Early detection of Alzheimer's disease based on the state-of-the-art deep learning approach: a comprehensive survey

Doaa Ahmed Arafa[1] · Hossam El-Din Moustafa[2] · Amr M. T. Ali-Eldin[1] · Hesham A. Ali[1]

Abstract

Alzheimer's disease (AD) is a form of brain disorder that causes functions' loss in a person's daily activity. Due to the tremendous progress of Alzheimer's patients and the lack of accurate diagnostic tools, early detection and classification of Alzheimer's disease are open research areas. Accurate detection of Alzheimer's disease in an effective way is one of the many researchers' goals to limit or overcome the disease progression. The main objective of the current survey is to introduce a comprehensive evaluation and analysis of the most recent studies for AD early detection and classification under the state-of-the-art deep learning approach. The article provides a simplified explanation of the system stages such as imaging, preprocessing, learning, and classification. It addresses broad categories of structural, functional, and molecular imaging in AD. The included modalities are magnetic resonance imaging (MRI; both structural and functional) and positron emission tomography (PET; for assessment of both cerebral metabolism and amyloid). It reviews the process of pre-processing techniques to enhance the quality. Additionally, the most common deep learning techniques used in the classification process will be discussed. Although deep learning with preprocessing images has achieved high performance as compared to other techniques, there are some challenges. Moreover, it will also review some challenges in the classification and preprocessing image process over some articles what they introduce, and techniques used, and how they solved these problems.

Keywords Alzheimer's disease (AD) · Convolution neural network (CNN) · Deep Learning (DL) · Imaging pre-processing · Neuroimaging classification

✉ Doaa Ahmed Arafa
  doaaarafa@mans.edu.eg

Extended author information available on the last page of the article

1 Introduction

Alzheimer's disease (AD) is a neurological disease that affects the disorder in brain function and destroys brain cells slowly that leads to a loss in memory and instability in human life [122]. AD pathogenesis is thought to be caused by the overproduction of amyloid-β (Aβ) and hyperphosphorylation of tau protein. This results in the accumulation of Aβ plaques and tau neurofibrillary tangles, which disrupt the nucleocytoplasmic transport between neurons leading to cell death, which causes loss in memory and learning [112]. Physicians diagnose patients concerning many requirements where imaging scanning is an essential part. The common symptoms are (1) loss of motion function, (2) speaking difficulties, and (3) memory problems [12].

In our time with economic development and the advent of computer technology and medical information processing technologies, doctors require fast and accurate ways to diagnose and detect the disease aiming to help patients and save their lives. Compared with the traditional ways for diagnosing the disease. Patients pass through several stages to be diagnosed with the disease, but this can be late due to the diagnosis stages, and patients become in a late stage [109]. So, early diagnosis of AD is very important for patients, that help him in taking precaution and help clinicians to detect the risk of the progress of AD, it provides AD patients with knowledge of the seriousness and encourages them to take preventive steps, such as lifestyle changes and drugs [87].

Researchers want to find a simple and accurate approach to identify Alzheimer's disease before symptoms appear. So, Early detection of AD has discovered the symptoms before reaching to risk stage. AD has several stages, one of these stages that appear disease in the prodromal stage is MCI. MCI is a stage of memory loss or other cognitive abilities loss (such as language or visual/spatial perception) in people who can still do most of their daily tasks independently [39].

Recently, most researchers' interest in a search in this field to care about it to improve the quality of a patient's life and discover drugs by tracking the pathological processes related to several stages of AD [112]. Due to AD is develop progressive disease so it has several stages, cognitive normal (CN), Mild cognitive impairment (MCI), Late Mild cognitive impairment (LMCI), and Alzheimer's disease (AD). There are various technologies for neuroimaging that help researchers in classification, the common one that uses to imaging brain tissue in magnetic Resonance Imaging (MRI) [6]. Deep learning is the most common and best technique used for the diagnosis and classification of disease with a large number of input data [6].

Many surveys have been published recently reviewing histopathological imag#e analysis comprising its history, and detailed information of general artificial intelligence techniques [29, 52, 54, 55, 136, 139]; the main limitation is the lack of surveys of histopathological image analysis that focused on Alzheimer's disease [29, 52, 54, 63, 152]. Accordingly, we present more image analysis from an Alzheimer's disease point of view in this survey.

The main objective of the current survey is to provide a comprehensive overview of the state-of-the-art image analysis and artificial intelligence techniques, specifically for histopathology images in AD, and their challenges. This survey focuses on 159 state-of-the-art related studies, where 110 papers concentrate mainly on Alzheimer's disease. Figure 1 depicts the corresponding statistical distribution of the studies used in the current survey.
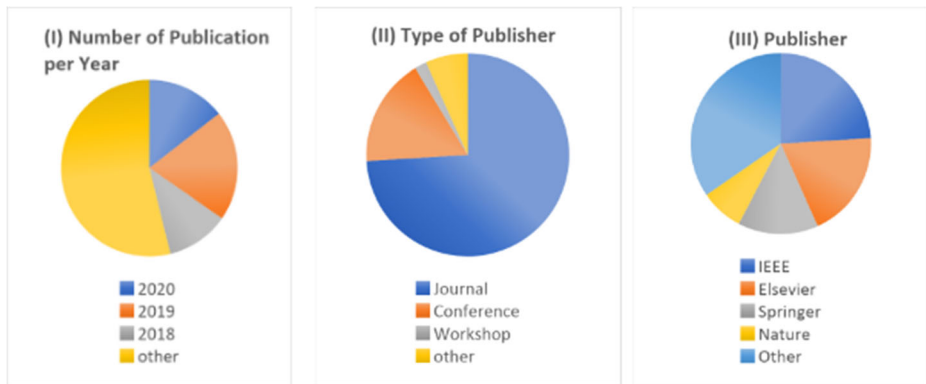
Fig. 1 Statistical Distribution of the Studies used in the current Survey. Left: (I) Number of Publications per Year, Middle: (II) Type of Publisher, and Right: (III) Publisher

1.1 Paper contributions

In summary, the current survey is to introduce a comprehensive evaluation and analysis of the most recent studies for AD early detection and classification under the state-of-the-art deep learning approach. Also, we present some preprocessing techniques that can enhance the quality of images and achieve the best performance of the classification process. Moreover, it will highlight different challenges throughout related studies and how they overcome them. In evaluating and analyzing the existing studies, several common trends and gaps have been identified.

The contributions of the current survey are summarized as follows:

- Introducing a process and stage of diagnosis of Alzheimer's disease from acquiring the data to how to classify disease.
- Presenting type and modalities of brain imaging such as MRI, PET, and CT and comparative advantage and disadvantage.
- Summarizing the most important basics and background related to the presented survey with a focus on deep learning as one of the most important recent trends to improve these systems.
- Categorizing the most important Research Challenges for each the most recent and significant.
- Presenting a comparison between various articles and present the contribution of each article, the most significant feature, and the advantage (and the disadvantage) of the solution that they used in solving a specific problem.
- Concluding the most open research points in this area.

1.2 Paper organization

The current survey is organized as follows: Section 2 presents different related works in the diagnosis of AD, Section 3 introduces an overview of ML and deep learning (DL), definitions, and challenges. Section 4 focuses on the diagnosis of Alzheimer's disease as an overview and highlights the various used methods. Section 5 explores the most important problems and

challenges associated with the early detection of AD and concludes remarks. We introduce some future possibilities in Section 6. We present our limitation in Section 7. Finally, the survey is concluded in Section 8.

2 Related work

Due to the importance of early diagnosis of AD, many researchers interest search in this field to solve the problem of AD. So, this section will present the most important research in this field. Table 1 introduces a comparison between different studies that diagnosis AD with different DL techniques. Jain et al. [66] proposed a transfer learning approach for classifying MRI images. They used PE SE CTL mathematical model to differentiate between 3 classes (AD, MCI, CN). Firstly, they collected data from ADNI datasets and make preprocessing data by using FreeSurfer (PE) to eliminate unnecessary information of MRI images. Preprocessing techniques that they used are 5 process motion correction, non-uniform intensity normalization, Talairach transform computation, intensity normalization, skull stripping. Then after preprocessing data, select the most important slices (SE) that have more information based on entropy. Lastly, they used the VGG16 pre-trained model and transfer learning to build a classification model (CTL). Their proposed technique achieves high accuracy of 95.3%, 99.14%, 99.3%, 99.22% for 3-way classification (AD vs MCI vs CN), AD vs CN, AD vs MCI, and CN vs MCI, respectively.

Ding et al. [35] introduced a CNN architecture by using an Inception v3 network trained on 90% of ADNI data and testing 10%. Fluorine 18 fluorodeoxyglucose PET images are processed by using the grid method, these images are acquired from the ADNI dataset. Otsu threshold was applied to detect brain voxel. Adam optimizer was used with a learning rate of 0.0001 and with batch size 8 for the training model. The model was trained by using 90% of the dataset (1921 images studies), this dataset includes 3 classes (AD, MCI, and no disease). The proposed architecture achieves 82% of specificity and 100% sensitivity.

In Chitradevi et al. [28], several optimization algorithms (Genetic Algorithm, particle Swarm Optimization Algorithm, Grey Wolf Optimization, and Cuckoo Search) were used to segment the brain into sub-region such as the hippocampus, white matter, and gray matter.

Table 1 A literature review of diagnosis Alzheimer Disease with Deep Learning Techniques

Reference	Dataset	Classifier	Image type	Size of image	Specificity	Sensitivity	accuracy
Jain et al. [66]	ADNI	2D-CNN	sMRI	224 *224 * 3	–	–	95.3%
Ding et al. [35]	ADNI	CNN (Inception v3)	fluorine 18 fluorodeoxy-glucose PET	512 * 512	82%	100%	–
Chitradevi et al. [28]	Chettinad Health City	CNN (AlexNet)	MRI	256 * 256	94	95	95%
Nawaz et al. [94]	OASIS	SVM	MRI	256 * 256	–	–	99.21%
Kundaram et al. [79]	ADNI	CNN	MRI	256 * 256	–	–	98.57%

They acquired images from Chettinad Health City which contain 200 images for AD and normal patients, images were processed with various methods such as skull stripping, enhance quality, and contrast enhancement. After segmentation, to get the performance of segmentation they validate the segmented region with a ground truth image. The validation measures Feature Similarity Index Metrics, Structure Similarity Index Metrics, dice similarity, Jaccard Index, Tanimoto, and volume similarity. To perform feature extraction and classification model CNN was used especially AlexNet. Grey Wolf Optimization shows the highest performance compared with other optimization techniques by achieving high accuracy with 95%.

Nawaz et al. [94] proposed three models and compared them to get which of them achieve high accuracy. In the first model, images are preprocessed and extracted the handcrafted features and make classification by using support vector machine, k-nearest neighbor, and Random Forest. Second model, training model on the preprocessed dataset from scratch by using CNN deep learning model. The third model, AlexNet is used to extract deep features, to determine the best classifier support vector machine, k-nearest neighbor, and Random Forest was fed with features. By comparing the three models, the deep features-based model achieved the best accuracy with a support vector machine classifier. By comparing result analysis support vector machine achieve the highest accuracy of 99.21% and the accuracy of k-nearest neighbor 57.32% and random forest achieve 93.97%.

Kundaram et al. [79] acquired data from the ADNI dataset and preprocessed images by rescaling it to 255. CNN models are used to trained and classify disease. Images were classified into 3 classes (AD, MCI, and NC), 9540 images are used for the training model. CNN model is consists of three convolutional layers, three max-pooling, four ReLU activation layers. They used different optimizers such as Adam, SGD, Adagrad, Nadam, Adadelta, Rmsprop. By comparing different optimizers with the proposed framework, Adagrad achieves the best accuracy with less loss. The proposed model achieved 98.57% accuracy on the ADNI dataset.

In Table 2 we compare the recent survey of AD with DL and our proposed survey. Description and limitation for each survey are presented to show the difference and similarities between other survey and our proposed survey.

3 Machine Learning and Deep Learning overview

Machine learning is a branch of artificial intelligence that has become precisely widespread, and valuable, in the last two decades. One definition of ML is that it is the semi-automated extraction of knowledge from data [15]. ML uses data to feed an algorithm that can understand the relationship between the input and the output. When the machine finishes learning, it can predict the value or the class of a new data point [64].

Deep learning is a subset of ML, as shown in Fig. 2. ML is an algorithm with the ability to learn without being explicitly programmed. Artificial intelligence is a technique of getting the machine to work and behave like humans. In DL, the learning phase is done through the artificial neural network [14]. An artificial neural network is an architecture where the layers are stacked on top of each other. There are different DL types such as convolutional neural networks (CNN), recurrent neural networks (RNN), and autoencoders [58].

Although the performance of ML models has progressively better in many functions, they still need guidance (e.g., human experts) to solve some problems. Developers should enhance the architecture or algorithm if incorrect predictions are returned. On the other hand, a DL

Table 2　Comparison between other surveys and our proposed survey

Reference	Description	Limitation
Zhang et al. [156]	Discussed techniques of DL like ANN, Stacked Auto-encoder, Deep Belief Networks, Deep Boltzmann Machine, CNN], RNN. Then, they discussed DL with brain disorder diseases, including AD, Parkinson's disease, Autism spectrum disorder, and Schizophrenia.	Preprocessing techniques and challenges are not discussed.
Zhao et al. [158]	Discussed most of the recent DL approaches such as CNN, AE, RNN. Then, they showed DL can be used in the analysis of brain MRIs and present the application of DL to Parkinson's disease, AD, and Schizophrenia. Finally present challenges that face DL models.	They discussed a few approaches of DL and when they didn't present image preprocessing techniques but show what the DL can be used for?.
Our proposed survey	(1)introduce different preprocessing techniques which processed on neuroimaging, (2) combine preprocessing techniques with most common DL methods in one survey, (3) compared different state-of-art research with their challenges in dealing with dataset and classification stage	we discussed only the most common preprocessing techniques and ML is not discussed in detail as DL (we discussed in detail in section 7)

model algorithm, whether the prediction is correct or not, is adaptive and learns from the features automatically on its own [89].

