesirable data from each picture. Dealing with this task using different norms like l1 and l2 and hybrid classification will help reach a better result. Other studies have used features computed from MRI images to discriminate between different cognitive states related to AD [55]. With the rise in prostate cancer, patient's attention to prostate image segmentation has been conducted. With the use of this data, a semiautomatic method has been designated [56]. This procedure consists of two new algorithms for better feature selection. Experiments on prostate CT images have shown the effect of this method for segmentation and regression tasks [57].
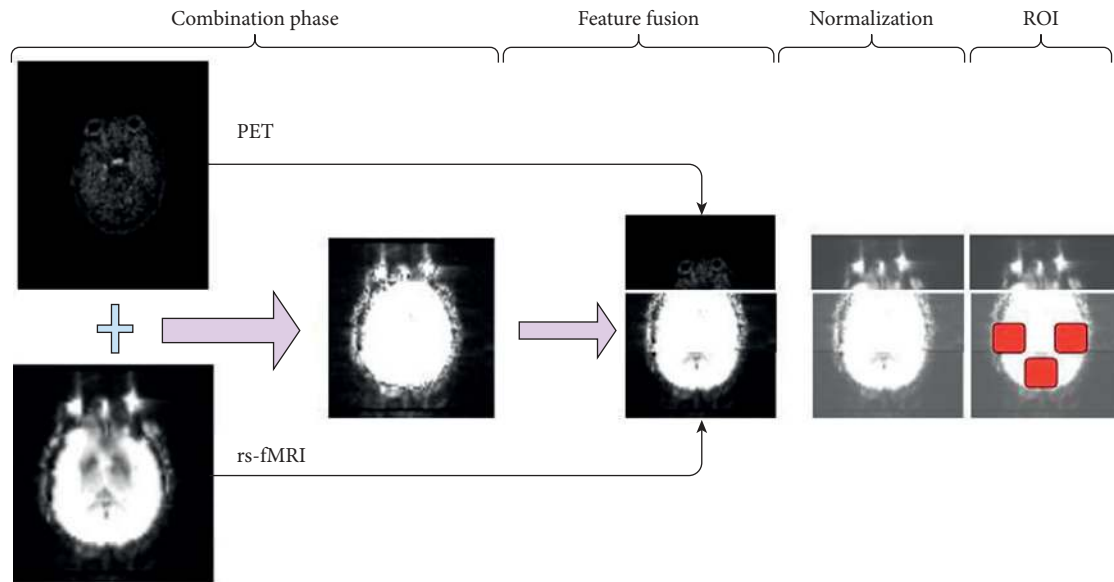
FIGURE 7: Combination of PET with rs-fMRI pictures in the processing of data preprocessing.

Principal Component Analysis (PCA) is another feature selection technique. PCA, in a nutshell, determines the alliance or axis responsible for the most significant amount of variance in the input data set. Respectively, PCA will determine a second axis that must be orthogonal to the first axis responsible for the largest total of remaining data variance. A standard matrix factorization method is called singular value decomposition (SVD) for dealing with this task. This technique will decompose the training data matrix into three different matrices. PCA will be the result of the multiplication of three multiplications. These three matrices will be A, B, and C, where C contains all the principal components used as a principal feature [58]. The whole structure of feature selection and PCA is shown in Figure 8. Effect of proper feature selection and dimension reduction for reaching the better performance of classifier algorithms is necessary. Fewer features will lead to less time consumption in order to train the different models. Using proper features and tuned model state-of-the-art result will be achieved in AD detection and classification.

4. Methods

In the last two decades, due to vast improvement of computational power, more conventional GPU and online platforms for the implementation of artificial intelligence (AI) systems have been developed. So, interest in the use of AI, ML, and DL to synthesize the applications for studying mental health has increased rapidly [59]. Different ML algorithms have been used for the prediction and classification of ARD. The main goal of these different algorithms was to separate different patients into AD, MCI, and NC classes [60]. In ML and DL fields, algorithms have been separated into three categories:

(i) Unsupervised algorithms

(ii) Semisupervised algorithms

(iii) Supervised algorithms

Supervised algorithms refer to algorithms targeting specific targets, and all samples have their target [61]. Most varieties of ML and DL algorithms belong to this category. In this category, gathering samples and labeling them would be tedious. DL can be categorized as a subfield of ML, usually used on big data as input. Different data structures like pictures, time-dependent data, and images can be used for AD classification tasks. It has attracted massive attention in the last few years, especially in image analysis [62]. Several DL architectures such as Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN), and autoencoder (AE) are some examples of these fields which have been used for ARD and classification. Semisupervised learning refers to algorithms that work with a data set that most of its labels are unclear. Most data will be labeled by knowledge about already known labels from data sets [63]. Unsupervised learning refers to algorithms in which labels of the whole data set are not clear. In most cases, most data do not have any labels, making this category very important [64].

4.1. Semisupervised Methods

4.1.1. K Nearest Neighbor (KNN). Most used cases of semisupervised learning algorithms rely on Euclidean distances. The famous semisupervised learning algorithms are K Nearest Neighbor (KNN) and its branch [65]. The work structure here is simply finding a cluster of labeled data, computing the average of this cluster, and then computing Euclidean distances of unlabeled data from this mean of labeled data. Finally, labeling data set based on their nearest known average cluster of data will be done. After labeling the data set, supervised algorithms such as CNN or DNN will classify the data set [66]. How semisupervised algorithms will deal with semilabeled data has been shown in
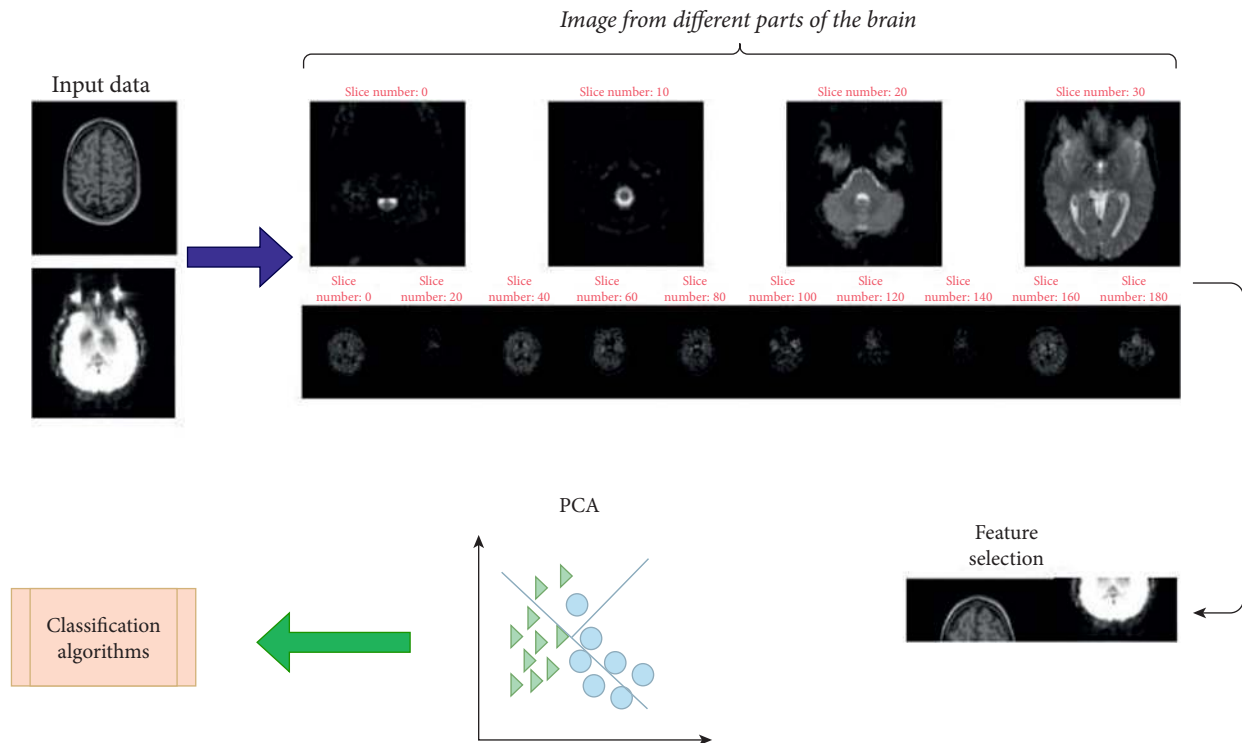
FIGURE 8: Diagram of feature selection and PCA.

Figure 9. Most of the different classification methods will repeat some examples from nature. Semisupervised algorithms will use the procedure of learning in humans. With different years of education, we humans will learn a little labeled information and solve unseen challenges. A variety of different models have been used as a semisupervised algorithm for classification. El-Yacoubi et al. [67] have proposed an approach based on generating each subject cluster and analyzing the correlation of these clusters with NC, AD, and MCI profiles. The main aim of their work was to find the optimal number of clusters and a subset of valuable features that help reach an excellent discriminative algorithm. They used a semisupervised algorithm based on normalized mutual information feature selection, which guides a clustering algorithm to optimize the choice for the number of optimal clusters and the discriminative power of each three-output class. Pohl and Davatzikos [68] used semisupervised algorithms for classification tasks which use both labeled and unlabeled data for training as all semi-supervised algorithms do.

They used clustering methods to deal with unlabeled data; then, for training the labeled data, they used the linear Laplacian support vector machine (LapSVM) [69]. Gorriz et al. [70] Proposed a novel case-based model selection method at their time, which syndicates hypothesis testing from a separate set of expected results and feature extraction. For the training and validation part, they have used a cross-validation strategy for avoiding overfitting. This proposed model will take advantage of proper feature selection. Using good features, this model tries to improve the network's performance on validation and test sets.