　　The major differences between ML and DL are summarized in Table 1. DL is a specific category (i.e., branch) of ML. ML extracts a relevant feature manually from input data. Then the extracted feature is used to update the model parameters that will help in the correct prediction (i.e., classification) process [146]. This does not apply with DL as relevant features are extracted automatically from data. In addition to that, DL implements end-to-end learning where data and the task required to be implemented are passed to the network [26]. As
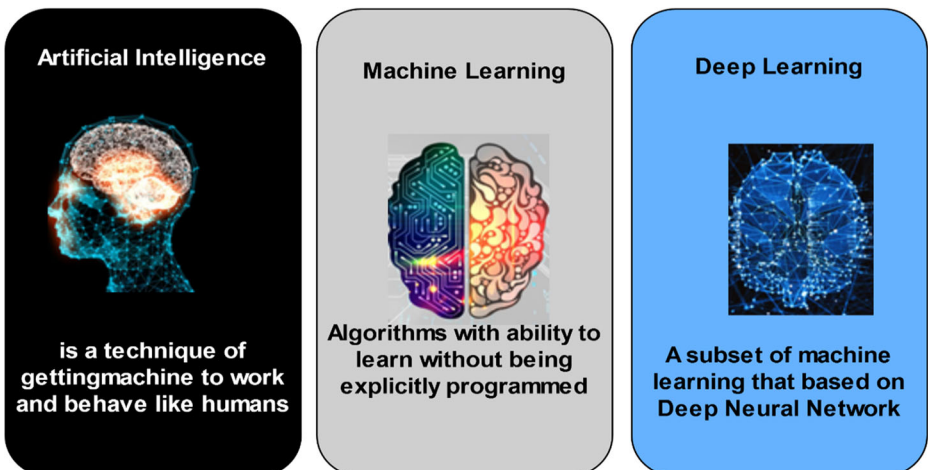
Fig. 2　Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) Definitions

mentioned, the learning process is done automatically from the features and hence no more manual modifications (i.e., enhancements) are required.

From Table 3. It compares ML and DL according to some factor data size, time of training data, and interpretability. ML provides several approaches and models that you can choose depending on your application, the size of the data you are processing, and the type of problem you want to solve [69]. To train the model, a high-performance DL application requires a very large amount of data (i.e., thousands of records) [51]. Over the last few years, DL has provided extensive applications in image recognition [101, 148], speech recognition [33, 97], medicine and pharmacy [84, 144], natural language processing [100, 155]. An extensive number of DL methods have been proposed recently [10], and these methods can be broadly classified into many algorithms that will be discussed in Section 3.

4 An overview on diagnosis of Alzheimer's disease

Diagnosis (i.e., classification) is an important area in computer science concerning the number of published articles recently [43, 151]. The AD diagnosis process passes through many stages to detect the disease and classify it as shown in Fig. 3.

From Fig. 3, the first stage is the data acquisition process for gathering the dataset required for diagnosis. The second stage is preprocessing the dataset to enhance the dataset quality and improve the performance of the classification task. The third stage is the splitting stage. The dataset can be split into training, testing, and validation subsets. The last stage is a learning system with proper and specific techniques to extract the features, learn from the data, update the parameters, and classify the disease to a specific class. The following subsections discuss these stages in detail.

4.1 Data acquisition stage

The first stage is the acquisition of raw data. Information in images describes inner aspects of the body that can be taken with different modalities or techniques. In this process, we can collect neuroimaging images that may utilize different physical principles. There are different modalities of neuroimaging such as (1) Magnetic Resonance Imaging (MRI), (2) Positron Emission Tomography (PET), (3) functional Magnetic Resonance Imaging (fMRI), and (4) Computed Tomography (CT) [138]. Selecting one of these modalities depends on the researcher's choice, task, and the used model. Datasets can be acquired from different organizations such as hospitals, clinical centers, radiology centers, and online websites [149].

Table 3 Difference between Machine Learning and Deep Learning Techniques

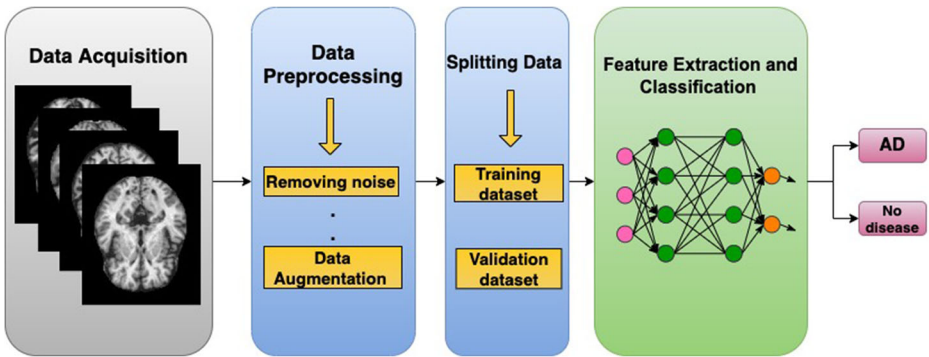| Factors	Machine Learning (ML)	Deep Learning (DL)
Data Size	Can handle small to medium datasets	Can handle large amounts of data
Training Time	Can take less time concerning the dataset size	Can take a long time concerning the dataset size
Interpretability	Easier to interpret the algorithms or models	Harder to interpret the algorithms or models

Fig. 3 The Process of Diagnosis Alzheimer Disease

4.1.1 Biomarkers and neuroimaging for Alzheimer's disease diagnosis

Conventionally, the clinical diagnosis of dementia has concentrated on (1) clinical assessment, (2) neuropsychological testing, and (3) the exclusion of other possible causes [102]. Normally, the diagnosing of AD can be accomplished with three altered methods, as depicted in Fig. 4.
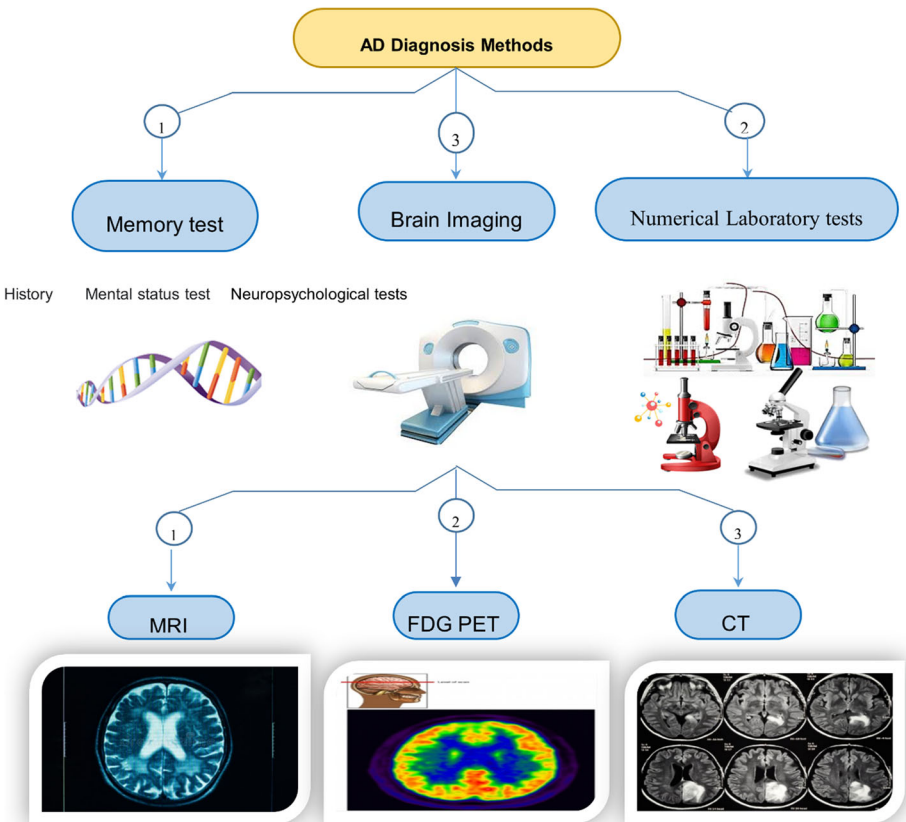

Fig. 4 Different Imaging Techniques for AD Diagnosis

From Fig. 4, these three methodologies are (1) Memory test through history and discussion, mental status test, and neuropsychological tests. (2) numerical laboratory, and (3) brain imaging scan [21]. There has been a revolution in the part played by neuroimaging in AD study and practice in the last years. Diagnostically, imaging has motivated from a slight exclusionary role to a crucial position [72]. In research, imaging is aiding address numerous scientific interrogations. Concurrently the probability of brain imaging has extended rapidly with new modalities and innovative ways of acquiring images and of analyzing them [141]. The definite modalities included are magnetic resonance imaging (MRI; both structural and functional) and positron emission tomography (PET; for assessment of both cerebral metabolism and amyloid).

Imaging Module and Types These modalities have different strengths and limitations and as a result, have different and often balancing roles and scope [108]. Although additional data are required, imaging is preliminary to offer prognostic information at this premature preclinical phase. The necessity for an earlier and more definite diagnosis will only grow as disease-modifying therapies are identified [40]. This will be particularly true if, as expected, these therapies work best (or only) when introduced at the preclinical stage. Table 4 summarizes the different brain imaging in AD [73].

In Table 4, it compares advantage and disadvantage between neuroimaging modalities and shows open research area for three modalities. Brain imaging has different scans depend on the type of disease, there are CT, MRI, and PET imaging. Structural imaging provides information about the shape, and volume of the brain like CT and MRI, functional imaging showing activity of the brain and how cells work.

4.2 Preprocessing techniques

Datasets, especially images, may contain noise and distortions. Radiography noise is generally caused by changes in the sensitivity of the detector, diminished illumination of the object (i.e., low contrast), photographic limitations, and spontaneous variations in the radiation signal [111]. So, it is essential to preprocess the data to enhance its quality or to optimize its geometric and intensity patterns [107]. Preprocessing lets researchers concentrate on a specific part of the brain and highlights the most vital information that is required in the classification process. Preprocessing techniques are many and the most common ones are depicted in Fig. 5.

Figure 5, divide AD diagnosis into two mainly step. Firstly, preprocessing techniques on images. Secondly, techniques which used in learning and classify diseases. Choosing one of these techniques according to the problem that the researcher wants to solve and the type of input data.

4.2.1 Intensity normalization

It is an essential preprocessing technique that is used for mapping the intensity of all image pixels to a reference scale [133]. In general, data collected from different sources, or the same source but at different points of time, may not have identical intensity ranges [120]. For example, normalization can calibrate different pixels to the normal distribution as depicted in Fig. 6 [115].

Table 4 Summary on the Different Brain Imaging in Alzheimer's Disease

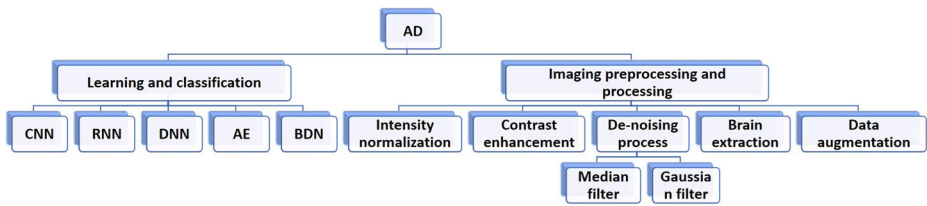|		Advantages	Disadvantages	Open Research Area
CT	Univariate	High-quality images. The information aids the health care provider by focusing attention away from unnecessary medical concerns. Modern CT scanners acquire this information in seconds depending on the examination [61].	The need for radiation exposure and the use of a contrast material (dye) in most cases, which may make it inappropriate for patients with significant kidney problems. Precise identification of small distal stones is occasionally difficult ([61]; [32]).	Longitudinal multimodality studies. Examining patients with additional unadorned cognitive impairment, as these techniques are right sensitive to head motion. Longitudinal functional imaging studies are needed to track the evolution of alterations in the fMRI activation pattern throughout the cognitive continuum from preclinical to prodromal to clinical AD. Imaging is progressively being combined in trial designs to evaluate the effects of therapy on fibrillary amyloid (through amyloid imaging), atrophy (by MRI), and metabolism (PET and fMRI). ([20]; [23]; [104]).
MRI	Structural	MRI measures of atrophy reflect cumulative neuronal damage which in turn is directly responsible for the clinical state. When compared with other imaging markers (and other biomarkers) cerebral atrophy has a strong connection with cognitive decay [30].	Shortages molecular specificity. It cannot right detect the histopathological hallmarks of AD. Atrophy patterns overlap with those of other diseases, and atypical types of AD show abnormal atrophy patterns as well. Individuals with claustrophobia and those who are more seriously impacted [56].	
	Functional	A noninvasive technique that delivers an indirect measure of neuronal activity, inferred from measuring changes in blood oxygen level-dependent contrast. Have documented the organization of the brain into multiple large-scale brain networks. Can determine the whole network of brain areas engaged when the person makes specific tasks ([157]; [19]).	Blood fMRI response is known to be variable across subjects, and very few studies examining the reproducibility of fMRI activation in older The responsiveness of the blood supply to the electrical signals that define neural transmission is poor [4].	
PET	FDG	Widely accepted to be a valid biomarker of overall brain metabolism to which ionic gradient maintenance for synaptic activity is the principal contributor.	It requires intravenous access and involves exposure to radioactivity, although at levels well below the significant known risk.	
	Amyloid	Determination of brain Aβ content to be moved from the pathology laboratory into the clinic. Amyloid imaging can detect cerebral β-amyloidosis and appears specific for this type of amyloid pathology, giving negative signals in pathologically confirmed cases of prion amyloid.	The widespread use of amyloid PET cost-effectiveness and availability [104].	

Fig. 5 Taxonomy of AD for Classification and Preprocessing Techniques

4.2.2 Contrast enhancement

Contrast enhancement is the difference between the highest and smallest pixel intensities as shown in Eq. 1. It improves the quality of images and increases the contrast of borders in the image that helps us to differentiate between organs. It improves the brightness of the image by expanding the range of pixel values as well [80].

$$g(x,y) = \frac{f(x,y) - fmin}{fmax - fmin} * levels \ of \ gray \tag{1}$$

where *fmin* is the minimum value, *fmax* is the maximum value, *f(x, y)* is the value of each pixel in the image, and *g(x, y)* is the enhanced pixel after that image contrast is applied [117]. Figure 7 shows a sample MRI axial brain image with high contrast.