4.1.2. Generative Adversarial Network (GANs). GANs were first developed and pioneered in 2014 and, from then until now, have gathered much attention on image generation tasks [71]. The use of GANs as semisupervised methods is one of the most capable areas of real-world application of GANs. In a nutshell, semisupervised GAN (SGAN) is a subset of GANs in which discriminator is a multiclass classifier, and its generator is an expanded CNN. Instead of distinguishing between only two classes, it learns to distinguish three or more classes with the production of fake images. Generator in SGAN is not the essential part, unlike other conventional GANs which have aimed to produce new high-quality data from the useless noisy data set [72]. The structure of SGAN has been shown in Figure 10. As shown in Figure 10, the generative part of the network will expand the dimension of noise to create fake images. The SGAN generator's aim is the same as in the original GAN. The generator of ordinal GANs will take a vector of random like Gaussian noise. It will produce fake examples or samples where the goal is very similar to natural images as input data of the training data set. The goal of the generator is the same in SGANs and GANs. The SGAN discriminator, on the other hand, differs obviously from the original GAN procedures. Discriminators of SGANs will get three sorts of inputs: fake examples produced by the generator model, real examples without any labels from the train set, and real examples with labels from the training data. Instead of binary classification, the SGAN discriminator's goal is to distinguish between real and fake examples and then use the labeled data and fake image to classify the input into different classes. In
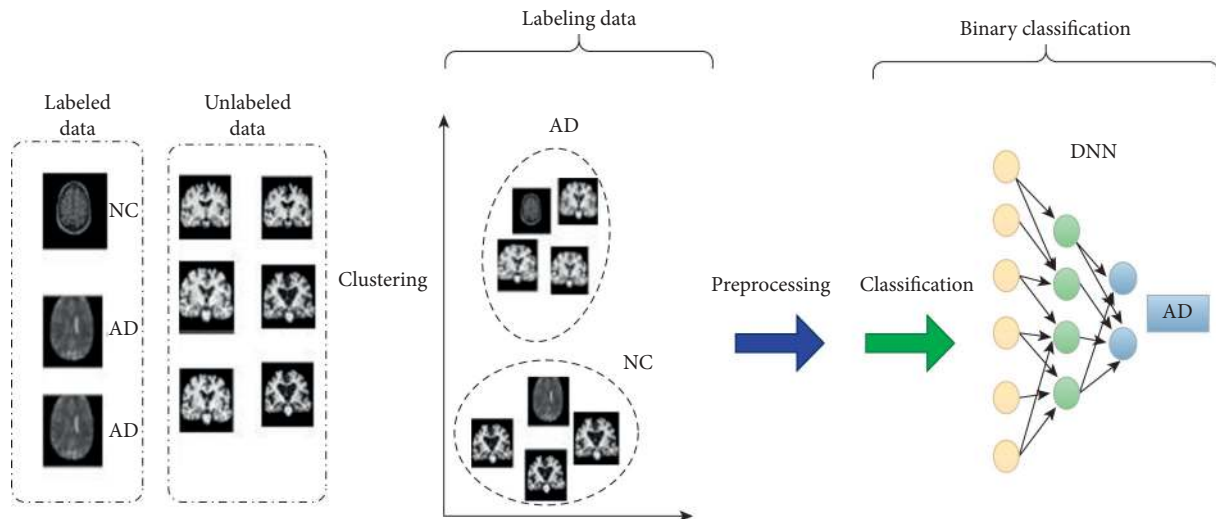
Figure 9: Semisupervised learning method for binary AD classification.
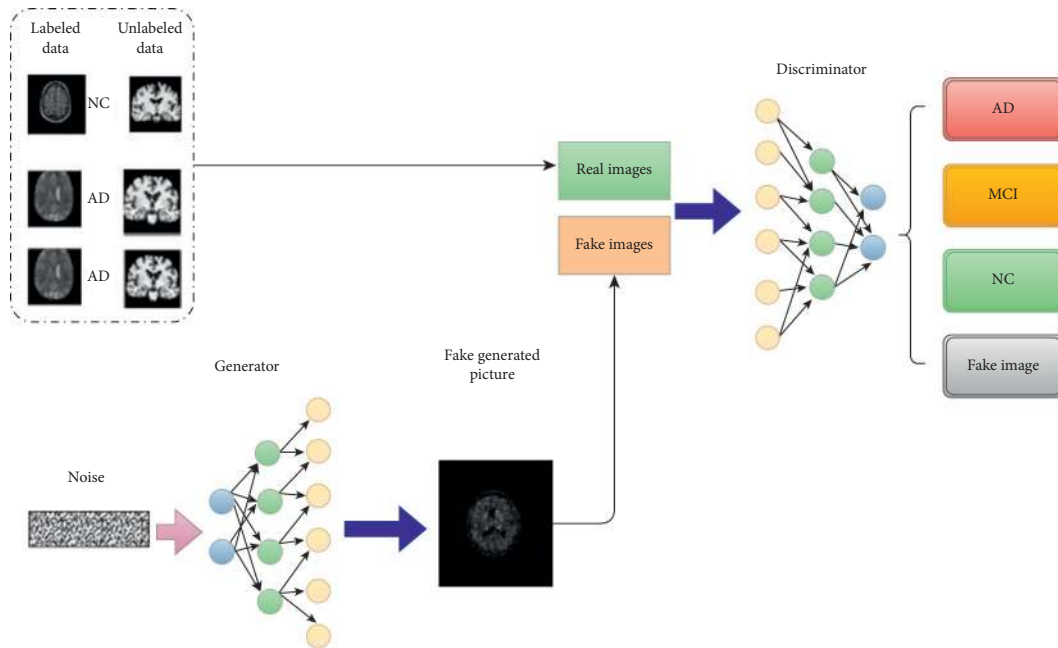
Figure 10: Semisupervised GAN learning method for binary AD classification.

research by Yu et al. [73], the authors used SGAN to predict MCI and AD. They proposed that THS-GAN is designed for semisupervised classification. They have used partially labeled data set input.

Then used the distribution of labels to predict the label for both labeled and unlabeled data and the newly generated samples. Their model can profit from the mechanical information of the brain. Also, they introduced high-order pooling, which helps to exploit more essential features by using the second-order statistics of the holistic MRI images. The result of THS-GAN demonstrates that the classification of MCI vs. NC has been done with 89.29% accuracy. AD vs. NC classification has been done with 95.92% accuracy.

4.2. Supervised Learning. Supervised methods have higher popularity because of their performance. In this section, supervised methods for AD detection have been reviewed. In real-world data sets, much of the data are not correctly labeled. So, the use of supervised methods will be restricted only to the labeled data set. This review first focused on DL algorithms and then used ML algorithms such as support vector machine (SVM) [74].

4.2.1. DNN. The DNN structure is similar to the traditional multilayer perceptron network structure, and with more perceptron layers, the structure of models will get deeper. This deep model can learn the more sophisticated pattern and

relations from input data [75]. The model with a deep layer can determine the best features for classification. DNNs have been used only as supervised methods. We separate the DL method into Conventional Neural Network (CNN), Recurrent Neural Network (RNN), and MLP (Vanilla or Residual Model). 2D CNN is the most used type of CNN for the classification and detection of ARD. The new research focused on using 3D and 4D convolutional layers to extract information from videos [76, 77]. First, we describe CNN because of the common use of this method for AD classification.

4.2.2. Convolutional Neural Network (CNN). The difference between DNN and CNN methods can be described in the connectivity of neurons in different CNN network layers. The connection in the primarily convolutional layer is not connected to all connections of the second layer. The first layer of convolutional layers extracts simple structure from images like orthogonal and diagonal lines. As the CNN structure goes further and gets deeper complex shapes like face and trapezoid shape can be extracted from images [78]. In each layer of CNN, CNN's top layers are connected to a restricted number of neurons in the next or precious layer located within a specific rectangle shape. This building structure of CNN permits the proposed model to focus on a small sublevel of features in the first hidden layer. Model uses them into more extensive higher-level features in the second hidden layer, and the same structure will be repeated until the last layer. This hierarchical structure is typical in real-world images, like ARD. Complete structural work of CNN consists of the following:

  (i) Specifying the convolutional kernels which are defined by a width and heights

 (ii) Specifying the number of input and output channels of each convolutional layer

(iii) Specifying the depth of each convolutional layer must be the same as the number of RGB colors in input data

The structure of CNN consists of a pooling layer that helps whole algorithms work with the smaller size of pictures and have the same performance as full pictures [79]. At the end of the convolutional layer, a flatten layer and stack of multilayer perceptron layers complete the whole structure of CNN. The structure of some CNN is shown in Figure 11.

As shown in Figure 11, CNN will increase the number of channels or depth of the input data while decreasing the widths and height of the input picture. The exact location of a feature is less important than its irregular location comparative to other features in the convolutional layer. It is the idea behind the use of a pooling layer in convolutional neural networks. The pooling layer will extract the essential features from the output of convolutional layers. With the pooling layer, the height and width of the input layer will decrease by a factor of the pooling layer's window. The use of the pooling layer will help to reduce overfitting and computational power, which is needed to train the CNN model. After the convolutional layer, a fully connected layer will be used. This model uses MLP as a feedforward network and

tries to classify input features [80]. Sarraf and Tofighi [81] used the fundamental type of ordinal CNN called vanilla CNN to exploit different patterns from input data to develop a proper model for AD diagnosis among elderly patients. They proposed a state-of-the-art DL-based procedure to distinguish AD from NC using MRI and fMRI. The use of proposed pipelines was performed on a GPU-based powerful device as the computing platform. They have categorized their input data into three parts of train, test, and validation. Their research use of fMRI data has been used for the first time in the DL model as an application to distinguish between AD, MCI, and NC. Spasov et al. [82] have proposed a DL to classify input data, combining sMRI, demographic, neuropsychological, and APOe4 genotyping pictures as input data for classification tasks. The innovation of their work consists of the DL model, which learns to distinguish between MCI vs. AD and AD vs. NC. All the analyses of this work were performed on a subset of the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The data set used in their research consists of 785 participants subcategorized as 192 AD, 409 MCI, and 184 NC. Their research found that the most helpful combination of input data included the sMRI images and the demographic, neuropsychological, and APOe4 data. More and more CNN algorithms have been developed for AD classification. For better understanding, the convolutional layer and the effect of each layer at its input are shown in Figure 12. As shown in Figure 12, as the network goes deeper and deeper, more structure of each data set will get extracted and difference between two input pictures will get more precise.

4.2.3. RNN with CNN. RNN is a subtype of DNN that remembers earlier time-series data and uses this information with present input for predicting the future. RNN is a structure of time-variant algorithms with repeated information along with its layers.