4.2.3 Denoising process

Median filter The median filter is a technique that is used to minimize noise without blurring the edges [110]. It is especially appropriate for the enhancement of the required MRI images. The median filter identifies pixels as noise by matching each pixel in the image to its neighboring pixels [117]. It contains a filter (i.e., kernel) with a specific size that passes through each pixel value in the image and replaces it with the corresponding median value. The median value is determined by sorting the surrounding pixels' values and then replaces the
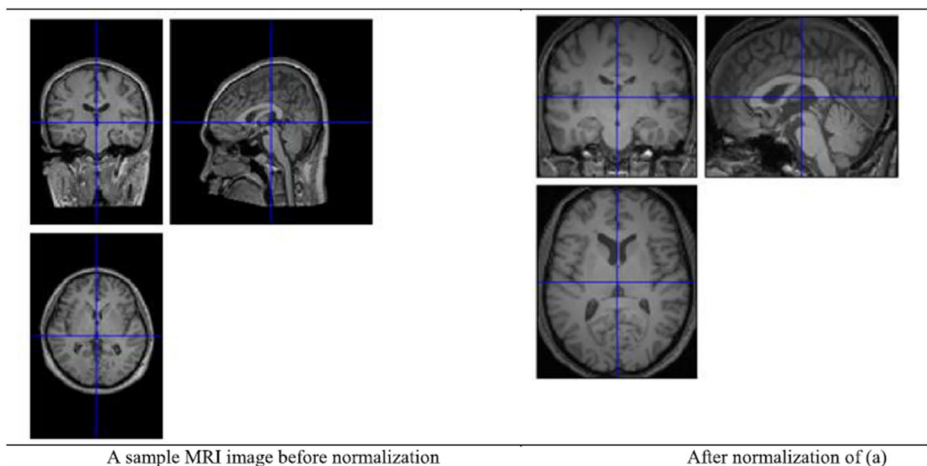

A sample MRI image before normalization   After normalization of (a)

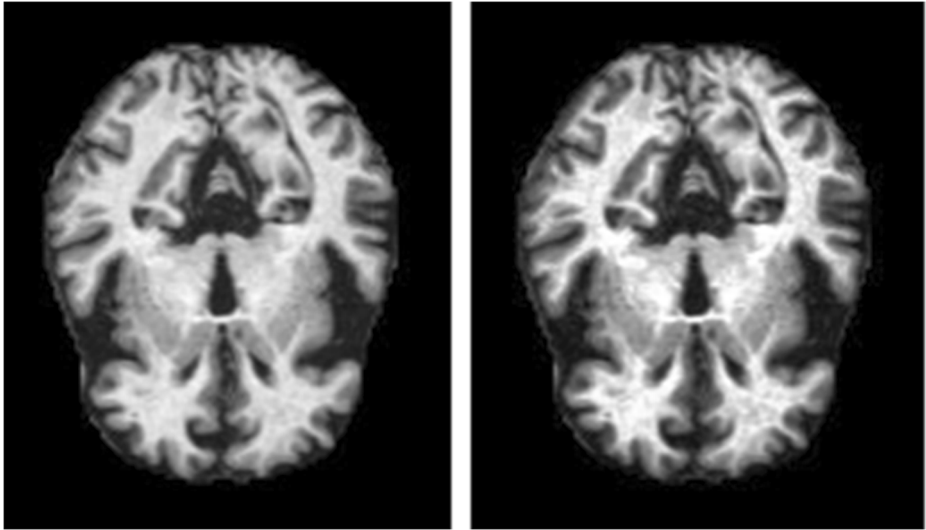**Fig. 6** Before and After Normalization Appliance [115]

Fig. 7 Applying the Contrast Enhancement Technique on a Sample Image. Left: Input Image and Right: Output Image

pixel with the corresponding middle pixel value [127]. Figure 8 shows the effect of applying the median filter on a sample image where the filter had s size of (3 * 3).

Gaussian filter Gaussian filtering is a technique that helps to denoising the image and is performed by detecting the size of the mask as shown in Eq. 2 [159].

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{\frac{-1}{2}*\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2}$$
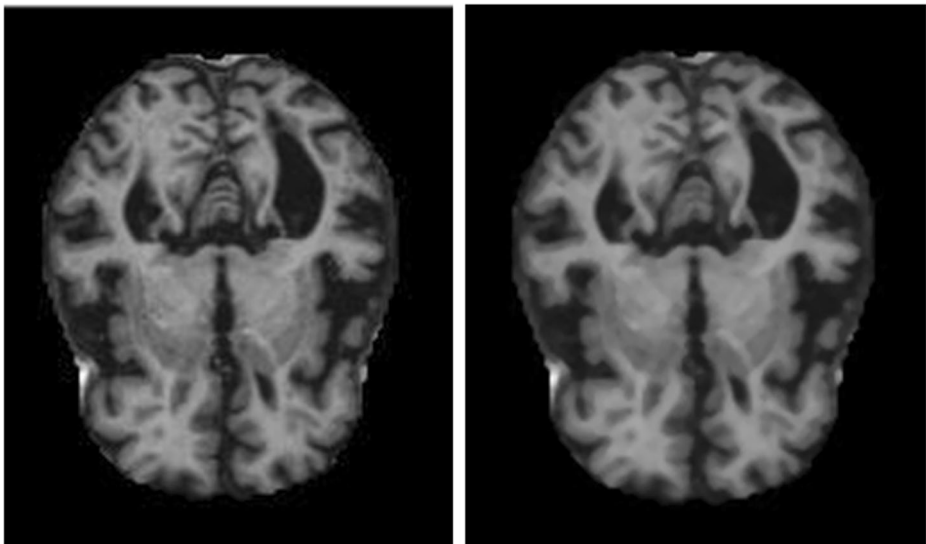

Fig. 8 Left: Before Removing Noise by the Median Filter and Right: After Applying the Median Filter

where σ is the standard deviation and defines how the Gaussian looks like, μ is the mean value, and x is the input value. Figure 9 shows the effect of applying the Gaussian filter on a sample image by using a mask size of (3 * 3).

4.2.4 Brain extraction and skull stripping

It is one of the preprocessing techniques that remove any non-brain tissues such as eyes, necks, and skulls. It segments these by using the dark space between the skull and brain occupied by the Cerebrospinal fluid (CSF). There are many tools presented in [75], that are used in brain extraction, such as the brain extraction tool ROBEX algorithm that uses machine learning. It may use DL also, but it will be a more extensive computational process and will need specific hardware to be able to run the algorithm. Khademi et al. [75] used the Random Forest classifier for brain extraction. They targeted to find a binary segmentation mask for the brain. After getting that mask, they multiplied it with the original image as shown in Fig. 10 [126].

4.2.5 Data augmentation

Data augmentation techniques help our model to avoid the overfitting problem [126]. The meaning of the overfitting is increasing in validation error with decreasing training error value so to build the best model, validation error must still decrease with training error [5]. After collecting your data, we can apply data augmentation to increase the images' diversity in each class. There are different techniques such as cropping, shifting, shearing, scaling, and zooming [95].

From Table 5, A literature review about preprocessing techniques and what is the methodology used to diagnosis AD. Also, we present the advantage and effectiveness of preprocessing on the images and show the performance of the training model.
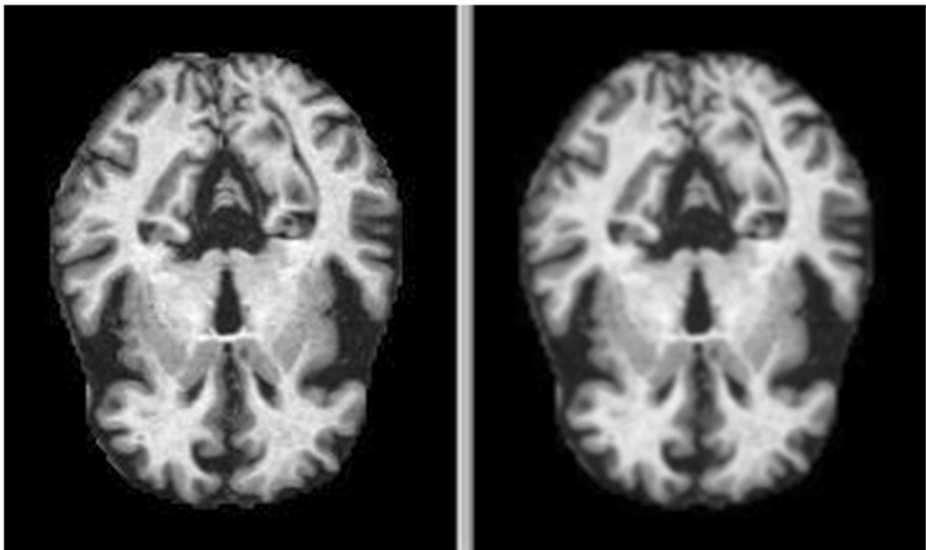

Fig. 9 Left: Before Removing Noise by the Gaussian Filter and Right: After Applying the Gaussian Filter
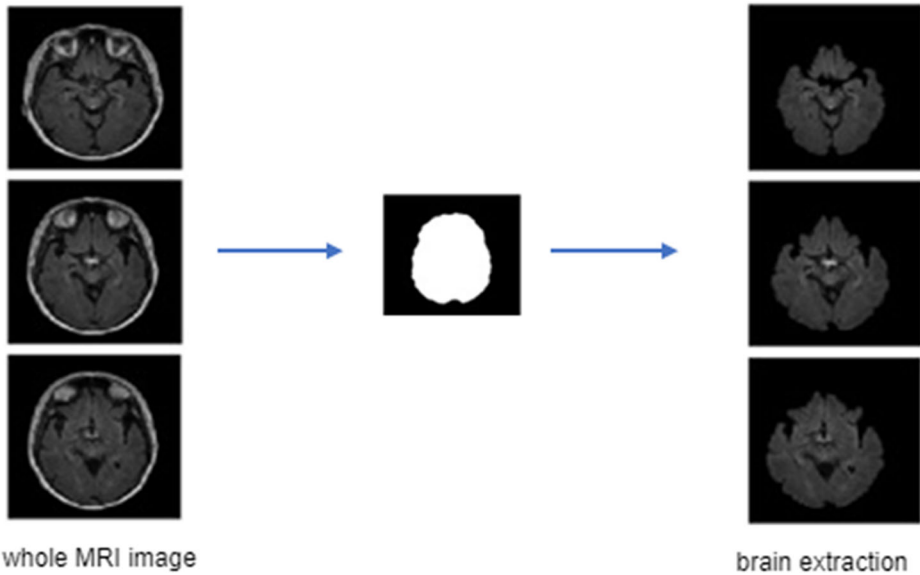
whole MRI image   brain extraction

Fig. 10 Brain Extraction with a Binary Segmentation Mask [75]

4.3 Classification

DL, as mentioned earlier, is a sub-field of ML [24]. DL is more effective than the traditional ways of ML because it extracts the features automatically [37]. Also, DL performs "end-to-end

Table 5 Compare different studies showing preprocessing techniques and classifiers

Study	Preprocessing techniques	Methodology	Advantages	Performance	
Ramzan et al. [112]	Intensity Normalization, brain extraction, high pass filter	2D CNN, ResNet-18, Transfer Learning (fine-tuning, off-the-shelf)	Removing noise from images and removes non-brain tissue from MRI images that increase and enhance the accuracy of learning.	ResNet-18 fine-tuning off-the-shelf	97.37% 97.88% 97.92%
Afzal et al. [5]	Data Augmentation	CNN (pre-trained AlexNet model)	Overcoming overfitting issues and improving the value of testing accuracy.	The model with Data Aug. Model without Data Aug.	98.41% 85.15%
Sarraf et al. [118]	High pass filter	CNN (GoogLeNet and LeNet-5)	To remove low-level noise from images they used a high pass filter with a frequency of 0.01 HZ.	GoogLeNet LeNet-5	98.84% 98.79%
Mehmood et al. [90]	Contrast enhancement and gaussian noise techniques	Siamese CNN(VGG16 model with extra one convolution layer in the model)	Improving contrast of datasets and improve performance of the model.	Proposed Siamese CNN	99.05%

learning" where raw data and tasks are provided to the network [17]. Most researchers depended on Convolutional Neural Network approaches for detecting Alzheimer's disease from MRI images compared to other techniques of DL such as Recurrent Neural Networks (as shown in Fig. 11(b)), Deep Neural Networks (as shown in Fig. 11(a)), Autoencoder (as shown in Fig. 11(c)), and Deep Belief Networks (as shown in Fig. 11(d)) [8, 62].

4.3.1 Deep neural network (DNN)

DNN, as shown in Fig. 11(a), has an input layer, output layer, and one (or more) hidden layers [50]. It is distinguished by dealing with complicated problems and understanding the relationship between input and output data, also able to model complex non-linear relationships [42]. It considers supervised learning techniques and is used in various areas of research to explore patterns between inputs that were unknown before [68]. It requires a large number of training data to extract the features of the labeled images [98].

4.3.2 Convolution neural network (CNN)

A CNN is one of the most successful techniques to perform image classification and recognition in neural networks. From Fig. 12, CNN is composed of several convolution layers, pooling layers, activation layers, fully connected layers, and a classifier layer.

The convolution layer is an essential layer that extracts the feature maps bypassing the learned filter (or kernel) with a specific size of the input image [42]. Then, it follows up the activation function that decides whether the neuron should be activated or not. It makes the nonlinear transformation to the input making it capable to learn and perform more complex tasks [2]. Activation functions have numerous types such as sigmoid, Tanh, and ReLU to
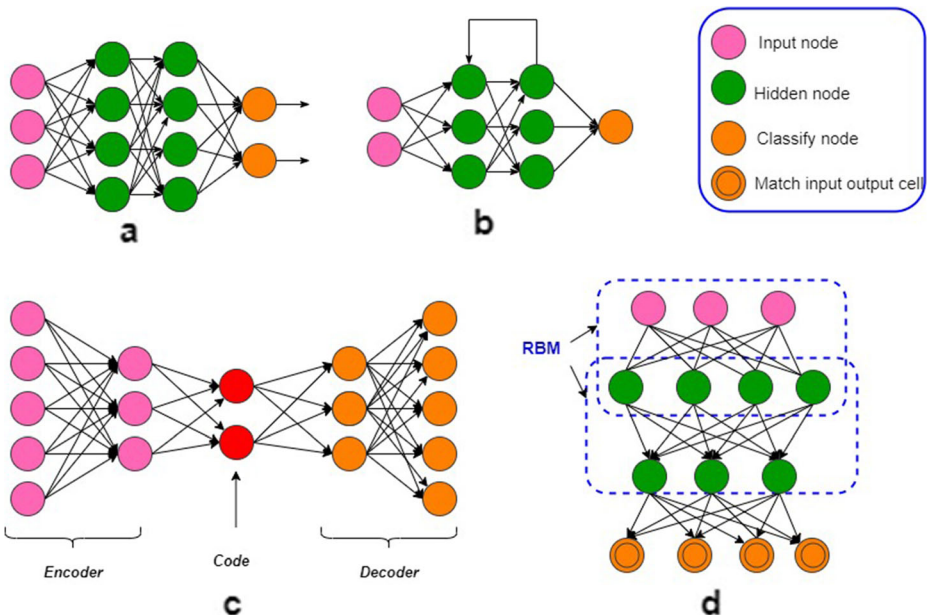

Fig. 11 **a**) DNN Architecture, **b**) RNN Architecture, **c**) Autoencoder, and **d**) Deep Belief Network
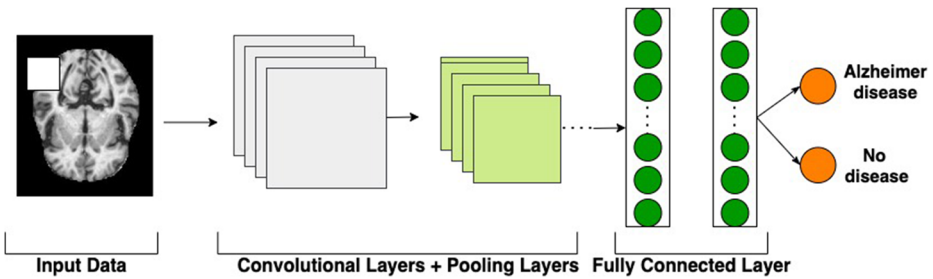
Fig. 12 Convolutional Neural Network Architecture

create feature maps [123]. Pooling layers reduce the dimensionality but keep the most important features. They can be considered as down-scalers [41]. A fully connected layer connects every neuron from the previous layer to all neurons in the current layer. Finally, the classifier layer selects a class (i.e., label) with the highest probabilities [7].

One important thing in CNN is that it can handle large datasets to get high performance in the classification task [53]. From the transfer learning point of view [140], CNN has several architectures that were trained on the ImageNet dataset including VGGNet, LeNet, GoogLeNet, ResNet, AlexNet [118]. With a pre-trained CNN model, the developer can benefit from the parameters (i.e., weights) of that model and transfer them to the new task [128, 150].

4.3.3 Recurrent neural network (RNN)

RNN is used in sequence- or time-series problems [91]. The most important advantage in that approach is the used memory and hidden state. Figure 11(b) shows a sample RNN architecture with an input, a hidden, and an output layer. The hidden state is effective in remembering the confident information about the problem sequence [106]. Another distinguishing characteristic of RNN is that they share the same parameters within each layer in the network unlike the feedforward networks [65]. The later networks, the feedforward networks, have different parameters for each node in the networks thus causing a large number of parameters [114].

RNN does not have a large number of layers and is not too deep compared to CNN's or DNNs [88]. However, the model is difficult to train and suffers from vanishing or exploding gradients limiting its application for modeling long-time activity sequence and temporal dependencies in sensor data [81]. The common applications of RNNs are natural language processing [154], speech recognition [57], and language translation [27].

Long short-term memory (LSTM) and Gated recurrent units (GRUs) are common architectures of RNN [36]. The main purpose of the LSTM is to maintain any error that occurs through the different layers and times [96]. It contains cells in the hidden layer, which have three gates: input, output, and forget gate. These gates are responsible for storing the information and regulating the flow of information to predict the output of the network [114]. This single-cell helps the model to decide which one to stock and when it can read and update the information through the gates [98]. GRUs use a hidden state and have two gates: reset gate and update gate. They control what and how much information will be retained [114]. Its performance, in many tasks, is better than LSTM [36, 42].

4.3.4 Autoencoder (AE)

AE is an unsupervised learning technique and can take an unlabeled dataset and compress it to feature encoded data. It is used for dimensionality reduction and consists of two major parts: an encoder and a decoder, as shown in Fig. 11(c) [98]. The encoder converts the input data to code (i.e., compressed data) and then the decoder rebuilds the code to the output which looks like the input [105]. Layers in the encoder part may be dense layers or convolution layers. The number of layers in the encoding part must be equal to the layers in the decoding part. The encoder reduces the dimensions of data, but it increases the dimensions of data [103]. The middle layer is called the bottleneck layer that compresses the representation of the input data [48].

AE has different types that improve the performance named (1) denoising AE, (2) sparse AE, and (3) contractive AE [1]. AE faces some problems such as a copy of the input layer to the hidden layer causes inefficient extraction of the meaningful features although it can retrieve the input in the output layer [134]. Denoising AE solves that problem by corrupting the inputs that the AE must then reconstruct or denoising [34]. This helps the model to recognize the feature from noisy input and hence can classify. The model does not copy the input to the output without learning features about data.

The sparse AE added many constraints to reduce the number of hidden nodes and limit nodes that were activated [135]. When the average of activation of the hidden nodes is close to zero this means nodes in the hidden layer are active and the other not active [147]. It can learn features by imposing some penalty, it is applying on the hidden layer. There are two ways to put sparsity constraint: (1) L1 regularization that is added to cost function which helps in preventing the overfitting problem [93] and (2) KL-Divergence constraint that is added to all hidden nodes to provide a low average activation value [11]. The main purpose of contractive AE is to support strong representation which will be able to extract useful information and be less sensitive to small variations in data [124].

4.3.5 Deep belief network (DBN)

DBN is a supervised learning technique that can link the unsupervised features, which are extracted from the stacked layer [70]. It is a generative graphical model and is constructed by a stack of Restricted Boltzmann machines (RBM) which can extract features and reconstruct the input. DBN has an undirected connection between the top of two layers as depicted in Fig. 11(d). DBN reduces the weight initialization by using RBM that helps the model overcome the overfitting problem [70].

DBN was created to analyze the apparent distribution between the input and the hidden layers in such a way that the lower layer node is connected directly, and the upper layer nodes are connected indirectly [92]. This model is helpful in a task that needs to extract features, involving biological data, and with classes that are not separated linearly [13].

Finally, the authors summarize the DL branches including the different models graphically in Fig. 13.

5 Research challenges

In this section, we present some of the research challenges that can be divided into two categories (1) the first is related to the data, (2) the second is related to the classification problem. The
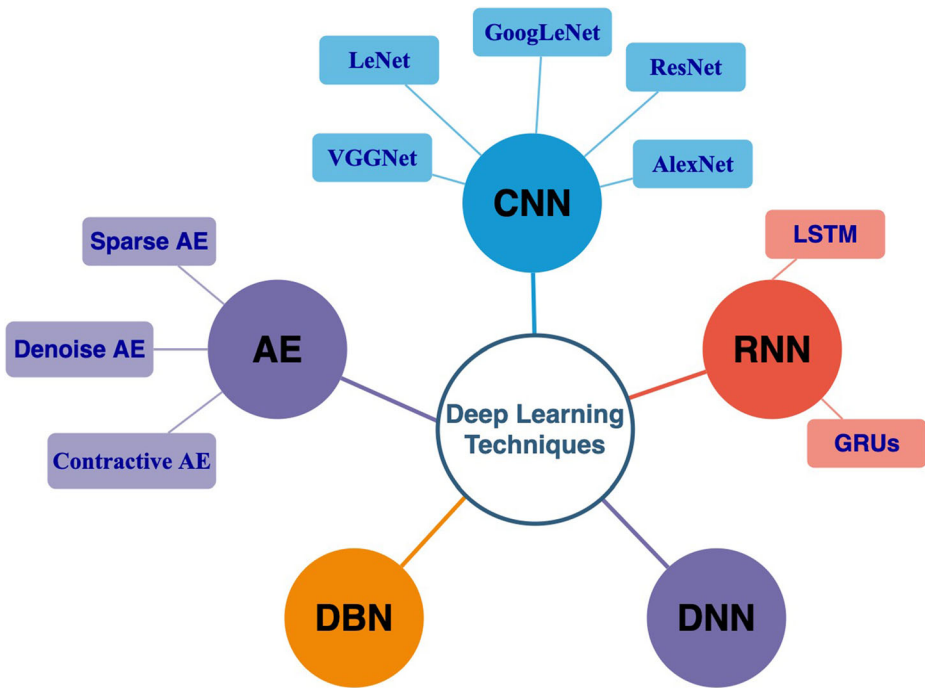
Fig. 13 Summarize section of Deel Learning Techniques

important challenges for each category are highlighted in the following subsections. The authors summarize the challenges section in Fig. 14 and discussed it in detail in the following subsections.

5.1 Availability of large datasets

To achieve the best result, DL techniques require a large number of datasets for the training process. Unavailability of data is a major challenge as it is difficult to acquire data from hospitals and clinical centers due to the privacy of patients. A set of online medical datasets available for a researcher for example ADNI (Alzheimer's Disease Neuroimaging Initiative), OASIS (Outcome and Assessment Information Set), COBRE (CENTR for biomedical research excellence), and the FBIRN (Function Biomedical Informatics Research Network). To overcome the small diversity of the available datasets, data augmentation is a technique that is used to increase the number of images without adding a new image by flipping, padding, rotation, etc. Table 6 summarizes the different methods used in different studies to solve the issue of limiting datasets.

In Table 6, we compare different studies how they overcome the limitation of availability of data. We compared their advantage, disadvantage, type of image modalities, datasets, and methods they used for overcoming this problem.

5.1.1 Alzheimer disease datasets

To have the ability to train systems and compare performance between architecture by using a different dataset, large datasets are required for training, testing, and validation. Table 7,
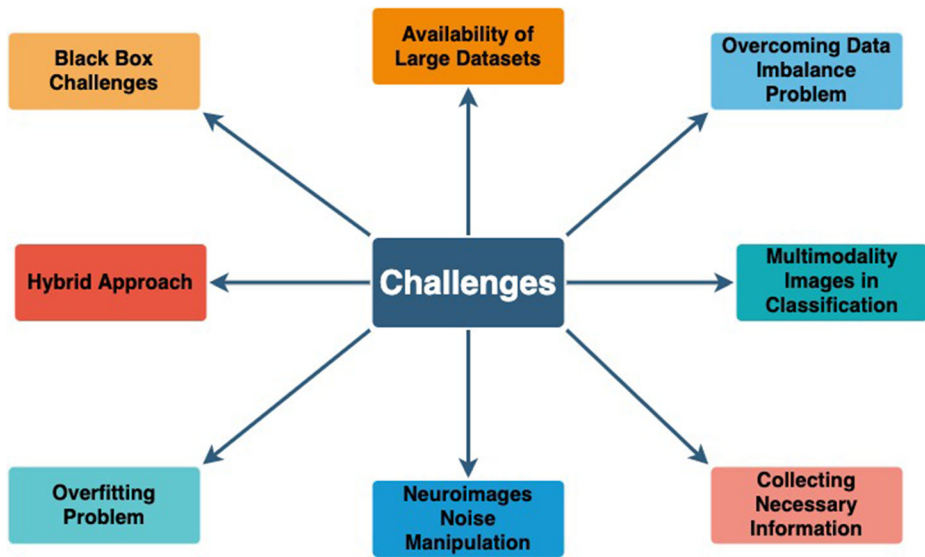
Fig. 14 summarize challenges of data and classification

Table 6 The Different Methods to Overcome Data Limitation

Study	Modality	Datasets	Contributions	Advantages	Disadvantages
Vu et al. [143]	MRI+ PET	ADNI	Combine different modalities of images PET and MRI images.	Increase the number of datasets.	Computational complexity and longtime training dataset.
Jain et al. [67]	sMRI	ADNI	Used transfer learning due to the amount of data is not available.	Overcome the problem of fewer amounts of data and reduce the computational cost.	–
Salehi et al. [116]	MRI	ADNI	Combine two datasets (ADNI1 Annual 2 Yr 3 T and ADNI1 Baseline 3 T) to increase the number of images.	Less computational process than using any techniques.	Different datasets may have unequal image sizes and formats.
Basaia et al. [16]	MRI	ADNI + Milan dataset	Combine two datasets ADNI and Milan dataset in the test phase.	Get higher accuracy in testing with multi datasets in (c-MCI vs HC, s-MCI vs HC, AD vs c-MCI, AD vs s-MCI, c-MCI vs s-MCI).	By using two datasets in classification AD vs HC has a lower accuracy than testing by ADNI dataset only.
Wang et al. [145]	MRI (DTI+ fMRI)	Beijing Xuanwu Hospital	Extract features from each modality separately and concatenates the features of each model to the classification and training process by using CNN.	Improve accuracy of classification by 92.06% instead of using only one modality and solve the problem of availability large data quantity.	Researchers were not given access to the classification features.

summarize the dataset of AD with their links. ADNI dataset is a multicenter study that aims to develop imaging, clinical, and genetic for tracking the growth of disease and to detect AD at the early stage (pre-dementia). There are 6 classes: Patient = 28, CN = 834, MCI = 671, EMCI = 340, LMCI = 185, AD = 450, SMC = 115.

OASIS is a collection of neuroimaging data sets that are open access for research and analysis. On XNAT Central, you may see and download 3-OASIS datasets. OASIS-1 is a collection of cross-sectional MRI data from 434 scan sessions including 416 participants. It was first released in 2007. OASIS-2 is a collection of longitudinal MRI data from 373 imaging sessions in 150 people. In 2010, it was made available. OASIS-3 is made up of 1098 subjects' cross-sectional and longitudinal MRI and PET data. It was released in the year 2018.

Alzheimer's Dataset (4 classes of images) is from open access Kaggle website that consisting MRI images. The dataset contains two files training and testing each of them containing 5121 and 1279 images respectively. There are 4 classes in training file: Mild Demented = 717, Very Mild Demented = 1792, Moderate Demented = 52, Non Demented = 2560.