The min different part of RNN which separates them from other DNN models is the structure of hidden states. This hidden state will help to extract useful information in the sequence of data. The same recurrent layer will be used as a stacked layer after each other. The use of this structure will be helpful in input data as video for the AD classification task. It results in the reduction of the complexity parameters, unlike the other DNN [83]. Two main subcategories of RNN algorithms are Long-Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). The main purpose of LSTM is to help to use and maintain the error signals through the structure of network models on the long sequence of data.

This maintenance of signal error will be done with short-term memory and long-term memory. The common activation functions used in RNN structure are sigmoid and tanh, which can help backpropagate error signal through different layers. In LSTM, different gates will be used for the preservation of information and ignoring redundant information. These gates consist of learning gates, forget gates, use gates, and remember gates. LSTM model is similar to a computer's memory. This cell's structure can be used
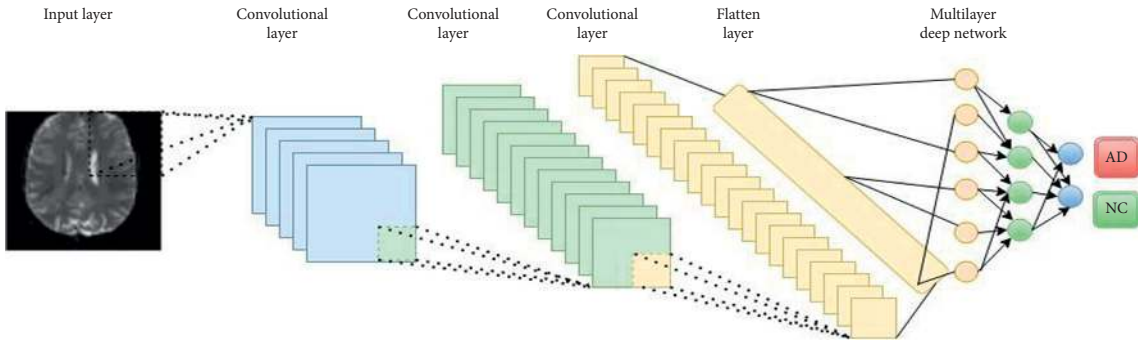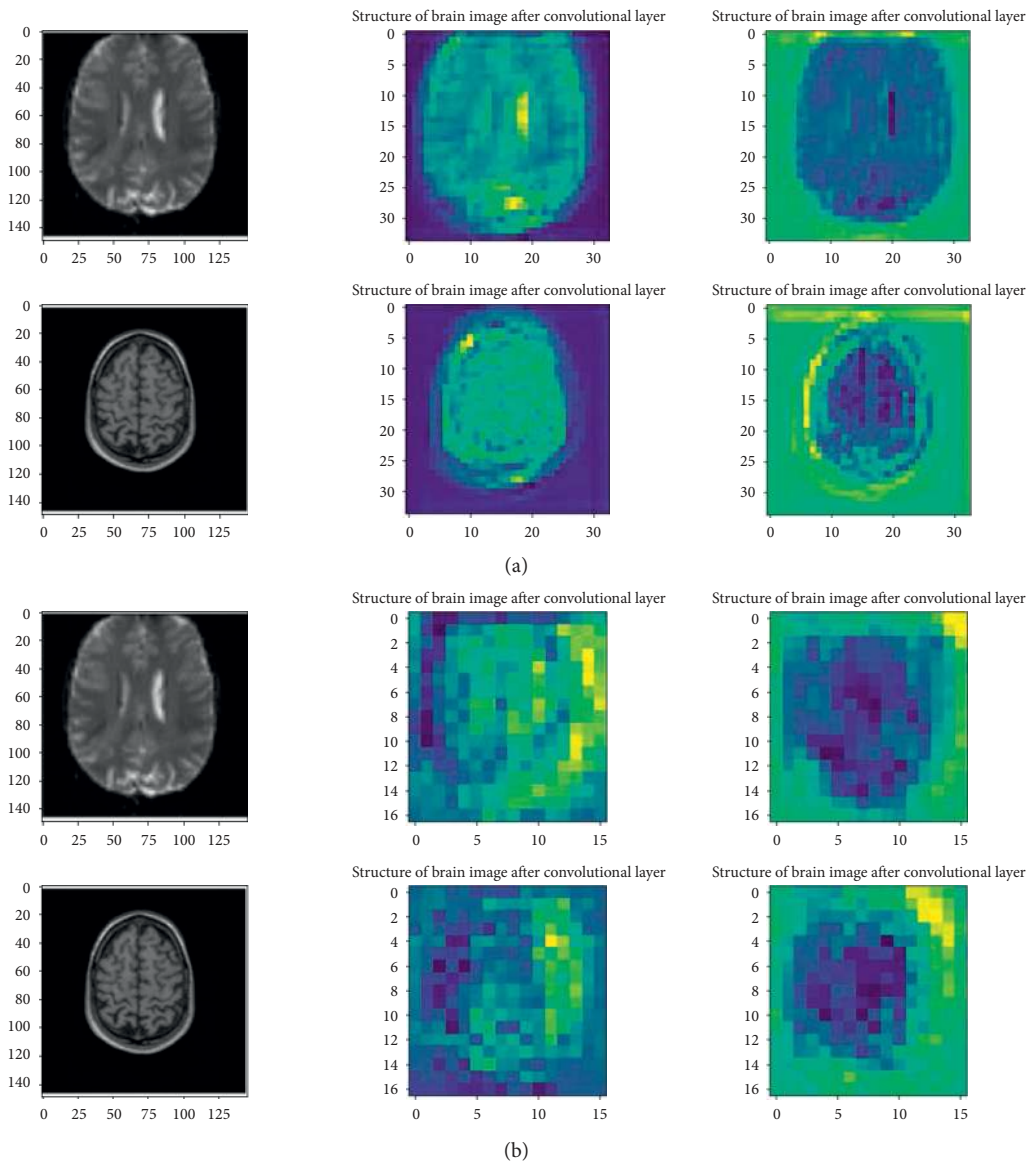
Figure 11: Structure of CNN.
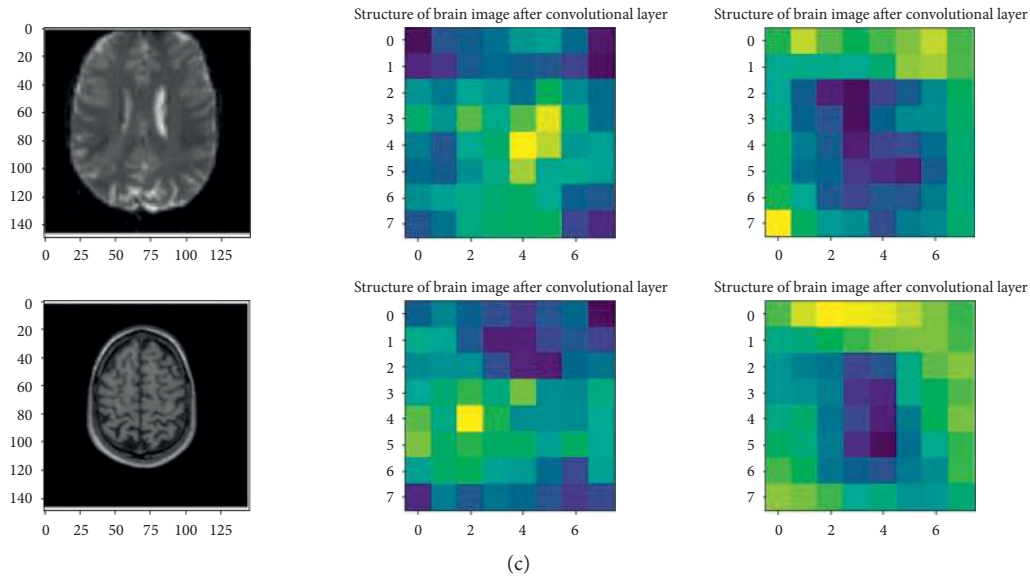

(a)

(b)

Figure 12: Continued.

FIGURE 12: Result of convolutional layer on input image for (a) first convolutional layer, (b) second convolutional layer, and (c) last convolutional layer.

independently to decide which information to store and which information to forget [84]. The structure of the RNN and LSTM network is shown in Figure 13. As shown in Figure 13, the LSTM structure is more complex, but the flow of loss throws each layer for a better optimization process. Karlekar et al. [85] had used language changes as a sign of a patient's cognitive functionality for early diagnosis of AD. In their work, they have used Natural Language Processing (NLP) for AD classification tasks. In the proposed work, they have used CNN, LSTM\RNN to distinguish between language samples from AD and other stages of AD. They have reached 91.1% accuracy with this newly conducted procedure.

4.2.4. Machine Learning Algorithm. The use of supervised ML algorithms for classification is not more different than a semisupervised algorithm for classification. The difference between these two algorithms appears in part of the labeling data set. For a supervised classifier, all data labels are specified, but in semisupervised one, all data labels are not clear. A method that has been used for a handful of researches is SVM. In research by Kloppel et al. [86], they used SVM with the linear kernel to classify MRI scans from proven AD patients and MCI in elderly cases with two different scanning equipment and neuroimaging techniques. Finally, they used these methods to differentiate between patients suffering from AD and patients with frontotemporal lobar degeneration. The result of the classification models consists of 89% accuracy. In another research by Montagne et al. [87], they proposed a model based on a noninvasive neuroimaging technique for early diagnosis of ARD. They have used SVM to classify Alzheimer's disease versus NC group of patients. They have reached better classification rates by focusing on parietal and temporal lobes of brains with SVM. Other ML methods used for AD classification

comprise using second-order derivation or Hessian of loss function for updating weight. Extreme Learning Machine (ELM) is another popular algorithm for classification too.

ELM is a learning algorithm conducted without using multiple different stacked layers and tuning this vast majority of hidden layer and input so that the computation time will decrease [87]. In ELM, unlike other DNN algorithms and SVM, the hidden layer parameters consist of weights and biases. One hidden layer does not need to be tuned after importing the data set and can be generated randomly before the training samples are acquired. This modification will help the network for faster learning processes at the cost of higher loss [88]. In an article by Lama et al. [89], they proposed an AD diagnosis approach using sMRI images to discriminate AD, mild MCI, and NC. They have used SVM and a regularized ELM for prediction. Lama et al. experimented on the ADNI data sets. They showed that regularized ELM with the feature selection techniques could significantly improve the classification accuracy of AD from MCI and NC subjects.

5. Comparison Based on the Different Types of Data

5.1. fMRI, sMRI, and rs-fMRI. In this section use of different algorithms based on fMRI and sMRI pictures has been surveyed. In research by Duc et al. [7], different variants of MRI like sMRI and rs-MRI scans of 331 subject patients for Mini Mental State Examination prediction have been used. In their work, 3-dimensional CNN, a method, has been developed for the mentioned classification task. Task linear regression, support vector regression, bagging-based ensemble regression, and tree regression were used. They have reached a test accuracy of 85.27% for the classification of AD versus NC. Also, it was mentioned that the SVM method with desired features has reached the lowest root mean
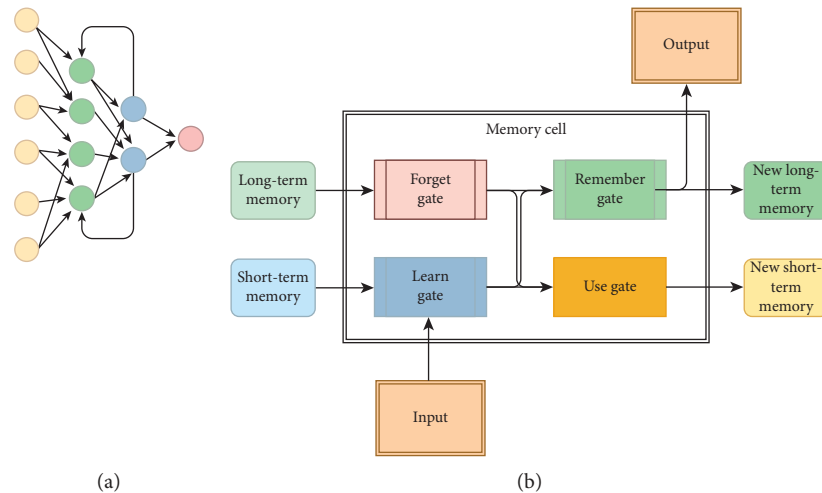
FIGURE 13: Structure of (a) RNN versus (b) LSTM.

square error of 3.27 and the highest R2 value of 0.63. As seen from this work structure, a perfect relation between input data and regression task had not been reported.

In another research, Ramzan et al. [90] have studied the effect of rs-fMRI for multiclass classification of AD and different stages of AD-related disease. They have used one of the famous structures of CNN, which is ResNet-18 [91]. The mentioned structure consists of skip connection for better use of error signal in backpropagation technique for optimization. Skip connection will help the model to get deeper with the vanishing of gradient signals. They have conducted model training from scratch with single-channel input. On the other hand, they also have performed transfer learning with and without an extended network architecture. Transfer learning helped them reach a state-of-the-art result with an average accuracy of 97.92% and 97.88% for all the AD stages prediction. Altinkaya et al. [92] have used superresolution on the low quality of input MRI picture and then used CNN for AD prediction. With superresolution, image processing time has been shortened, and images with high-quality features have been obtained. The result of their study concluded that the performance of proposed methods increased day by day. The resulting accuracy with CNN for AD detection using 302 MRI and fMRI instances was 99.9%.

Korolev et al. [93] have used 3DCNN and achieved better performance without including preprocessing steps like feature extraction in their proposed architecture. These proposed methods consist of CNN and the residual NN. In terms of performance metric Area under Receiver Operating Characteristic (AROC), receiver operating characteristics (ROC) curves, and accuracy have been evaluated. They have proposed a branch of CNN which is called Vox CNN and ordinal ResNet. AD vs. NC achieved the best result with AUC 0.88 and acc 0.79 using VoxCNN and AUC 0.87 with acc 0.80 using ResNet.

Li et al. [94] have used hippocampal MRI as input data. They surveyed 2146 subjects for prediction of MCI in each subject and how this stage will progress and lead to dementia in a time-to-event analysis setup. This study focused on the hippocampus region of the brain. They have reached 0.813 AUROC. For better results use of whole-brain MRI data can help different DL models to reach better classification tasks.

Yang et al. [95] have used SGANs with clustering as the novel semisupervised deep-clustering method. They have surveyed 8,146 scans which included NC, those with MCI, and dementia cognitively. One-fourth of their input data consists of sMRI data. Their proposed method has separated patients into four types of peoples: NC, MCI, relatively more significant memory impairment, and advanced dementia. Results of their work confirmed that the Smile-GAN model was able to cluster participants with 99.9% accuracy even with very severe confusing patterns as inputs. Zeng et al. [88] have proposed a total baseline for the diagnosis of AD. Their proposed model consists of MRI image preprocessing, feature extraction, PCA, and SVM algorithm developments at the end. For optimization, a particle swarm optimization algorithm was proposed to optimize the SVM parameters instead of traditional optimization methods. With their proposed model, they successfully conducted a classification of AD and MCI using MRI scans from the ADNI data set. The proposed algorithm in their research has a state-of-the-art performance compared to other presented methods [87]. Dua et al. [96] used MRI scans from different online repositories to create a diverse data set. They have deployed CNN, RNN, and LSTM individually and as ensemble methods together. The results of the proposed work show that, with the combination of CNN with RNN and CNN with LSTM, an accuracy of 89.75% has been achieved. Meanwhile, they have also used ensemble method with the bagging strategy of the first models. They achieved an accuracy of 92.22%. According to the structure of video and 4D, CNN's use of RNN with this specific data structure has shown promising results [97]. It has been shown that use of 3DCNN and stacked bidirectional LSTM could help the researcher to reach state-of-the-art performance based on accuracy and loss [97]. In research by Kruthika and Maheshappa [76], a DNN has used capsule networks, a branch of CNN. This method will use the benefit of not using

any max-pooling layers. The structure of the capsule net has been developed to use spatial information in the pictures. They have proposed a 3DCNN, which works with video structure too. Their work proposed a method based on CNN and capsule networks for AD prediction using MRI data as input. In the end, they have used the proposed model for creating an application for the AD detection task. They found that both the 3D capsule network method and CNN with pretrained 3D autoencoder improved the predictive performance compared to other structures of CNN trained from scratch. In another research, multiclass classification between AD, MCI, and NC has seen a state-of-the-art CNN called MCADNNet [98]. Using the mentioned method, the researcher reached 92.6% accuracy with distinguishable accuracy of 97% for MCI classification versus AD subjects. Furthermore, even after applying the decision-making algorithm, accuracy rates of 99.77% and 97.5% were achieved for MRI and fMRI pipelines to classify AD versus NC. Amini et al. [99] used CNN and ML models for finding the severity of AD. They have used fMRI pictures as input data set. Also, they have used a sophisticated procedure for converting raw fMRI input data into a valuable data set for training CNN and ML models. The performance of the proposed CNN model for different stages of AD classification was 98.1%, 92.4%, 97.0%, and 100% precision for the low, mild, moderate, and severe status of Alzheimer's patients. The absolute accuracy for their work was reported at 96.7%. They have used only vanilla CNN with one convolutional layer and three fully connected layers. Mentioned structure without any pooling layer is a variant of CNN, which helps to extract features from low dimension feature maps [100].

5.2. PET. The use of PET as a biomarker for AD-related detection is a relatively new procedure. Ozsahin et al. [101] have proposed amyloid-beta plaques combination as a marker. They have implied that mentioned marker should be taken for granted as a "start" of the degenerative process in the brain of most cases. This symptom can be seen earlier than other clinical symptoms, which will appear later in AD subjects. They have used DNN with 18F-florbetapir PET data for automatic classification of four patient groups into four different classes of people, which consist of the following:

(i) AD

(ii) Late MCI

(iii) Early MCI

(iv) Significant memory concern

(v) NC

With this specification, even early stages of AD disease can be detected. Their work on 30% of data as test cases based on sensitivity, specificity, and accuracy was measured as 92.4%, 84.3%, and 87.9%, respectively. Other parts of the second experiment, which consist of the classification of NC versus late MCI images, resulted in sensitivity, specificity, and accuracy of 62.9%, 70.0%, and 66.4%. Other experiments in the classification of NC versus early MCI images have shown the sensitivity of 60.0%, specificity of 60.0%, and

accuracy of 60.0%. Finally, NC has significant memory concern as other classes have been used for classification, and the result of their work showed sensitivity, specificity, and accuracy values of 60.0%, 45.7%, and 52.9%.

The use of ML for improving AD detection in the analysis of digital biomarkers within PET imaging has emerged recently. Islam et al. [102] have proposed a 3DCNN for AD diagnosis using only PET scans. They have reached the classification accuracy of 88.76% for NC versus AD categories. Their experiment also developed a regular CNN model using axial, coronal, and sagittal segments from each subject's brain from PET data; in the end, they have achieved 71.45% accuracy for NC versus AD classification.

In another research by Vasan et al. [103], F-FDG-PET brain images from the ADNI repository have been used to train the models for testing the result of the training model; they have used an independent test set from individual patients. They have used a branch of CNN which is called InceptionV3 [104]. Using this model, they reached AUC to predict AD, MCI, and NC, of 92%, 0.63%, and 0.73%, respectively. The AUROC for this classification task was 0.98, 0.52, and 0.84 to predict AD, MCI, and NC, respectively. The reported sensitivity consists of 100%, 43%, and 35% for AD, MCI, and NC prediction, respectively. As it is shown, the best result has been achieved for AD classification.

Lu et al. [105] used Fluoro-deoxy-glucose positron emission tomography (FDG-PET) to detect the brain's metabolic activity in the different subjects as data set. They proposed a novel DL method at the time for analyzing the FDG-PET. They used this information to classify MCI subjects with symptomatic AD and distinguish them from other subjects with MCI stages. The result of their work shows 82.51% accuracy of classification just using measures from a single modality. Because of the similarity between these two stages, their work can distinguish between two similar stages rather than AD versus NC.