MIRIAD is a database of volumetric MRI brain-scan of 46 Alzheimer's sufferers and 23 healthy elderly people. It includes a total of 708 scans and should be of particular interest for work on longitudinal biomarkers and image analysis. It is also an open-access dataset.

Table 7, presents available datasets with their links, the number of available images, shows the number of classes,s and is open access for the researcher or not?.

5.2 Overcoming data imbalance problem

Data imbalance is one of the problems that face researchers when they solve any problem by using DL. We may describe this as the distribution of examples across classes is not equal. For example, in the AD dataset, the number of images that have a disease is larger than images with no disease in which will cause an imbalance in the dataset [71].

There are different suggestions to solve this problem as shown in Table 8. Under-sampling and oversampling are two common techniques to enhance the data imbalance problem. Under-sampling is done by deleting data randomly which have enough amount of data in class to balance between classes [113]. It helps in obtaining an equal number of samples in classes and fast training time. Although the simplicity of that approach, there is a high probability that the

Table 7 Datasets for Alzheimer's disease

Dataset	Links	Data size	Availability	No of classes
ADNI	https://ida.loni.usc.edu/login.jsp?project=ADNI	2623	Open access	6 classes (AD, CN, MCI, EMCI, LMCI, SMC)
OASIS	https://www.oasis-brains.org/	–	Open access	2 classes
Alzheimer's Dataset (4 classes of images)	https://www.kaggle.			

com/tourist55/alzheimers-dataset-4-class-of-images5121Open access4 classesMIRIADhttp://miriad.drc.ion.ucl.ac.uk/708Open access–

data which we are deleting may have essential information or features for the classification of this class.

Over-sampling means increasing the amount of data by copying the existing sample. So, to achieve balanced classes, increase the size of the minority class. This process is done on the minority class which has a smaller number of data than other classes. Overfitting is the major issue that occurs with over-sampling [113].

In Table 8, we compare different studies on how they overcome the problem of data imbalance. We compared their advantage, disadvantage, type of image modalities, datasets, and methods they used for overcoming how to make different classes have the same number of images.

5.3 Multimodality images in classification

There is different scan type for neuroimaging such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI) and Computed Tomography (CT). Learning the model used in this different modality is another challenge. Learning Heterogeneity data may cause less performance because all input data are

Table 8 The Different Methods to Balance the Different Classes

Study	Modality	Datasets	Contributions	Advantages	Disadvantages
Afzal et al. [5]	MRI	OASIS (NC= 98, AD= 70, MCI= 28, LMCI =2)	Imbalance classes may cause an overfitting problem and insufficient training of the dataset. Data augmentation is used with different shapes like image rotation, crop from (top, bottom, left, center, and right), and whole crop.	Balancing datasets and overcome the overfitting problem and limiting of data available.	In most cases, data augmentation information isn't very accurate.
Khvostikov et al. [78]	sMRI + DTI	ADNI (AD= 48, MCI= 108, NC= 58) For each patient, there are sMRI and DTI modalities	Because of different class capacities, there is an imbalance in datasets that cause an overfitting problem so, to imbalance these classes they proposed a data augmentation process.		Adding DTI modalities in small data may cause the network to be noisier.
Farooq et al. [44]	MRI	ADNI (AD= 33, LMCI =22, MCI =49, NC= 45)	To balance the presented data, they applied a data augmentation technique to balance between classes.	Enhance the quality of the model in training data.	The risk of the overfitting problem may happen in over-sampling.
Mehmood et al. [90]	MRI	OASIS (AD =23, LMCI =87, MCI =105, NC =167)	Due to the overfitting problem in the dataset that was imbalanced, they used data augmentation techniques to balance the dataset.	Improve the model learning rate.	The best class distribution is unknown.

different and combine [82]. To solve this problem every single modality is learning separately by multi hidden layers to extract features then the second stage is combining features from the last hidden layer of each modality and then learn a model to classify labels by using combined features from the last stage.

A combination of features from different modalities has high performance than a training model with a single modality. Each neuroimage modality can offer new different details for the disease that make classification more effective [18]. Table 9 shows a summary of the different methods of the papers used in solving a combination of different modalities of the images.

In Table 9, we compare different studies which used different modalities of images in the learning model. We compared their advantage, disadvantage, type of image modalities, datasets, which classifier they used, and methods they used for overcoming this problem.

5.4 Collecting necessary information

MRI images may have some information that is unnecessary for diagnosis of AD that increases the time of processing and training data, computational process, and cause less efficiency in the training of our classification model [90]. The solution for this challenge is using techniques for preprocessing data before training it as shown in Table 10. It summarizes some of the related articles that solve this problem. It is worth mentioning that FreeSurfer is a free tool from the internet used for processing images like skull stripping, segmentation for the essential part of the brain for diagnosis of Alzheimer's disease [22].

In Table 10, we compare different studies on how they collecting important information in images and how to discard other information. We compared their advantage, disadvantage, type of image modalities, datasets, and methods they used for overcoming this problem.

5.5 Neuroimages noise manipulation

Adversarial noise may be found in neuroimages and this reduces the performance of the classification process. To minimize noise, as we mentioned before in Fig. 5, we present some preprocessing techniques that help in removing the noise from images as shown in Table 11. It summarizes the different methods that are used in removing noise from neuroimages. Gaussian filter, median filter, and many filters that remove noise from images increase the quality of training of the classification model.

In Table 11, we compare different studies which have datasets or images with noise. We compared their advantage, disadvantage, type of image modalities, datasets, which classifier they used, and methods they used for overcoming neuroimaging noise.

5.6 Overfitting problem

Overfitting happens when the model is trained on data that have noise and more unuseful details. The size of data used for training may not be enough is also one of the reasons that cause overfitting in this case to solve this problem we need to increase the amount of data [59]. To overcome this problem, the dropout can be used to drop the units (i.e., hidden and visible) in a neural network as shown in Table 12 [130]. In it, some articles used that method and others solve that problem in another way. This means removing the units temporarily from the network. Thereby, choosing a random sample of neurons to train rather than train whole neurons in the network. This can make the learning of the hidden layer better.

Table 9 The Different Methods to Solve Multimodalities Images in the Dataset

Study	Classifier	Modality	Datasets	Contribution	Advantage	Disadvantage
Vu et al. [143]	Sparse AE+ CNN	MRI+PET	ADNI	Pretraining each modality by using Autoencoder and then using CNN by extracting feature standalone and the combine it by using a fully connected layer.	Improver accuracy than training each modality standalone and finding the most discriminant features.	Due to the large features extracted, it consumes more time and computational complexity in training.
Lee et al. [82]	RNN	cerebrospinal fluid (CSF) biomarker + MRI+ Cognitive score+demographic data	ADNI	They extracted features separately for each modality and at the end, they combine features for four modalities.	Increase predictive power so its result shows better performance. The most advantageous thing the model can use irregular data without preprocessing.	The performance enhancement is lower because of the shortage of positive samples.
Liu et al. [86]	SVM	MRI+PET	ADNI	The aforementioned feature selection method is used to calculate the kernel to select a feature. Then they combine kernels of features linearly and make classification by SVM.	Effectively combine the complementary information from multi modalities data. Improve accuracy of classification.	Explanation of kernel classifier is hard.
Venugopalan et al. [142]	3D-CNN, stacked denoising autoencoder	Clinical data+ genetic data + MRI	ADNI	Every modality is trained separately. They used 3D-CNN to extract features from MRI images and stacked denoising auto-encoder to extract features from clinical and genetic data. They concatenated the extracted feature from both classifiers and passed it to the classification layer.	Multiple data modality merging will provide a comprehensive view of the AD staging research.	In this model, the challenges are the small size of ADNI data.
Aderghal et al. [3]	CNN with transfer learning	DTI-MD and sMRI	ADNI (ADNI Go + ADNI 2)	Different modalities are preprocessed by the alignment of MD corresponding to sMRI to select the same hippocampal region. Statistical Parametric Mapping software (SPM8) is used to transform MD and sMRI.	Increase the number of images that enhance classification and solve the problem of insufficient samples.	The resolution of normalized sMRI and MD is low.

Table 10 The Different Methods of Removing Unnecessary Data from the Image

Study	Modality	Datasets	Contribution	Advantage	Disadvantage
Ramzan et al. [112]	fMRI	ADNI	FSL-BET toolbox to remove non-brain tissue, using FSL-MCFLIRT toolbox for removing motion correction from the image.	Improved the quality of classification.	Increased processing time.
Jain et al. [67]	sMRI	ADNI	Using the FreeSurfer tool to remove unnecessary details and intensity normalization.	Improved training of classification model.	This tool takes more time to run.
Feng et al. [45]	MRI	ADNI	Skull stripping using Computational Anatomy Toolbox12 (CAT12) and spatially normalization by using statistical parametric mapping 12 (SPM12).	CAT12 toolbox distinguishes with fast running as compared to FreeSurfer tool and lower processing time.	The effectiveness of Skull striping segmentation depends on the user-defined parameters that do not appropriate in low contrast images.
Farooq et al. [44]	MRI	ADNI	Using SPM 8 tool to skull stripping and gray matter segmentation. Python Nibabel package to convert GM volume to JPEG slices.	Focuses on important features in the image and enhances dataset performance.	
Mehmood et al. [90]	MRI	OASIS	Extract the necessary parts of the image by segmentation using K-mean Clustering and skull stripping.		K-mean clustering has an issue that affects the result to predict the k value and does not work well with clusters that have different sizes.

In Table 12, we compare different studies in which their architecture suffers from an overfitting problem. We compared their advantage, which classifier they used, methods they used for overcoming overfitting problems, and shows the performance of each study.

5.7 Hybrid approach

The hybrid model is defined as combining more than an approach or technique to achieve high performance or enhance the training and classification process [99]. One example of a hybrid method is feature selector and CNN model, feature selector with the pre-trained model or transfer learning, and hyperparameters optimizer with any model of DL. As shown in Table 13, some of the articles combined more than one learning method or classification technique to enhance the overall classification of the disease.

Table 11 The Different Methods of Removing the Noise from Neuroimages

Study	Classifier	Datasets	Contribution	Advantage	Disadvantage
Ramzan et al. [112]	2D CNN, ResNe-t-18, Transfer Learning	ADNI	Used high pass filter and spatial smoothing.	Removing noise from images and enhance learning.	Taking processing time with a large dataset.
Sarraf et al. [118]	CNN (GoogLe-Net and LeNet)	ADNI	Used FSL-MCFLIRT to correct motion artifact. Hanning windowed is used to make slice time corrections. To remove low-level noise from images they used a high pass filter with a frequency of 0.01 HZ.	Improve the quality of image that improves the accuracy of classification.	
Mehmood et al. [90]	VGG16 model with extra one convolu-tion layer in the model	OASIS	They used contrast enhancement and gaussian noise techniques because through the acquisition process some of the noise will be in the image such as nonlinear light intensity that affects the performance and quality of classification.	Improving contrast of datasets and improve performance of the model.	
Al-Naami et al. [9]	ANN	Jordanian Hospital's datasets	They applied a low pass filter. Morphological filters to extract statistical output.	Helping in removing noise from images that confuse and make the performance less in classification.	The morphological filter is dependent on structuring element choice.
Kazemi et al. [74]	CNN (AlexNet)	ADNI	Spatial smoothing is used to reduce the noise level, it is done by using a Gaussian kernel of 5 mm full width at half-maximum. A high pass filter with 0.01 HZ is used to remove low-level signals.	Preprocessing increases the sensitivity of the analysis.	In the case of spatial smoothing be larger than the activated region causes loss in the main signal that may reduce the effectiveness of classification.

In Table 13, we compare different studies which used a hybrid approach in their studies. We compared methods they used for combining more than one model, shows the performance of each model, type of image, datasets that they used, which classifier they used, and our comments about their methods.

Table 12 The Different Methods for Overcoming the Overfitting Problem in Classification

Study	Classifier	Contribution	Advantage	Accuracy
Mehmood et al. [90]	CNN	Data augmentation applied to images to prevent overfitting problems and applied transfer learning.	Prevent overfitting and Improve the performance of the classifiers.	99.2% by using ADNI in testing and 98.2% by using ADNI and Milan datasets.
Jain et al. [67]	VGG16	Build vgg16 model and used dropout neurons	Avoid overfitting problem	95.73% in 3-way classification 99.14% (AD vs CN) 99.3% (AD vs MCI) 99.22% (CN vs MCI)
Basaia et al. [16]	Deep learning model	Used several strategies to fight the overfitting problem. Batch normalization Dropout of neuron in the hidden layer L2 regularization	Accelerate the speed of training in neural networks and each neuron is stronger in casual variations	86% by using structural MRI+ cognitive measure + APOe4+ demographics in input data
Aderghal et al. [3]	VGG16 model with extra one convolution layer in the model	Applied data augmentation technique to overcome overfitting problem	Effective in understand maximum features	Achieved 99.05%
Spasov et al. [129]	CNN with transfer learning	They used a different dataset that causes sparse in the dataset. Data augmentation techniques are used to overcome the overfitting problem.	Improve learning performance and increase the accuracy of classification	Accuracy is achieved with 92.5% (AD vs NC) 85% (AD vs MCI) 80% (MCI vs NC)

5.8 Black box challenges

One of the trending issues is the black box. Neural networks, which can be thought of as black boxes that convert input into output, are often used in machine learning methods [49]. Although math used to construct a neural network is straightforward how the output arrived is exceedingly complicated, ML algorithms get a bunch of data as input, identify patterns, and build a predictive model but understanding how the model worked is an issue [85].

Although DL has the most success in achieving high performance close to human in classification and predicting process, operate as black boxes [49]. It doesn't offer a specific reason or explanation for choosing a specific feature in the training process or why this achieves high or low performance or how the training data's relations are reflected in the feature selection as shown in Table 14 [25].

In Table 14, we compare different studies about black box challenges. We compared their model, shows the performance of each model, the type of image, datasets that they used, which classifier they used, their contribution, and our comments about their methods.