Adeli et al. [106] proposed a semisupervised algorithm for first dealing with little labeled data and the vast majority of unable data. The discriminative classification method based on the least-squares formulation of linear discriminant analysis has been used in work. The use of a linear discriminant model has helped them to deal with noisy and unwanted data. They have surveyed Parkinson's and AD diseases as neurodegenerative diseases. They have used their framework to create an application for neurodegenerative disease diagnosis. They have reached 92.1% accuracy for AD versus NC classifications. In another work by Gamberger et al. [105], they have analyzed 5-year longitudinal outcomes and biomarker data from 562 subjects with MCI from ADNI. The mentioned algorithm identified homogenous clusters of MCI subjects with evidently diverse predictive cognitive courses. In the end, they have reported high sensitivity and specificity for the classification of AD versus NC.

In another research by Liu et al. [107], they used FDG-PET data as input for training. They have used DL models with the combination of GRU and convolutional layers. Their proposed model has abled them to use intrasegment and intersegment to classify different stages of AD disease. They have used the different frames of video and then

developed a CNN model on these images. In the structure of the proposed model, the convolutions layer has been used only to capture the valuable features from input image data. After reaching good data, they used GRU and RNN to learn and integrate the intersegment features for classification tasks. They reached 95.3% for AD versus NC classification and 83.9% for MCI versus NC classification based on AUC metrics. Their proposed method reached 91.2% accuracy, 91.4% sensitivity, and 91.0 specificity for AD versus NC condition, respectively.

5.3. Combination of Data. Combining PET and different branches of MRI data with the aim of reaching a diverse data set with complementary effects has gained attention. Gupta et al. [35] have used a combination of four different biomarkers: FDG-PET, sMRI, the level of protein in cerebrospinal fluid (CSF) data, and apolipoprotein (APOE) genotype. They have used data set from ADNI as their baseline data set. In total, they have surveyed 158 patients whose all of the mentioned input data are available for each of these subjects. In their study, patients were divided into 38 subjects of AD, 82 subjects of MCI groups, and the remaining 38 subjects in the NC group. With these categories of data set, their data set was imbalanced. They used a kernel-based multiclass SVM classifier with a grid-search method and truncated PCA to determine the best feature to use for training. They have reached AUROC 98.33%, 93.59%, 96.83%, 94.64%, 96.43%, and 95.24% for AD versus NC, MCIs versus MCIc, AD versus MCIs, AD versus MCIc, NC versus MCIc, and NC versus MCIs classification. The accuracy, sensitivity, and F1-score of AD and NC classification were 98.42%, 100%, and 98.42%. The result of their work has shown some state-of-the-art results for AD classifications. The result of their work has shown a significant rise in accuracy in different stages from previous works.

Youssofzadeh et al. [34] used MRI and PET in combination due to their complementary nature as the input data set. They have used a multimodal imaging ML to enhance AD classification performance metrics. They have used 58 AD subjects, 108 MCI subjects, and 120 NC subjects from the Australian imaging, biomarkers, and lifestyle data set. For classification of AD versus NC, MCI versus NC, and AD versus NC, they reached 95.7%, 95.8%, and 95.1% accuracy, respectively. Also, they have found multikernel learning regression analysis for excellent predictions of diagnosis of AD in subjected samples with a relation factor of 0.86. In addition, they have reached significant correlations between developed methods and delayed memory recall scores with a relation factor of 0.62.

In another research, Li et al. [108] used whole-brain images as input and designed a disease-image-specific neural network for the classification task of AD subjects. They have used MRI and PET scans as input data for classification tasks. Also, they have used feature-consistent GANs to produce some images for better results of classification. Using this branch of GAN, they have encouraged the proposed model to use these produced images and authentic images as a consistent data set for final prediction. Their work

conducted a state-of-the-art performance in both AD identification and MCI conversion prediction tasks at times. They have reached 34.18% in terms of the PR-AUC score.

In another study, Dukart et al. [109] used MRI and fluorodeoxyglucose FDG-PET to improve the detection accuracy of differentiation subjects with AD complication and frontotemporal lobar neuron connection failure. They have used an SVM classifier for this task. SVM classification has used combined information from different ROI of each subject from FDG-PET and MRI based on comprehensive quantitative meta-analyses. For the ADNI data set, accuracy rates of 88% have been achieved. In another work by Triebkorn et al. [110], they have used an amyloid-beta marker from each PET data from each patient. Also, they have used MRI, specifically amyloid-beta binding tracer PET, and tau protein (Tau) binding PET from 33 participants of ADNI3. Their work aims to classify AD MCI and NC using SVM and Random Forest classifiers together. They have reached 90.5% accuracy for NC versus AD and MCI classification. For AD versus NC and MCI classification, they have reached 78.3% accuracy. The sensitivity of this classifier method for NC versus AD and MCI classification was 90.5%. They also have reached AD versus HC and MCI classification of 78.3%.

In research by Kim and Lee [111], they proposed an autoencoder and sparse ELM to classify AD versus NC and MCI. They have used MRI, PET, and CSF pictures from 93 individual subjects. They have extracted volume and means ROIs as input features. At last, they have used a stacked sparse ELM autoencoder for the classification task. The use of an autoencoder for changing the input space into smaller later space has been done. For evaluating the proposed model, they have used 10-fold cross-validation. The classification result has shown more than 96% and 86.44% accuracy for classifying AD versus NC and MCI versus NC subjects.

5.4. Comparison between Different Modalities. In this section, we compared different neuroimaging modalities according to their performance. Comparison has been made based on different metrics such as accuracy, precision, recall, F1-score, and area under the receiver AROC. Accuracy is given in the following:

$$\text{sensitivity} = \left[\frac{TP}{TP + FN}\right] \times 100,$$

$$\text{specificity} = \left[\frac{TN}{TN + FP}\right] \times 100,$$

$$\text{positive predictive value (PPV)} = \left[\frac{TP}{TP + FP}\right] \times 100,$$

$$\text{negative predictive value (NPV)} = \left[\frac{TN}{TN + FN}\right] \times 100,$$

$$\text{accuracy (ACC)} = \left[\frac{TP + TN}{TP + TN + FP + FN}\right] \times 100.$$

(1)

TABLE 1: Summary of methods and neuroimaging techniques for binary classification of AD versus NC results.

Num.	Reference	Year	Method data	Neuroimaging	Accuracy	Sensitivity	AUROC
1	Duc et al. [7]	2020	3DCNN	rs-fMRI	85.27%	—	—
2	Ramzan et al. [90]	2020	CNN (ResNet-18)	fMRI	97.88%	—	—
3	Altinkaya et al. [92]	2020	CNN	fMRI	99.9 (on 303 samples)	—	—
4	Korolev et al. [93]	2017	CNN (ResNet)	MRI	80%	—	0.87
5	Li et al. [101]	2019	CNN	MRI	—	—	0.82
6	Yang et al. [95]	2021	GANs	MRI	99%	—	—
7	Dua et al. [96]	2020	CNN-LSTM (RNN)	MRI	92.22%	91.92%	
8	Zeng et al. [88]	2018	CNN (PSO)	MRI	76.85%	—	—
9	Kruthika and Maheshappa [76]	2019	3DCNN (capsule net)	sMRI	92.98%	—	0.98
10	Sarraf et al. [112]	2019	CNN (MCADNNet)	fMRI + MRI	99%	95%	—
11	Ozsahin et al. [102]	2019	ANN	PET	87.9%	92.4%	—
12	Islam and Zhang [113]	2019	3DCNN	PET (FDG)	88.76%	—	—
13	Vasan et al. [103]	2020	CNN (Inceptionv3)	PET	92%	100%	0.98
14	Lu et al. [105]	2017	ANN	PET (FDG)	82.51%	—	—
15	Adeli et al. [106]	2018	RF	PET	92%	—	0.94
16	Liu et al. [110]	2021	CNN-RNN (GRU)	PET	91.2%	92.4%	0.94
17	Gupta et al. [34]	2017	SVM	PET + MRI	92%	98.42%	0.98
18	Youssofzadeh et al. [109]	2013	MKML	PET + MRI	95.7%	—	—
19	Li et al. [108]	2018	GANs	PET (FDG) + MRI	—	—	0.32
20	Dukart et al. [107]	2018	SVM	PET (FDG) + MRI	78.3%	90.5%	—
21	Kim and Lee [111]	2018	ANN (AE)	PET (FDG) + MRI	96%	—	—
22	Billones et al. [114]	2016	CNN (DemNET)	MRI	98.33%	98.99%	—
23	Talo et al. [115]	2019	CNN (DeepResNet-50)	MRI	95.23%	97.16%	—
24	Yu et al. [73]	2020	THS-GAN	MRI	95.92%	—	—
25	Moradi et al. [116]	2015	Semisupervised model	MRI	87%	82%	
26	Kloppel et al. [86]	2004	SVM	MRI	88%	—	—
27	Montagne et al. [87]	2013	SVM	PET	82%	81%	—
28	Lama et al. [89]	2017	ELM	Smri	83.38%	93.01%	0.85
29	Lin et al. [117]	2021	GANs + 3D VGG	PET (FDG) + MRI	74.1%	75.00%	0.92
30	Zhou et al. [118]	2021	GANs + DNN	MRI	88.73%	63.09%	0.932
31	Baydargil et al. [119]	2021	GANs	PET	96.03%	—	0.7521
32	Venugopalan et al. [120]	2021	DNN + RF	PET + MRI	89%	96%	—
33	Zhang et al. [121]	2021	3DCNN	MRI	97.35%	97.10%	0.9970
34	Mehmood et al. [122]	2021	CNN	MRI	98.73%	—	—
35	Raju et al. [123]	2021	CNN (VGG16 based model)	MRI	99.2%	98.5%	—
36	Subramoniam [124]	2021	CNN (ResNet 101)	MRI	99.71%	0.99	1
37	Lella et al. [125]	2021	ELM	MRI	89.25%	0.78	—
38	Acharya et al. [126]	2019	KNN	MRI	94.54%	0.96	—
39	Amini et al. [99]	2021	SVM + PCA	fMRI	85.8%	0.95	0.92
40	Amini et al. [99]	2021	KNN + PCA	fMRI	77.5%	0.98	0.93
41	Amini et al. [99]	2021	CNN	fMRI	96.7%	0.98	1

Based on mentioned metrics, an evaluation of binary and multiclass classification can be done. As shown in Table 1, different methods used MRI as input data. Most of them have reached a good accuracy too. Using PET as data set for the sake of AD detection only based on accuracy is lower than MRI as input. Combined modalities have shown better performance based on accuracy than single data. The best algorithms for prediction consist of a different branch of CNN.

6. Results

Alzheimer's disease is the main cause of memory loss and dementia in people older than 65. AD consists of a gradually progressive neurodegenerative disease process consisting of gradual synaptic integrity and loss of cognitive functions. For detecting AD, a marker as a sign of progressive neurodegenerative is needed. This marker has a more critical role for AD detection than other diseases because beta-
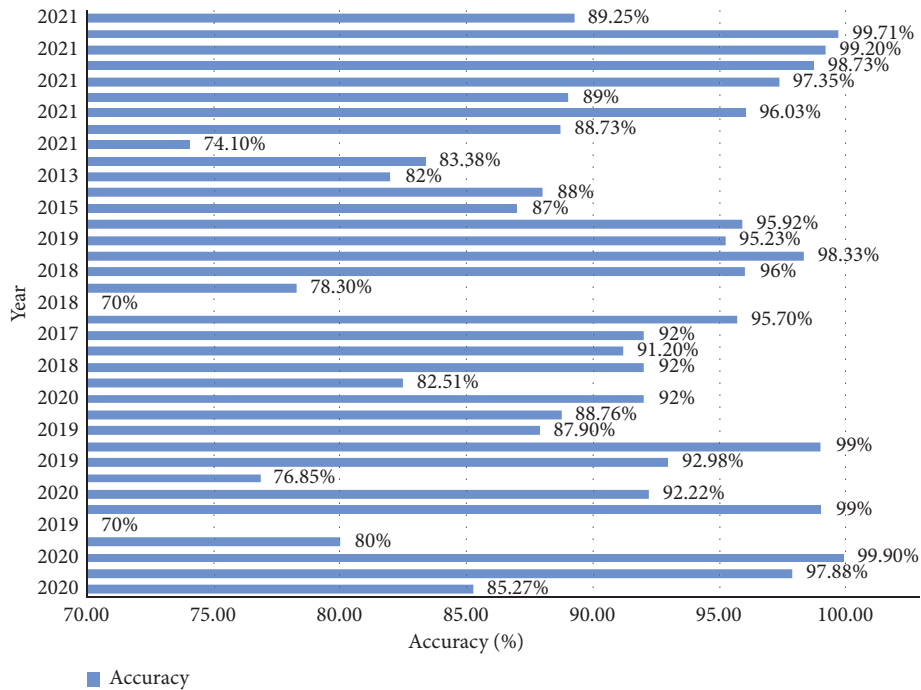
FIGURE 14: Different accuracy based on different years.

amyloid can form a complex structure with some metal ions. For assessing this marker, some neuroimaging techniques like fMRI and sMRI, PET, PET (FDG), and combination of these data have been used (see Table 1).

AI and its branches have helped with AD detection in patients with various methods. Some of these algorithms consist of ANN, CNN, and GANs as supervised and semi-supervised algorithms. Other famous ML algorithms such as KNN, SVM, and RF have been used too. This review article discussed different neuroimaging techniques and has shown some effects of some of the DL methods on this input data for AD detection. In the end, we have to summarize these different techniques and compare them based on accuracy, sensitivity, and AUROC. The results showed that combined neuroimaging techniques are a newly open field, and DL methods for detecting AD on average are of high accuracy. For people older than 65, if we want a model to be sensitive and its prediction considers low false-negative outcomes, MRI and CNN base methods have shown better results. In full use of the combination, methods have much more unspecific areas for research. As shown in Table 1 more than half of the reviewed article was published in 2020 and 2021. This year's focus on GANs and using VGG and resent structure with combination data from different repositories have gained more attention. Some articles have reached 99% accuracy with old popular CNN for AD versus NC classification task only. With proceeding years, using different methods for reaching a good accuracy has paid off, and most of the articles have reached a state-of-the-art accuracy. The result of accuracy versus years is shown in Figure 14. Time-series signals related to specific parts of the brain as input data and use of CNN or ML models for AD classification are another center of attention for researchers [100, 127]. As

discussed, we focused on AD versus NC classification improvements as the main priority. The reputation of authors or the number of citations for each research has not been considered in this article.

7. Conclusion and Discussion

Different ML techniques for AD classification and detection have been reviewed in this article. For gathering different images, data time consumption and the financial issue should be considered. Also, the use of combined data has shown promising performance. So, this is a trade-off that should solve this. Most of the research has worked on the ADNI data set due to its convenient access and variety of stages of the data set. AD versus NC and MCI classification has been surveyed in most of the researches. The use of different data set will add more controversy and variety to the data set. This variety is an essential part of each data set. So, for better and more comprehensive results use of a different variety of data must be considered. Also, for better early AD detection, different AD stages as a data set should be considered.

Data Availability

Data used in this paper's preparation were obtained from the ADNI database (http://adni.loni.usc.edu/).

Disclosure

The funding sources have no involvement support in the study design, collection, analysis, or interpretation of data,

writing of the manuscript, or in the decision to submit the manuscript for publication.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] K. Blennow, M. J. de Leon, and H. Zetterberg, "Alzheimer's disease," *The Lancet*, vol. 368, pp. 387–403, 2006.

[2] J. A. Weber, D. H. Baxter, S. Zhang et al., "The microRNA spectrum in 12 body fluids," *Clinical Chemistry*, vol. 56, no. 11, pp. 1733–1741, 2010.

[3] M. Brzyska, A. Bacia, and D. Elbaum, "Oxidative and hydrolytic properties of β-amyloid," *European Journal of Biochemistry*, vol. 268, no. 12, pp. 3443–3454, 2001.

[4] Y. Y. Syed and E. Deeks, "[18F]Florbetaben: a review in β-amyloid PET imaging in cognitive impairment," *CNS Drugs*, vol. 29, no. 7, pp. 605–613, 2015.

[5] M. E. Schmidt, L. Janssens, D. Moechars et al., "Clinical evaluation of [18F] JNJ-64326067, a novel candidate PET tracer for the detection of tau pathology in Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 47, no. 13, pp. 3176–3185, 2020.

[6] T. A. Pascoal, S. Mathotaarachchi, S. Mathotaarachchi et al., "Amyloid and tau signatures of brain metabolic decline in preclinical Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 45, no. 6, pp. 1021–1030, 2018.

[7] N. T. Duc, S. Ryu, M. N. I. Qureshi, M. Choi, K. H. Lee, and B. Lee, "3D-deep learning based automatic diagnosis of Alzheimer's disease with joint MMSE prediction using resting-state fMRI," *Neuroinformatics*, vol. 18, no. 1, pp. 71–86, 2020.

[8] C. Y. Sander, "Simultaneous PET/fMRI for imaging neuroreceptor dynamics," Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2014.

[9] H. F. Wehrl, M. Hossain, K. Lankes et al., "Simultaneous PET-MRI reveals brain function in activated and resting state on metabolic, hemodynamic and multiple temporal scales," *Nature Medicine*, vol. 19, no. 9, pp. 1184–1189, 2013.

[10] D. P. Veitch, M. W. Weiner, P. S. Aisen et al., "Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's disease neuroimaging initiative," *Alzheimer's Dementia*, vol. 15, no. 1, pp. 106–152, 2019.

[11] E. E. Brown, S. Kumar, T. K. Rajji, B. G. Pollock, and B. H. Mulsant, "Anticipating and mitigating the impact of the COVID-19 pandemic on Alzheimer's disease and related dementias," *The American Journal of Geriatric Psychiatry*, vol. 28, no. 7, Article ID 712721, 2020.

[12] L. H. Nicholas, K. M. Langa, J. P. Bynum, and J. W. Hsu, "Financial presentationof Alzheimer disease and related dementias," *JAMA Internal Medicine*, vol. 181, no. 2, Article ID 220227, 2021.

[13] N. Wang, M. Chen, and K. P. Subbalakshmi, "Explainable CNN-attention networks (C-attention network) for automated detection of Alzheimer's disease," 2020, https://arxiv.org/abs/2006.14135.

[14] G. Folego, M. Weiler, R. F. Casseb, R. Pires, and A. Rocha, "Alzheimer's disease detection through whole-brain 3D-CNN MRI," *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020.

[15] J. Bagherzadeh and H. Asil, "A review of various semi-supervised learning modelswith a deep learning and memory approach," *Iran Journal of Computer Science*, vol. 2, no. 2, pp. 65–80, 2019.

[16] H.-m. Lv, D. Zhao, and X. b. Chi, "Deep learning for early diagnosis of Alzheimer's disease based on intensive Alex-Net," *Journal of Computational Science*, vol. 44, pp. 50–60, 2017.

[17] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 173–183, 2017.

[18] K. Trojachanec, I. Kitanovski, I. Dimitrovski, and S. Loshkovska, "Longitudinal brain MRI retrieval for Alzheimer's disease using different temporal information," *IEEE Access*, vol. 6, pp. 9703–9712, 2018.

[19] S. Kumar, C. Dabas, and S. Godara, "Classification of brain MRI tumor images: a hybrid approach," *Procedia Computer Science*, vol. 122, pp. 510–517, 2017.

[20] J. Sun, B. Wang, Y. Niu et al., "Complexity analysis of EEG, MEG, and fMRI in mild cognitive impairment and Alzheimer's disease: a review," *Entropy*, vol. 22, no. 2, p. 239, 2020.

[21] W. Li and P. C. van Zijl, "Quantitative theory for the transverse relaxation time of blood water," *NMR in Biomedicine*, vol. 33, no. 5, p. 427, 2020.

[22] K. Sakurai, T. Shintani, N. Jomura, T. Matsuda, A. Sumiyoshi, and T. Hisatsune, "Hyper bold activation in dorsal raphe nucleus of APP/PS1 Alzheimer's disease mouse during rewARD -oriented drinking test under thirsty conditions," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.

[23] G. Castellazzi, M. G. Cuzzoni, M. Cotta Ramusino et al., "A machine learning approach for the differential diagnosis of alzheimer and vascular dementia fed by MRI selected features," *Frontiers in Neuroinformatics*, vol. 14, p. 25, 2020.