Table 13 Summarize Different Hybrid Approaches in AD

Study	Methodology	Modality	Datasets	Contribution	Comments	Model	Performance	
							AD vs NC	pMCI vs sMCI
Cui et al. [31]	CNN+RNN (three bidirectional gated recurrent units)	MRI	ADNI	Combined two models, first 3D-CNN for extract spatial features of images and after that, they used RNN to extract longitudinal feature. They made a comparison between CNN, RNN, and a combination of both models.	This approach ignores the correlation data between consecutive features. They used RNN to solve this issue.	CNN RNN CNN + RNN RNN	88.99 85.01 91.33	70.22 68.49 71.71
Dua et al. [38]	CNN, RNN, Long Short-Term Memory (LSTM)	MRI	OASIS	They divide data into 2 sample cross-section image datasets (OASIS-1) for training CNN and longitudinal dataset (OASIS-2) for training RNN and LSTM.	The bagged ensemble model decreases the variance and bias and achieves high accuracy than others.	CNN CNN with bagged model RNN RNN with bagged model LSTM LSTM with bagged model Simple ensemble Bagged ensemble	75.31 80.77 86.11 89.24 87.5 90.57 89.75 92.22	– – – – – – –
Li et al. [83]	Hybrid of CNN and RNN	MRI	ADNI	Hybrid CNN and RNN are used to classify AD by using hippocamps patch extracted from MR images. They combined DenseNet with the bidirectional GRU. The hybrid neural network is used with the internal hippocampus, whole hippocampal patches, and the decomposition of the hippocampus.	In CNN information may be lost when the features transformed from low to high levels, so they used DenseNet to connect each layer to every layer.	Internal-DenseNet-RNN Whole hippo DenseNet Hybrid	85.4 87.4 89.1	70.0 71.3 72.5
Shakarami et al.		PET	ADNI	They proposed the approach of AlexNet CNN and added to it a fully connected	In the testing method, they used majority voting. After classifying by using SVM,		86.64	–

Table 13 (continued)

Study	Methodology	Modality	Datasets	Contribution	Comments	Model	Performance AD vs NC	pMCI vs sMCI
[121]	Improved AlexNet-SVM			layer and using SVM as classifier instead of Softmax. They used a 2D slice in the training method to choose high-quality and best slices.	they count slice that belongs to each class and use majority voting on slices and show the result.	Classification with the proposed AlexNet-SVM Diagnosis with the proposed method and using majority voting	96.39	–
Khagi et al. [77]	CNN and FS by ReliefF, Laplacian, Mutinffs, and UDFS. classification by SVM and KNN	MRI	OASIS	They extracted features by using the proposed CNN and combined these features with other features that were extracted manually by using Gray Level Co-occurrence matrix (GLCM). They used feature selection algorithms to rank features and classification using SVM and KNN.	ReliefF and UDFS are the best algorithms in performance and have less execution time with few seconds.	SVM KNN	99.84 99.75	– –

Table 14 The different articles for black-box challenges

Study	Methodology	Modality	Datasets	Contribution	Comments	Performance	
						Model	Accuracy
Folego et al. [46]	3D-CNN	sMRI	ADNI, CAD–Dementia	3D-CNN to ADNI dataset and used CADDementia challenges in evaluating the step. Regularization of L1 and L2 norms are used to avoid overfitting problem. The different models of CNN (LeNet-5, VGG 512, GoogLeNet, ResNet) are used. VGG 512 is the best performance model which named is ADNet. They considered the domain adaption approach named ADNet-DA which trained the system with a dataset and evaluated it on another dataset.	L2 achieves the best result with all models except GoogLeNet achieves the best result with L1. CADDementia challenges need 19 h of computation which their proposed approach was achieved 10 times faster than required.	ADNet ADNet-DA	51.4 52.3
Khagi et al. [76]	Diverging 3D-CNN	MRI, PET	ADNI	Encoder-decoder based, converging, diverging, and equivalent each of them changes in kernel size to see the effect or performance of the model in classification. Diverging architecture (DivNet) is the proposed model which achieves the best testing accuracy. Different datasets were used with the proposed model, the combination of PET and MRI datasets have less accuracy than MRI. So, the MRI dataset is used rather than PET and MRI.	The number of the layer in diverging CNN model is chosen depending on the testing result after adding new layers starting from two layers to six layers. After 4 layers of convolution, the testing accuracy starts decreasing at both five and six layers. So, the diverging model is composed of 4 layers of convolution layer. The convolution layer is increased by one and testing the result.	Diverging with MRI DivNet with PET DivNet with MRI+ PET	94.59 66.6 82.12
Fu'adah et al. [47]	CNN (AlexNet)	MRI	Alzheimer's Dataset (4 classes of images)	Proposed AlexNet architecture to classify different classes of AD. 75% of the dataset is used in the training model. They used Adam optimizer with various value of learning rate 0.0001, .001, 0.01, and 0.1. At every learning rate value, accuracy and loss were measured, learning rate with 0.0001 achieved the best accuracy and loss performance. The best result of the model was at a 0.0001 learning rate.	According to the result for this model, they showed an increasing value of learning rate that may cause raising in the value of loss that named overshooting.	AlexNet	95

Table 14 (continued)

Study	Methodology	Modality	Datasets	Contribution	Comments	Performance	
						Model	Accuracy
Yildirim et al. [153]	CNN	MRI	Alzheimer's Dataset (4 classes of images)	They used Resnet50 architecture instead of the training model from scratch, the last five layers from Densenet50 are removed and added 10 layers instead of removed layers.	They classify disease in 4 classes nondemented, ModerateDemented, MilldDemented, VeryMildDemented.	Densenet201	87
						VGG16	78
						AlexNet	86
						Resnet50	78
						Proposed	90

6 Future directions

By revising the most recent literature on early diagnosis of Alzheimer's disease, it was concluded that. To achieve overall improvement and upgrading the accuracy of diagnosis using a computer application, the following points must be taken into account:

- One of the challenges is collecting brain-balanced and sufficient data related to Alzheimer's disease [60, 119, 125, 131].
- Most of the recently deployed methods and techniques correlated to DL, including deep sparse multi-task learning [131], stacked auto-encoder [137], and sparse regression models [132], each is attempting to overcome the aforementioned challenges. In [119], proposed a deep architecture to remove the features without load redundant information by using sparse multi-task learning in a hierarchy.
- Deep learning segmentation (e.g., U-Nets) can be injected in the process to specify only the region of interest.
- Combining two various conceptual methods of sparse regression and DL to diagnose AD can be effective [125]. Also, one of the promising techniques is the manifold-based learning method.
- Data augmentation and scaling techniques can help to improve the overall state-of-the-art performance.

7 Limitations

All studies have both strengths and weak points so, we have several limitations. First, we mentioned only the most common preprocessing techniques (intensity normalization, contrast enhancement, De-noising process, brain extraction, and data augmentation) that are used with neuroimaging. Second, we discussed only five techniques of DL (DNN, ANN, CNN, AE, and DBN) although there are many approaches we mentioned the most common one with a diagnosis of AD. Third, ML is not discussed in detail as DL. Finally, we mentioned four datasets, not all of them, and the current study works on ten years of study.

8 Conclusions and future work

AD is a cumulative neurological disorder that is the most common form of late dementia. AD causes nerve cell death and tissue loss in the brain, resulting in a substantial decrease in brain volume over time and impairment of most of its functions. In this paper, we started with the big difference between traditional ML and DL, followed by the stage of diagnosis of AD. In the diagnosis of AD, we need to preprocess images to enhance the quality of learning, so we show some preprocessing techniques used with images. And also, we presented different methods of DL that are most common in the classification process such as CNN, RNN, DNN, AE, and DBN. Although the importance of the classification of disease by using DL, there are challenges for dealing with the dataset.so, we presented a review of literature for every challenge and show their suggestion to solve these problems. The novelty of the current survey can be summarized in (1) introduce different preprocessing techniques which processed

Table 15 Table of Abbreviations

Abbreviation	Definition
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AE	Autoencoder
CN	Cognitive Normal
CNN	Convolution Neural Network
COBRE	Center for Biomedical Research Excellence
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DBN	Deep Belief Network
DL	Deep Learning
DNN	Deep Neural Network
FBIRN	Function Biomedical Informatics Research Network
fMRI	Functional Magnetic Resonance Imaging
FS	Feature Selection
GPU	Graphical Processing Unit
GRUs	Gated Recurrent Units
LMCI	Late Mild Cognitive Impairment
LSTM	Long Short-Term Memory
MCI	Mild Cognitive Impairment
ML	Machine Learning
MRI	Magnetic Resonance Imaging
OASIS	Outcome and Assessment Information Set
PET	Positron Emission Tomography
PPMI	Parkinson's Progression Markers Initiative

on neuroimaging, (2) combine preprocessing techniques with the most common DL methods in one survey, (3) compared different state-of-art research with their challenges in dealing with dataset and classification stage. In future studies, to classify AD with the proper dataset, we can use (1) Abstract CNN Models, (2) apply transfer learning only, (3) apply both transfer learning with abstract CNN models, (4) use feature selector to select feature separately and after that using CNN models, (5) use feature selector to select feature separately and after that using transfer learning, (6) compare the performance of two models that mentioned in 4 and 5 points, and (7) using hyperparameters optimizer with one of the models like CNN or transfer learning.

Table of Abbreviations

Table 15 presents the used abbreviations in the current survey with the corresponding definitions. They are sorted in alphabetical order.

Authors agreement We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

Corresponding author We understand that the declared corresponding author is the sole contact for the editorial process (including Editorial Manager and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions, and final approval of proofs. We confirm that we have provided a current, correct email address that is accessible by the Corresponding Author.

Intellectual property We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, concerning intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

Data availability We confirm that no datasets were users, generated, or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no competing interests nor conflict of interests for the current study.

References

1. Aamir M, Nawi NM, Mahdin HB, Naseem R, Zulqarnain M (2020) Auto-encoder variants for solving handwritten digits classification problem. Int J Fuzzy Logic Intell Syst 20(1):8–16
2. Activation function neural network (2020) https://www.geeksforgeeks.org/activation-functions-neural-networks/
3. Aderghal K, Khvostikov A, Krylov A, Benois-Pineau J, Afdel K, Catheline G (2018) Classification of Alzheimer disease on imaging modalities with Deep CNNs using cross-modal transfer Learning. 2018 IEEE 31st international symposium on computer-Based medical systems (CBMS). doi:https://doi.org/10.1109/cbms.2018.00067
4. Advantages and disadvantages of functional MRI (2019) https://www.ed.ac.uk/clinical-sciences/edinburgh-imaging/research/themes-and-topics/medical-physics/imaging-techniques/functional-mri
5. Afzal S, Maqsood M, Nazir F, Khan U, Aadil F, Awan KM, … Song O-Y (2019) A data augmentation based framework to handle class imbalance problem for Alzheimer's stage detection. IEEE access, 1–1. doi:https://doi.org/10.1109/access.2019.2932786
6. Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ (2017) Deep Learning for brain MRI segmentation: state of the art and future directions. J Digit Imaging 30(4):449–459. https://doi.org/10.1007/s10278-017-9983-4
7. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In 2017 international conference on engineering and technology (ICET) (pp. 1-6). IEEE. https://doi.org/10.1109/ICEngTechnol.2017.8308186
8. Aldweesh A, Derhab A, Emam AZ (2020) Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues. Knowl-Based Syst 189:105124. https://doi.org/10.1016/j.knosys.2019.105124
9. Al-Naami B, Gharaibeh N, Kheshman AA (2013) Automated detection of Alzheimer disease using region growing technique and artificial neural network World Acad Sci Eng Technol Int J Biomed Biol Eng, 7(5)
10. Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, van Essen BC, Awwal AAS, Asari VK (2019) A state-of-the-art survey on deep learning theory and architectures. Electronics 8(3):292. https://doi.org/10.3390/electronics8030292

11. Al-Qatf M, Lasheng Y, Alhabib M, Al-Sabahi K (2018) Deep Learning approach combining sparse Autoen-coder with SVM for network intrusion detection. IEEE Access 6:1–52856. https://doi.org/10.1109/access.2018.2869577

12. Amen DG (2015) Change your brain, Change Your Life (Revised and Expanded): The Breakthrough Program for Conquering Anxiety, Depression, Obsessiveness, Lack of Focus, Anger, and Memory Problems. Harmony

13. An N, Jin L, Ding H, Yang J, Yuan J (2020) A deep belief network-based method to identify proteomic risk markers for Alzheimer disease. arXiv preprint arXiv:2003.05776

14. Ayodele TO (2010) Types of machine Learning algorithms. New Adv Mach Learn. https://doi.org/10.5772/9385

15. Bardis MD, Houshyar R, Chang PD, Ushinsky A, Glavis-Bloom J, Chahine C, … Chow DS (2020) Applications of artificial intelligence to prostate multiparametric MRI (mpMRI): Current and emerging trends. Cancers 12(5):1204. https://doi.org/10.3390/cancers12051204

16. Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, Filippi M (2018) Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage: Clinical, 101645. https://doi.org/10.1016/j.nicl.2018.101645

17. Bashar A (2019) Survey on evolving deep learning neural network architectures. J Artificial Intell 1(02):73–82. https://doi.org/10.36548/jaicn.2019.2.003

18. Baskar D, Jayanthi VS, Jayanthi AN (2019) An efficient classification approach for detection of Alzheimer's disease from biomedical imaging modalities. Multimed Tools Appl 78(10):12883–12915. https://doi.org/10.1007/s11042-018-6287-8

19. Bi X, Zhao X, Huang H, Chen D, Ma Y (2020) Functional brain network classification for Alzheimer's disease detection with deep features and extreme learning machine. Cogn Comput 12(3):513–527 https://link.springer.com/article/10.1007/s12559-019-09688-2

20. Borghesani PR, Johnson LC, Shelton AL, Peskind ER, Aylward EH, Schellenberg GD, Cherrier MM (2008) Altered medial temporal lobe responses during visuospatial encoding in healthy APOE* 4 carriers. Neurobiol Aging 29(7):981–991. https://doi.org/10.1016/j.neurobiolaging.2007.01.012

21. Brown J (2015) The use and misuse of short cognitive tests in the diagnosis of dementia. J Neurol Neurosurg Psychiatry 86(6):680–685. https://doi.org/10.1136/jnnp-2014-309086

22. Brown EM, Pierce ME, Clark DC, Fischl BR, Iglesias JE, Milberg WP, McGlinchey RE, Salat DH (2020) Test-retest reliability of FreeSurfer automated hippocampal subfield segmentation within and across scanners. Neuroimage 210:116563. https://doi.org/10.1016/j.neuroimage.2020.116563