[24] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nature Reviews Neuroscience*, vol. 8, no. 9, pp. 700–711, 2007.

[25] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995.

[26] T. Wu, L. Wang, Y. Chen, C. Zhao, K. Li, and P. Chan, "Changes of functional connectivity of the motor network in the resting state in Parkinson's disease," *Neuroscience Letters*, vol. 460, no. 1, pp. 6–10, 2009.

[27] T. Jones and D. W. Townsend, "History and future technical innovation in positronemission tomography," *Journal of Medical Imaging*, vol. 4, no. 1, pp. 110–113, 2017.

[28] G. Muehllehner and J. S. Karp, "Positron emission tomography," *Physics in Medicine Biology*, vol. 51, no. 13, Article ID R117, 2006.

[29] A. Gustavsson, M. Svensson, F. Jacobi et al., "Cost of disorders of the brain in europe 2010," *European Neuropsychopharmacology*, vol. 21, no. 10, pp. 718–779, 2011.

[30] R. E. Carson, "Tracer kinetic modeling in PET," in *Positron Emission Tomography*, pp. 127–159, 2005.

[31] A. Drzezga, D. Altomare, C. Festari et al., "Diagnostic utility of 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) in asymptomatic subjects at increased risk for Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 45, no. 9, pp. 1487–1496, 2018.

[32] M. Colom, B. Vidal, and L. Zimmer, "Is there a role for GPCR agonist radiotracers in PET neuroimaging?" *Frontiers in Molecular Neuroscience*, vol. 12, p. 255, 2019.

[33] A. J. Reader, S. Ally, F. Bakatselos et al., "One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays," *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 693–699, 2002.

[34] V. Youssofzadeh, B. McGuinness, L. P. Maguire, and K. Wong-Lin, "Multi-kernellearning with dartel improves combined MRI-PET classification of Alzheimer's disease in AIBL data: group and individual analyses," *Frontiers in Human Neuroscience*, vol. 11, p. 380, 2017.

[35] Y. Gupta, R. K. Lama, and G. R. Kwon, "Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers," *Frontiers in Computational Neuroscience*, vol. 13, p. 72, 2019.

[36] Y. Fan, S. M. Resnick, X. Wu, and C. Davatzikos, "Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study," *Neuroimage*, vol. 41, no. 2, pp. 277–285, 2008.

[37] D. Zhang and D. Shen, "Alzheimer's disease neuroimaging initiative, multi-modalmulti-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease," *NeuroImage*, vol. 59, no. 2, pp. 895–907, 2012.

[38] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Alzheimer's disease neuroimaging initiative, multimodal classification of Alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 55, no. 3, pp. 856–867, 2011.

[39] K. Ritter, J. Schumacher, M. Weygandt, R. Buchert, C. Allefeld, and J. D. Haynes, "Alzheimer's disease neuroimaging initiative, multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers," *Alzheimer's Dementia: Diagnosis, Assessment Disease Monitoring*, vol. 1, no. 2, pp. 206–215, 2015.

[40] A. Drzezga, H. Barthel, S. Minoshima, and O. Sabri, "Potential clinical applications of PET/MR imaging in neurodegenerative diseases," *Journal of Nuclear Medicine*, vol. 55, no. 2, pp. 47S–55S, 2014.

[41] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, p. 95, 2007.

[42] J. Young, M. Modat, M. J. Cardoso, A. Mendelson, D. Cash, and S. Ourselin, "Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment," *NeuroImage: Clinical*, vol. 2, pp. 735–745, 2013.

[43] Y. Asim, B. Raza, A. K. Malik, S. Rathore, L. Hussain, and M. A. Iftikhar, "A multimodal, multi-atlas-based approach for Alzheimer detection via machine learning," *International Journal of Imaging Systems and Technology*, vol. 28, no. 2, pp. 113–123, 2018.

[44] H. P. W. Schlemmer, B. J. Pichler, M. Schmand et al., "Simultaneous MR/PET imaging of the human brain: feasibility study," *Radiology*, vol. 248, no. 3, pp. 1028–1035, 2008.

[45] N. J. Tustison, B. B. Avants, P. A. Cook et al., "N4ITK: improved N3 bias correction," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[46] T. Hara, K. Inagaki, N. Kosaka, and T. Morita, "Sensitive detection of mediastinal lymph node metastasis of lung cancer with 11C-choline PET," *Journal of Nuclear Medicine*, vol. 41, no. 9, pp. 1507–1513, 2000.

[47] L. Fan, H. Li, J. Zhuo et al., "The human brain netome atlas: a new brain atlas based on connectional architecture," *Cerebral Cortex*, vol. 26, no. 8, pp. 3508–3526, 2016.

[48] T.-E. Kam, H. Zhang, Z. Jiao, and D. Shen, "Deep learning of static and dynamic brain functional networks for early MCI detection," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 478–487, 2019.

[49] L. Kang, J. Jiang, J. Huang, and T. Zhang, "Identifying early mild cognitive impairment by multi-modality MRI-based deep learning," *Frontiers in Aging Neuroscience*, vol. 12, p. 206, 2020.

[50] T. E. Kam, H. Zhang, and D. Shen, "A novel deep learning framework on brain functional networks for early MCI diagnosis," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 293–301, Granada, Spain, September 2018.

[51] S. H. Hojjati, A. Ebrahimzadeh, A. Khazaee, and A. Babajani-Feremi, "Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI," *Computers in Biology and Medicine*, vol. 102, pp. 30–39, 2018.

[52] Y. Du and Y. Fan, "Group information guided ICA for fMRI data analysis," *NeuroImage*, vol. 69, pp. 157–197, 2013.

[53] W. E. Klunk, R. A. Koeppe, J. C. Price et al., "The centiloid project: standardizing quantitative amyloid plaque estimation by PET," *Alzheimer's Dementia*, vol. 11, no. 1, pp. 1–15, 2015.

[54] R. Haase, L. A. Royer, P. Steinbach et al., "CLIJ: GPU-accelerated image processing for everyone," *Nature Methods*, vol. 17, no. 1, pp. 5-6, 2020.

[55] S. I. Dimitriadis, D. Liparas, and M. N. Tsolaki, "Random forest feature selection, fusion and ensemble strategy: combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and Alzheimer's disease patients: from the Alzheimer's disease neuroimaging initiative (ADNI) database," *Journal of Neuroscience Methods*, vol. 302, pp. 14–23, 2018.

[56] Y. Shi, Y. Gao, S. Liao, D. Zhang, Y. Gao, and D. Shen, "A learning-based CT prostatesegmentation method via joint transductive feature selection and regression," *Neurocomputing*, vol. 173, pp. 317–331, 2016.

[57] M. Guinin, S. Ruan, B. Dubray, L. Massoptier, and I. Gardin, "Notice of removal: feature selection and patch-based segmentation in MRI for prostate radiotherapy," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 2663–2667, Phoenix, AZ, USA, August 2016.

[58] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: principal component analysis," *Nature Methods*, vol. 14, no. 7, pp. 641-642, 2017.

[59] S. Graham, C. Depp, E. E. Lee et al., "Artificial intelligence for mental health and mental illnesses: an overview," *Current Psychiatry Reports*, vol. 21, no. 11, pp. 1–18, 2019.

[60] C. J. Brown and G. Hamarneh, "Machine learning on human connectome data fromMRI," 2016, https://arxiv.org/abs/1611.08699.

[61] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–16, 2019.

[62] E. Barbará-Morales, J. Pérez-González, K. C. Rojas-Saavedra, and V. Medina-Bañuelos, "Evaluation of brain tortuosity measurement for the automatic multimodal classification of subjects with Alzheimer's disease," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 4041832, 11 pages, 2020.

[63] H. Huo, Z. Rong, O. Kononova et al., "Semi-supervised machine-learning classification of materials synthesis procedures," *Npj Computational Materials*, vol. 5, no. 1, pp. 1–7, 2019.

[64] J. F. Rodriguez-Nieva and M. Scheurer, "Identifying topological order through unsupervised machine learning," *Nature Physics*, vol. 15, no. 8, pp. 790–795, 2019.

[65] K. Shankar, S. K. Lakshmanaprabu, A. Khanna, S. Tanwar, J. J. Rodrigues, and N. R. Roy, "Alzheimer detection using group grey wolf optimization based features with convolutional classifier," *Computers Electrical Engineering*, vol. 77, pp. 230–243, 2019.

[66] K. R. Kruthika and H. D. Maheshappa, "Multistage classifier-based approach for Alzheimer's disease prediction and retrieval," *Informatics in Medicine Unlocked*, vol. 14, pp. 34–42, 2019.

[67] M. A. El-Yacoubi, S. Garcia-Salicetti, C. Kahindo, A. S. Rigaud, and V. Cristancho Lacroix, "From aging to early-stage Alzheimer's: uncovering handwriting multimodal behaviors by semi-supervised learning and sequential representation learning," *Pattern Recognition*, vol. 86, pp. 112–133, 2019.

[68] D. H. Ye, K. M. Pohl, and C. Davatzikos, "Semi-supervised pattern classification: application to structural MRI of Alzheimer's disease," in *Proceedings of the 2011 International Workshop on Pattern Recognition in NeuroImaging*, pp. 1–4, Seoul, Korea, May 2011.

[69] V. Sindhwani, P. Niyogi, M. Belkin, and S. Keerthi, "Linear manifold regularization for large scale semi-supervised learning," in *Proceedings of the 22nd ICML Workshop on Learning with Partially Classified Training Data*, vol. 28, Bonn, Germany, August 2005.

[70] J. M. Gorriz, J. Ramirez, J. Suckling et al., "A semi-supervised learning approach for model selection based on class-hypothesis testing," *Expert Systems with Applications*, vol. 90, pp. 40–49, 2017.

[71] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent progress on generative adversusrial networks (GANs): a survey," *IEEE Access*, vol. 7, pp. 36322–36333, 2019.

[72] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversusrialnetworks: algorithms, theory, and applications," 2020, https://arxiv.org/abs/2001.06937.

[73] W. Yu, B. Lei, M. K. Ng, A. C. Cheung, Y. Shen, and S. Wang, "TensorizingGAN with high-order pooling for Alzheimer's disease assessment," 2020, https://arxiv.org/abs/2008.00748.