23. Busche MA, Eichhoff G, Adelsberger H, Abramowski D, Wiederhold KH, Haass C, Staufenbiel M, Konnerth A, Garaschuk O (2008) Clusters of hyperactive neurons near amyloid plaques in a mouse model of Alzheimer's disease. Science 321(5896):1686–1689. https://doi.org/10.1126/science.1162844

24. Çayir A, Yenidoğan I, Dağ H (2018) Feature extraction based on deep learning for some traditional machine learning methods. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 494-497). IEEE. https://doi.org/10.1109/UBMK.2018.8566383

25. Chakraborty S, Tomsett R, Raghavendra R, Harborne D, Alzantot M, Cerutti F, … Gurram P (2017) Interpretability of deep learning models: A survey of results. 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). doi:https://doi.org/10.1109/uic-atc.2017.8397411

26. Chen Z, Zhang T, Ouyang C (2018a) End-to-end airplane detection using transfer learning in remote sensing images. Remote Sens 10(1):139. https://doi.org/10.3390/rs10010139

27. Chen J, Yan S, Wong KC (2018b) Verbal aggression detection on twitter comments: convolutional neural network for short-text sentiment analysis. Neural Comput & Applic:1–10. https://doi.org/10.1007/s00521-018-3442-0

28. Chitradevi D, Prabha S (2020) Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease. Appl Soft Comput 86:105857. https://doi.org/10.1016/j.asoc.2019.105857

29. Choi BK, Madusanka N, Choi HK, So JH, Kim CH, Park HG, Bhattacharjee S, Prakash D (2020a) Convolutional neural network-based mr image analysis for Alzheimer's disease classification. Current Med Imag 16(1):27–35. https://doi.org/10.2174/1573405615666191021123854

30. Choi BK, Madusanka N, Choi HK, So JH, Kim CH, Park HG, Bhattacharjee S, Prakash D (2020b) Convolutional neural network-based mri image analysis for Alzheimer's disease classification. Current Med Imaging 16(1):27–35. https://doi.org/10.2174/1573405615666191021123854

31. Cui R, Liu M (2019) RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. Comput Med Imaging Graph 73:1–10. https://doi.org/10.1016/j.compmedimag.2019.01.005

32. de Marvao A, Dawes TJ, O'Regan DP (2020) Artificial intelligence for cardiac imaging-genetics research. Front Cardiovasc Med 6:195. https://doi.org/10.3389/fcvm.2019.00195
33. Deng L, Platt JC (2014) Ensemble deep learning for speech recognition. In Fifteenth annual conference of the international speech communication association
34. Denoising autoencoder (n.d.). https://paperswithcode.com/method/denoising-autoencoder
35. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, Lituiev D, Copeland TP, Aboian MS, Mari Aparici C, Behr SC, Flavell RR, Huang SY, Zalocusky KA, Nardo L, Seo Y, Hawkins RA, Hernandez Pampaloni M, Hadley D, Franc BL (2018) A Deep Learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. Radiology 180958:456–464. https://doi.org/10.1148/radiol.2018180958
36. DiPietro R, Hager GD (2020) Deep learning: RNNs and LSTM. In handbook of medical image computing and computer assisted intervention (pp. 503-519). Academic press. https://doi.org/10.1016/B978-0-12-816176-0.00026-0
37. Dong B, Wang X (2016) Comparison deep learning method to traditional methods using for network intrusion detection. In 2016 8th IEEE international conference on communication software and networks (ICCSN) (pp. 581-585). IEEE. https://doi.org/10.1109/ICCSN.2016.7586590
38. Dua M, Makhija D, Manasa PYL, Mishra P (2020) A CNN–RNN–LSTM Based amalgamation for Alzheimer's disease detection. J Med Biol Eng 40:688–706. https://doi.org/10.1007/s40846-020-00556-1
39. Dubois B, Picard G, Sarazin M (2009a) Early detection of Alzheimer's disease: new diagnostic criteria. Dialogues Clin Neurosci 11(2):135–139. https://doi.org/10.31887/DCNS.2009.11.2/bdubois
40. Dubois B, Picard G, Sarazin M (2009b) Early detection of Alzheimer's disease: new diagnostic criteria. Dialogues Clin Neurosci 11(2):135. https://doi.org/10.31887/DCNS.2009.11.2/bdubois
41. Ebrahim D, Ali-Eldin AM, Moustafa HE, Arafat H (2020) Alzheimer disease early detection using convolutional neural networks. In 2020 15th international conference on computer engineering and systems (ICCES) (pp. 1-6). IEEE. DOI: https://doi.org/10.1109/ICCES51560.2020.9334594.
42. Ebrahimighahnavieh MA, Chiong DR (2019) Deep Learning to detect Alzheimer's disease from neuro-imaging: a systematic literature Review. Comput Methods Prog Biomed 105242. https://doi.org/10.1016/j.cmpb.2019.105242
43. Esmaeilzadeh S, Belivanis DI, Pohl KM, Adeli E (2018) End-to-end Alzheimer's disease diagnosis and biomarker identification. In: International workshop on machine Learning in medical imaging, vol 11046. Springer, Cham, pp 337–345. https://doi.org/10.1007/978-3-030-00919-9_39
44. Farooq A, Anwar S, Awais M, Rehman S (2017) A deep CNN based multi-class classification of Alzheimer's disease using MRI. 2017 IEEE international conference on imaging systems and techniques (IST). doi:https://doi.org/10.1109/ist.2017.8261460
45. Feng W, Halm-Lutterodt NV, Tang H, Mecum A, Mesregah MK, Ma Y, Li H, Zhang F, Wu Z, Yao E, Guo X (2020) Automated MRI-Based Deep Learning model for detection of Alzheimer's disease process. Int J Neural Syst 30(06):2050032. https://doi.org/10.1142/s012906572050032x
46. Folego G, Weiler M, Casseb RF, Pires R, Rocha A (2020) Alzheimer's disease detection through whole-brain 3D-CNN MRI. Frontiers in Bioengineering and Biotechnology 8. https://doi.org/10.3389/fbioe.2020.534592
47. Fu'adah YN, Wijayanto I, Pratiwi NKC, Taliningsih FF, Rizal S, Pramudito MA (1844) Automated classification of Alzheimer's disease Based on MRI image Processing using convolutional neural network (CNN) with AlexNet architecture. J Phys Conf Ser 2021(1):012020. https://doi.org/10.1088/1742-6596/1844/1/012020
48. Glossary of Deep Learning: Autoencoder (2017) https://medium.com/deeper-learning/glossary-of-deep-learning-autoencoder-1044ec82c300
49. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018) A survey of methods for explaining black box models. ACM Comput Surveys (CSUR) 51(5):1–42. https://doi.org/10.1145/3236009
50. Gulhare KK, Shukla SP, Sharma LK (2017) Deep neural network classification method to Alzheimer's disease detection. Int J Adv Res Comput Sci Softw Eng 7(6):1–4
51. Guo K, Xu T, Kui X, Zhang R, Chi T (2019) iFusion: towards efficient intelligence fusion for deep learning from real-time and heterogeneous data. Information Fusion 51:215–223. https://doi.org/10.1016/j.inffus.2019.02.008
52. Gupta Y, Lee KH, Choi KY, Lee JJ, Kim BC, Kwon G-R (2019) Alzheimer's disease diagnosis Based on cortical and subcortical features. J Healthcare Eng 2019:1–13. https://doi.org/10.1155/2019/2492719
53. Han D, Liu Q, Fan W (2018) A new image classification method using CNN transfer learning and web data augmentation. Expert Syst Appl 95:43–56. https://doi.org/10.1016/j.eswa.2017.11.028
54. Hazarika RA, Kharkongor K, Sanyal S, Maji AK (2020) A comparative study on different skull stripping techniques from brain magnetic resonance imaging. In: In international conference on innovative

computing and communications. Springer, Singapore, pp 279–288. https://doi.org/10.1007/978-981-15-1286-5_24

55. Hazarika RA, Maji AK, Sur SN, Paul BS, Kandar D (2021) A survey on classification algorithms of brain images in Alzheimer's disease Based on feature extraction techniques. IEEE Access 9:58503–58536. https://doi.org/10.1109/access.2021.307255

56. Herrera LJ, Rojas I, Pomares H, Guillén A, Valenzuela O, Baños O (2013) Classification of MRI images for Alzheimer's disease detection. In 2013 international conference on social computing (pp. 846-851). IEEE. https://doi.org/10.1109/SocialCom.2013.127

57. Hori T, Chen Z, Erdogan H, Hershey JR, Le Roux J, Mitra V, Watanabe S (2017) Multi-microphone speech recognition integrating beamforming, robust feature extraction, and advanced DNN/RNN backend. Comput Speech Lang 46:401–418. https://doi.org/10.1016/j.csl.2017.01.013

58. Hosseini MP, Lu S, Kamaraj K, Slowikowski A, Venkatesh HC (2020) Deep learning architectures. In: Deep learning: concepts and architectures. Springer, Cham, pp 1–24. https://doi.org/10.1007/978-3-030-31756-0_1

59. How to Avoid Overfitting in Deep Learning Neural Networks (2019) https://machinelearningmastery.com/introduction-to-regularization-to-reduce-overfitting-and-improve-generalization-error/

60. Hu C, Ju R, Shen Y, Zhou P, Li Q (2016) Clinical decision support for Alzheimer's disease based on deep learning and brain network. In: Communications (ICC), 2016 IEEE international conference on, IEEE. pp. 1–6. https://doi.org/10.1109/ICC.2016.7510831.

61. Ibrahim A, Mohammed S, Ali HA, Hussein SE (2020a) Breast cancer segmentation from thermal images based on chaotic Salp swarm algorithm. IEEE Access 8:122121–122134. https://doi.org/10.1109/ACCESS.2020.3007336

62. Ibrahim A, Mohammed S, Ali HA, Hussein SE (2020b) Breast cancer segmentation from thermal images based on chaotic Salp swarm algorithm. IEEE Access 8:122121–122134. https://doi.org/10.1109/ACCESS.2020.3007336

63. Irankhah E (2020) Evaluation of early detection methods for Alzheimer's. Bioprocess Eng 4(1):17–22. https://doi.org/10.11648/j.be.20200401.13

64. Islam T, Manivannan D (2017) Predicting application failure in cloud: a machine Learning approach. 2017 IEEE international conference on cognitive computing (ICCC). https://doi.org/10.1109/ieee.iccc.2017.11

65. Jain A, Zamir AR, Savarese S, Saxena A (2016) Structural-rnn: Deep learning on spatio-temporal graphs. In proceedings of the ieee conference on computer vision and pattern recognition (pp. 5308-5317)

66. Jain R, Jain N, Aggarwal A, Jude Hemanth D (2019a) Convolutional neural network-based Alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 57:147–159. https://doi.org/10.1016/j.cogsys.2018.12.015

67. Jain R, Jain N, Aggarwal A, Jude Hemanth D (2019b) Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 57:147–159. https://doi.org/10.1016/j.cogsys.2018.12.015

68. JayaLakshmi ANM, Kishore KK (2018) Performance evaluation of DNN with other machine learning techniques in a cluster using apache spark and MLlib. J King Saud Univ-Comput Inform Sci. https://doi.org/10.1016/j.jksuci.2018.09.022

69. Jemwa GT, Aldrich C (2005) Improving process operations using support vector machines and decision trees. AICHE J 51(2):526–543. https://doi.org/10.1002/aic.10315

70. Jo T, Nho K, Saykin AJ (2019) Deep Learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci 11. https://doi.org/10.3389/fnagi.2019.00220

71. Johnson JM, Khoshgoftaar TM (2019) Survey on deep learning with class imbalance. J Big Data 6(1):1–54. https://doi.org/10.1186/s40537-019-0192-5

72. Johnson KA, Fox NC, Sperling RA, Klunk WE (2012a) Brain imaging in Alzheimer disease. Cold Spring Harbor Perspectives in Medicine 2(4):a006213–a006213. https://doi.org/10.1101/cshperspect.a006213

73. Johnson KA, Fox NC, Sperling RA, Klunk WE (2012b) Brain imaging in Alzheimer disease. Cold Spring Harbor Perspect Med 2(4):a006213. https://doi.org/10.1101/cshperspect.a006213

74. Kazemi Y, Houghten S (2018) A deep learning pipeline to classify different stages of Alzheimer's disease from fMRI data. 2018 IEEE conference on Computational Intelligence in bioinformatics and Computational biology (CIBCB). doi:https://doi.org/10.1109/cibcb.2018.8404980

75. Khademi A, Reiche B, DiGregorio J, Arezza G, Moody AR (2019) Whole volume brain extraction for multi-Centre. Multi-Disease FLAIR MRI Datasets Magnetic Resonance Imaging 66:116–130. https://doi.org/10.1016/j.mri.2019.08.022

76. Khagi B, Kwon GR (2020) 3D CNN design for the classification of Alzheimer's disease using brain MRI and PET. IEEE Access 8:217830–217847. https://doi.org/10.1109/ACCESS.2020.3040486

77. Khagi B, Kwon GR, Lama R (2019) Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques. Int J Imaging Syst Technol 29(3): 297–310. https://doi.org/10.1002/ima.22316

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

79. Kundaram SS, Pathak KC (2020) Deep Learning-Based Alzheimer disease detection. Proceedings of the fourth international conference on microelectronics, computing and communication systems, 587–597. https://doi.org/10.1007/978-981-15-5546-6_50

80. Labeeb YA, Morsy M, Abo-Elsoud MEA (2015) Preprocessing technique for enhancing the DICOM kidney images. Int J Eng Res Technol (Ijert) 04(08):2278–0181

81. Lechner M, Hasani R (2020) Learning long-term dependencies in irregularly-sampled time series. Advances in neural information Processing systems, 33. arXiv:2006.04418v4

82. Lee G, Nho K, Kang B, Sohn K-A, Kim D (2019) Predicting Alzheimer's disease progression using multi-modal deep learning approach. Sci Rep 9(1):1952. https://doi.org/10.1038/s41598-018-37769-z

83. Li F, Liu M (2019) A hybrid convolutional and recurrent neural network for Hippocampus analysis in Alzheimer's disease. J Neurosci Methods 323:108–118. https://doi.org/10.1016/j.jneumeth.2019.05.006

84. Li H, Yu L, Tian S, Li L, Wang M, Lu X (2017) Deep learning in pharmacy: the prediction of aqueous solubility based on deep belief network. Autom Control Comput Sci 51(2):97–107. https://doi.org/10.3103/s0146411617020043

85. Lipton ZC (2018) The mythos of model interpretability: in machine learning, the concept of interpretability is both important and slippery. Queue 16(3):31–57

86. Liu F, Wee C-Y, Chen H, Shen D (2013) Inter-modality relationship constrained multi-task feature selection for AD/MCI classification. Lecture Notes Comput Sci:308–315. doi:https://doi.org/10.1007/978-3-642-40811-3_39

87. Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D (2014) Early diagnosis of Alzheimer's disease with deep learning," 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, pp. 1015–1018, DOI: https://doi.org/10.1109/ISBI.2014.6868045.