[74] H. Bhavsar and M. H. Panchal, "A review on support vector machine for data classification," *International Journal of Advanced Research in Computer Engineering Technology*, vol. 1, p. 10, 2012.

[75] F. Li, L. Tran, K. H. Thung, S. Ji, D. Shen, and J. Li, "Robust deep learning for improved classification of AD/MCI patients," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, pp. 240–247, Boston, MA, USA, September 2014.

[76] K. R. Kruthika and H. D. Maheshappa, "CBIR system using capsule networks and 3D CNN for Alzheimer's disease diagnosis," *Informatics in Medicine Unlocked*, vol. 14, pp. 59–68, 2019.

[77] A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, "3DCNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies," 2018, https://arxiv.org/abs/1801.05968.

[78] Y. D. Li, Z. B. Hao, and H. Lei, "Survey of convolutional neural network," *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508–2515, 2016.

[79] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3693–3702, Honolulu, HI, USA, July 2017.

[80] D. Yu, H. Wang, P. Chen, and Z. Wei, "Mixed pooling for convolutional neuralnetworks," in *Proceedings of the International Conference on Rough Sets and Knowledge Technology*, pp. 364–375, Shanghai, China, October 2014.

[81] S. Sarraf and G. Tofighi, "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, Article ID 070441, 2016.

[82] S. Spasov, L. Passamonti, A. Duggento, P. Lio, and N. Toschi, "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease," *Neuroimage*, vol. 189, pp. 276–287, 2019.

[83] J. Dakka, P. Bashivan, M. Gheiratmand, I. Rish, S. Jha, and R. Greiner, "Learning neural markers of schizophrenia disorder using recurrent neural networks," 2017, https://arxiv.org/abs/1712.00512.

[84] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTMcells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[85] S. Karlekar, T. Niu, and M. Bansal, "Detecting linguistic characteristics of Alzheimer's dementia by interpreting neural models," 2018, https://arxiv.org/abs/1804.06440.

[86] S. Kloppel, C. M. Stonnington, C. Chu et al., "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp. 681–689, 2008.

[87] C. Montagne, A. Kodewitz, V. Vigneron, V. Giraud, and S. Lelandais, "3D local binary pattern for PET image classification by SVM, Application to early Alzheimer disease diagnosis," in *Proceedings of the 6th International Conference on Bio-Inspired Systems and Signal Processing (BIOSIGNALS 2013)*, Barcelona, Spain, February 2013.

[88] N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, and Y. Li, "A new switching delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 320, pp. 195–202, 2018.

[89] R. K. Lama, J. Gwak, J. S. Park, and S. W. Lee, "Diagnosis of Alzheimer's diseasebased on structural MRI images using a regularized extreme learning machine and PCA features," *Journal of Healthcare Engineering*, vol. 2017, Article ID 5485080, 2017.

[90] F. Ramzan, M. U. G. Khan, A. Rehmat et al., "A deep learning approach for automated diagnosis and multiclass classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks," *Journal of Medical Systems*, vol. 44, no. 2, pp. 1–16, 2020.

[91] M. Pak and S. Kim, "A review of deep learning in image recognition," in *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, pp. 1–3, IEEE, Kuta Bali, Indonesia, August 2017.

[92] E. Altinkaya, K. Polat, and B. Barakli, "Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: a comprehensive review," *Journal of the Institute of Electronics and Computer*, vol. 1, no. 1, pp. 39–53, 2020.

[93] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3D

brain MRI classification," in *Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging*, pp. 835–838, Melbourne, Australia, April 2017.

[94] H. Li, M. Habes, D. A. Wolk, and Y. Fan, "A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data," *Alzheimer's Dementia*, vol. 15, no. 8, pp. 1059–1070, 2019.

[95] Z. Yang, I. M. Nasrallah, H. Shou et al., "Disentangling brain heterogeneity via semi-supervised deep-learning and MRI: dimensional representations of Alzheimer's disease," 2021, https://arxiv.org/abs/2102.12582.

[96] M. Dua, D. Makhija, P. Y. L. Manasa, and P. Mishra, "A CNN–RNN–LSTM based amalgamation for Alzheimer's disease detection," *Journal of Medical and Biological Engineering*, vol. 40, no. 5, pp. 688–706, 2020.

[97] C. Feng, A. Elazab, P. Yang et al., "Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBi-LSTM," *IEEE Access*, vol. 7, pp. 63605–63618, 2019.

[98] A. Sarraf, M. Azhdari, and S. Sarraf, "A comprehensive review of deep learning architectures for computer vision applications," *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 77, no. 1, pp. 1–29, 2021.

[99] M. Amini, M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of Alzheimer's disease severity with fMRI images using robust multitask feature extraction method and convolutional neural network (CNN)," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 5514839, 15 pages, 2021.

[100] M. Amini, M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of Alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using EEG signal," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 5511922, 17 pages, 2021.

[101] I. Ozsahin, B. Sekeroglu, and G. S. Mok, "The use of back propagation neural networks and 18F-florbetapir PET for early detection of Alzheimer's disease using Alzheimer's disease neuroimaging initiative database," *PLoS One*, vol. 14, no. 12, Article ID e0226577, 2019.

[102] J. Islam and Y. Zhang, "Understanding 3D CNN behavior for Alzheimer's disease diagnosis from brain PET scan," 2019, https://arxiv.org/abs/1912.04563.

[103] D. Vasan, M. Alazab, S. Wassan, H. Naeem, B. Safaei, and Q. Zheng, "IMCFN: image-based malware classification using fine-tuned convolutional neural network architecture," *Computer Networks*, vol. 171, Article ID 107138, 2020.

[104] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease," *Medical Image Analysis*, vol. 46, pp. 26–34, 2018.

[105] D. Gamberger, N. Lavrac, S. Srivatsa, R. E. Tanzi, and P. M. Doraiswamy, "Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease," *Scientific Reports*, vol. 7, no. 1, pp. 1–12, 2017.

[106] E. Adeli, K. H. Thung, L. An et al., "Semi-supervised discriminative classification robust to sample-outliers and feature noises," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 515–522, 2018.

[107] M. Liu, D. Cheng, and W. Yan, "Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images," *Frontiers in Neuroinformatics*, vol. 12, p. 35, 2018.

[108] W. Li, Y. Wang, Y. Cai, C. Arnold, E. Zhao, and Y. Yuan, "Semi-supervised rare disease detection using generative adversusrial network," 2018, https://arxiv.org/abs/1812.00547.

[109] J. Dukart, K. Mueller, H. Barthel, A. Villringer, O. Sabri, and M. L. Schroeter, "Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI," *Psychiatry Research: Neuroimaging*, vol. 212, no. 3, pp. 230–236, 2013.

[110] P. Triebkorn, L. Stefanovski, K. Dhindsa et al., "Multi-scale brain simulation with integrated positron emission tomography yields hidden local field potential activity that augments machine-learning classification of Alzheimer's disease," *BioRxiv*, 2021.

[111] J. Kim and B. Lee, "Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine," *Human Brain Mapping*, vol. 39, no. 9, pp. 3728–3741, 2018.

[112] S. Sarraf, D. D. Desouza, J. A. Anderson, and C. Saverino, "MCADNNet: recognizing stages of cognitive impairment through efficient convolutional fMRI and MRI neural network topology models," *IEEE Access*, vol. 7, pp. 155584–155600, 2019.

[113] Y. Ding, J. H. Sohn, M. G. Kawczynski et al., "A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain," *Radiology*, vol. 290, no. 2, pp. 456–464, 2019.

[114] C. D. Billones, O. J. L. D. Demetria, D. E. D. Hostallero, and P. C. Naval, "DemNet: a convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment," in *2016 IEEE Region 10 Conference*, pp. 3724–3727, Singapore, November 2016.

[115] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using MRI images," *Computerized Medical Imaging and Graphics*, vol. 78, Article ID 101673, 2019.

[116] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *Neuroimage*, vol. 104, pp. 398–412, 2015.

[117] W. Lin, W. Lin, G. Chen, H. Zhang, Q. Gao, and Y. Huang, "Bidirectional mapping of brain MRI and PET with 3D reversible GAN for the diagnosis of Alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, p. 357, 2021.

[118] X. Zhou, S. Qiu, P. S. Joshi et al., "Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversusrial learning," *Alzheimer's Research & Therapy*, vol. 13, no. 1, pp. 1–11, 2021.

[119] H. B. Baydargil, J. S. Park, and D. Y. Kang, "Anomaly analysis of Alzheimer's disease in PET images using an unsupervised adversusrial deep learning model," *Applied Sciences*, vol. 11, no. 5, Article ID 2187, 2021.

[120] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021.

[121] J. Zhang, B. Zheng, A. Gao, X. Feng, D. Liang, and X. Long, "A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification," *Magnetic Resonance Imaging*, vol. 78, pp. 119–126, 2021.

[122] A. Mehmood, S. Yang, Z. Feng et al., "A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images," *Neuroscience*, vol. 460, pp. 43–52, 2021.

[123] M. Raju, M. Thirupalani, S. Vidhyabharathi, and S. Thilagavathi, "Deep learning based multilevel classification of Alzheimer's disease using MRI scans," in *IOP Conference Series: Materials Science and Engineering*, vol. 1084, no. 1, Bristol, UK, IOP Publishing, Article ID 012017, 2021.

[124] M. Subramoniam, "Deep learning based prediction of Alzheimer's disease from magnetic resonance images," 2021, https://arxiv.org/abs/2101.04961.

[125] E. Lella, A. Pazienza, D. Lofù, R. Anglani, and F. Vitulano, "An ensemble learning approach based on diffusion tensor imaging measures for Alzheimer's disease classification," *Electronics*, vol. 10, no. 3, p. 249, 2021.

[126] U. R. Acharya, S. L. Fernandes, J. E. Wei Koh et al., "Automated detection of Alzheimer's disease using brain MRI images–a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, pp. 1–14, 2019.

[127] Y. Höller, K. H. Butz, A. C. Thomschewski et al., "Prediction of cognitive decline in temporal lobe epilepsy and mild cognitive impairment by EEG, MRI, and neuropsychology," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 8915961, 16 pages, 2020.