88. Lu H, Li Y, Chen M, Kim H, Serikawa S (2018) Brain intelligence: go beyond artificial intelligence. Mobile Networks Appl 23(2):368–375. https://doi.org/10.1007/s11036-017-0932-8

89. Machine learning and deep learning (2020) https://www.zendesk.com/blog/machine-learning-and-deep-learning/

90. Mehmood A, Maqsood M, Bashir M, Shuyuan Y (2020) A Deep Siamese convolution neural network for multi-class classification of Alzheimer disease. Brain Sci 10(2):84. https://doi.org/10.3390/brainsci10020084

91. Ming Y, Cao S, Zhang R, Li Z, Chen Y, Song Y, Qu H (2017) Understanding hidden memories of recurrent neural networks. In 2017 IEEE conference on visual analytics science and technology (VAST) (pp. 13-24). IEEE. https://doi.org/10.1109/VAST.2017.8585721

92. Morabito FC, Campolo M, Ieracitano C, Mammone N (2019) Deep Learning approaches to electrophysiological multivariate time-series analysis. Artificial Intell Age Neural Networks Brain Comput:219–243. https://doi.org/10.1016/b978-0-12-815480-9.00011-6

93. Murugan P, Durairaj S (2017) Regularization and optimization strategies in deep convolutional neural network. arXiv preprint arXiv:1712.04711

94. Nawaz H, Maqsood M, Afzal S, Aadil F, Mehmood I, Rho S (2020) A deep feature-based real-time system for Alzheimer disease stage detection. Multimed Tools Appl 80:35789–35807. https://doi.org/10.1007/s11042-020-09087-y

95. Ng YS, Xue W, Wang W, Qi P (2019) Convolutional neural networks for food image recognition: An experimental study. In proceedings of the 5th international workshop on multimedia assisted dietary management (pp. 33-41). https://doi.org/10.1145/3347448.3357168

96. Nicholson C (2019) A beginner's guide to lstms and recurrent neural networks. Skymind. Saatavissa: https://wiki.pathmind.com/lstm. Hakupäivä, 6, 2019

97. Noda K, Yamaguchi Y, Nakadai K, Okuno HG, Ogata T (2014) Audio-visual speech recognition using deep learning. Appl Intell 42(4):722–737. https://doi.org/10.1007/s10489-014-0629-7

98. Noor MBT, Zeina NZ, Kaiser MS, Al Mamun S, Mahmud M (2020) Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. Brain Informatics 7:11. https://doi.org/10.1186/s40708-020-00112-2

99. O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Hernandez GV, Krpalkova L, … Walsh J (2019) Deep learning vs. traditional computer vision. In: Science and Information Conference. Springer, Cham, pp 128–144. https://doi.org/10.1007/978-3-030-17795-9_10

100. Otter DW, Medina JR, Kalita JK (2020) A survey of the usages of Deep Learning for Natural Language Processing. IEEE Transactions on Neural Networks and Learning Systems 32:1–21. https://doi.org/10.1109/tnnls.2020.2979670

101. Pak M, Kim S (2017) A review of deep learning in image recognition. 2017 4th international conference on computer applications and information Processing technology (CAIPT). doi:https://doi.org/10.1109/caipt.2017.8320684

102. Panza F, Frisardi V, Capurso C, D'Introno A, Colacicco AM, Imbimbo BP, Santamato A, Vendemiale G, Seripa D, Pilotto A, Capurso A, Solfrizzi V (2010) Late-life depression, mild cognitive impairment, and dementia: possible continuum? Am J Geriatr Psychiatry 18(2):98–116. https://doi.org/10.1097/JGP.0b013e3181b0fa13

103. Pathirage CSN, Li J, Li L, Hao H, Liu W, Ni P (2018) Structural damage identification based on autoencoder neural networks and deep learning. Eng Struct 172:13–28. https://doi.org/10.1016/j.engstruct.2018.05.109

104. Paudel YN, Angelopoulou E, Piperi C, Othman I, Aamir K, Shaikh M (2020) Impact of HMGB1, RAGE, and TLR4 in Alzheimer's disease (AD): from risk factors to therapeutic targeting. Cells 9(2):383. https://doi.org/10.3390/cells9020383

105. Peng J, Guan J, Shang X (2019) Predicting Parkinson's disease genes based on node2vec and autoencoder. Front Genet 10:226. https://doi.org/10.3389/fgene.2019.00226

106. Pereira L, Pinto M, Shah K, Khan S (2020) Show and tell: a neural visual story-teller. Available at SSRN 3565282. https://doi.org/10.2139/ssrn.3565282

107. Perumal S, Velmurugan T (2018) Preprocessing by contrast enhancement techniques for medical images. Int J Pure Appl Mathematics 118(18):3681–3688

108. Pichler BJ, Judenhofer MS, Pfannenberg C (2008) Multimodal imaging approaches: pet/ct and pet/mri. Molecular Imaging I:109–132. https://doi.org/10.1007/978-3-540-72718-7_6

109. Pierce AL, Bullain SS, Kawas CH (2017) Late-onset Alzheimer disease. Neurol Clin 35(2):283–293. https://doi.org/10.1016/j.ncl.2017.01.006

110. Qiu F, Berglund J, Jensen JR, Thakkar P, Ren D (2004) Speckle noise reduction in SAR imagery using a local adaptive median filter. GIScience & Remote Sens 41(3):244–266. https://doi.org/10.2747/1548-1603.41.3.244

111. Rajeshwari S, Sharmila TS (2013) Efficient quality analysis of MRI image using preprocessing techniques. IEEE Conference Inform Commun Technol. https://doi.org/10.1109/cict.2013.6558127

112. Ramzan F, Khan MUG, Rehmat A, Iqbal S, Saba T, Rehman A, Mehmood Z (2020) A Deep Learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. J Med Syst 44:37. https://doi.org/10.1007/s10916-019-1475-2

113. Random Oversampling and Undersampling for Imbalanced Classification. (2021) https://machinelearningmastery.com/random-oversampling-and-undersampling-for-imbalanced-classification/

114. Recurrent neural networks (2020). https://www.ibm.com/cloud/learn/recurrent-neural-networks

115. Sajedi H, Pardakhti N (2019) Age prediction Based on brain MRI image: a survey. J Med Syst 43(8):279. https://doi.org/10.1007/s10916-019-1401-7

116. Salehi AW, Baglat P, Sharma BB, Gupta G, Upadhya A (2020) A CNN model: earlier diagnosis and classification of Alzheimer disease using MRI. 2020 international conference on smart electronics and communication (ICOSEC). doi:https://doi.org/10.1109/icosec49089.2020.9215402

117. Sandoub G, Atta R, Ali HA, Abdel-Kader RF (2021) A low-light image enhancement method based on bright channel prior and maximum colour channel, IET Image Processing. Institution Eng Technol 118(18):3681–3688

118. Sarraf S, Tofighi G (2016a) Classification of Alzheimer's disease structural MRI data by deep learning convolutional neural networks. arXiv preprint arXiv:1607.06583

119. Sarraf S, Tofighi G (2016b) Classification of Alzheimer's disease using fmri data and deep learning convolutional neural networks. arXiv preprint. https://arxiv.org/abs/1603.08631

120. Shah M, Xiao Y, Subbanna N, Francis S, Arnold DL, Collins DL, Arbel T (2011) Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. Med Image Anal 15(2):267–282. https://doi.org/10.1016/j.media.2010.12.003

121. Shakarami A, Tarrah H, Mahdavi-Hormat A (2020) A CAD system for diagnosing Alzheimer's disease using 2D slices and an improved AlexNet-SVM method. Optik, 164237. doi:https://doi.org/10.1016/j.ijleo.2020.164237

122. Shankar GM, Walsh DM (2009) Alzheimer's disease: synaptic dysfunction and Aβ. Mol Neurodegener 4(1):48. https://doi.org/10.1186/1750-1326-4-48

123. Sharma S, Sharma S (2017) Activation functions in neural networks. Towards Data Sci 6(12):310–316

124. Shen C, Qi Y, Wang J, Cai G, Zhu Z (2018) An automatic and robust features learning method for rotating machinery fault diagnosis based on contractive autoencoder. Eng Appl Artif Intell 76:170–184. https://doi.org/10.1016/j.engappai.2018.09.010

125. Shi J, Zheng X, Li Y, Zhang Q, Ying S (2017) Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE J Biomed Health Inform 22:173–183. https://doi.org/10.1109/JBHI.2017.2655720

126. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48. https://doi.org/10.1186/s40537-019-0197-0

127. Shrestha S (2014) Image denoising using new adaptive based median filters. arXiv preprint arXiv: 1410.2175

128. Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyás B (2020) 3D deep learning on medical images: a review. Sensors 20(18):5097. https://doi.org/10.3390/s20185097

129. Spasov S, Passamonti L, Duggento A, Liò P, Toschi N (2019) A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. NeuroImage 189:276–287. https://doi.org/10.1016/j.neuroimage.2019.01.031

130. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Machine Learn Res 15(1):1929–1958

131. Suk HI, Lee SW, Shen D (2016) A. S. D. N. initiative. Deep sparse multi-task learning for feature selection in Alzheimer's disease diagnosis. Brain Struct Funct 221(5):2569–2587. https://doi.org/10.1007/s00429-015-1059-y

132. Suk H-I, Lee S-W, Shen D (2017) A S D N Initiative Deep ensemble learning of sparse regression models for brain disease diagnosis. Med Image Anal 37:101–113. https://doi.org/10.1016/j.media.2017.01.008

133. Sun X, Shi L, Luo Y, Yang W, Li H, Liang P, Li K, Mok VCT, Chu WCW, Wang D (2015) Histogram-based normalization technique on human brain magnetic resonance images from different acquisitions. Biomed Eng Online 14(1):1–17. https://doi.org/10.1186/s12938-015-0064-y

134. Sun W, Shao S, Zhao R, Yan R, Zhang X, Chen X (2016) A sparse auto-encoder-based deep neural network approach for induction motor faults classification. Measurement 89:171–178. https://doi.org/10.1016/j.measurement.2016.04.007

135. Sun C, Ma M, Zhao Z, Tian S, Yan R, Chen X (2018) Deep transfer Learning Based on sparse auto-encoder for remaining useful life prediction of tool in manufacturing. IEEE transactions on industrial informatics, 1–1. doi:https://doi.org/10.1109/tii.2018.2881543

136. Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M, Lin CT (2020) Machine Learning techniques for the diagnosis of Alzheimer's disease. ACM Trans Multimed Comput Commun Appl 16(1s):1–35. https://doi.org/10.1145/3344998

137. Tao S, Zhang T, Yang J, Wang X, Lu W (2015) Bearing fault diagnosis method based on stacked autoencoder and softmax regression. In: Control conference (CCC), 2015 34th Chinese. IEEE. https://doi.org/10.1109/ChiCC.2015.7260634

138. Theodore WH, Dorwart R, Holmes M, Porter RJ, DiChiro G (1986) Neuroimaging in refractory partial seizures: comparison of PET, CT, and MRI. Neurology 36(6):750–759. https://doi.org/10.1212/WNL.36.6.750

139. Trambaiolli LR, Lorena AC, Fraga FJ, Kanda PAM, Anghinah R, Nitrini R (2011) Improving Alzheimer's disease diagnosis with machine Learning techniques. Clin EEG Neurosci 42(3):160–165. https://doi.org/10.1177/155005941104200304

140. Transfer learning and the art of using Pre-trained Models in Deep Learning. (2017) https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model/

141. Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, … Jack CR (2009) MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology 73(4):294–301. https://doi.org/10.1212/WNL.0b013e3181af79fb

142. Venugopalan J, Tong L, Hassanzadeh HR, Wang MD (2021) Multimodal deep learning models for early detection of Alzheimer's disease stage. Sci Rep 11(1):1–13. https://doi.org/10.1038/s41598-020-74399-w

143. Vu TD, Yang HJ, Nguyen VQ et al (2017) Multimodal learning using convolution neural network and sparse autoencoder. BigComp 2017, Jeju, pp 309–312. https://doi.org/10.1109/BIGCOMP.2017.7881683

144. Wang F, Casalino LP, Khullar D (2018a) Deep Learning in medicine—promise, Progress, and challenges. JAMA Intern Med. https://doi.org/10.1001/jamainternmed.2018

145. Wang Y, Yang Y, Guo X, Ye C, Gao N, Fang Y, Ma HT (2018b) A novel multimodal MRI analysis for Alzheimer's disease Based on convolutional neural network. 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC). doi:https://doi.org/10.1109/embc.2018.8512372

146. Wei J, Chu X, Sun XY, Xu K, Deng HX, Chen J, Wei Z, Lei M (2019) Machine learning in materials science. InfoMat 1(3):338–358. https://doi.org/10.1002/inf2.12028

147. Wen L, Gao L, Li X (2017) A new Deep transfer Learning Based on sparse auto-encoder for fault diagnosis. IEEE transactions on systems, man, and cybernetics: systems, 1–9. doi:https://doi.org/10.1109/tsmc.2017.2754287

148. Wu M, Chen L (2015) Image recognition based on deep learning. In 2015 Chinese automation congress (CAC) (pp. 542-546). IEEE. https://doi.org/10.1109/cac.2015.7382560

149. Wu O, Winzeck S, Giese AK, Hancock BL, Etherton MR, Bouts MJ, … MRI-GENIE and GISCOME Investigators (2019) Big data approaches to phenotyping acute ischemic stroke using automated lesion segmentation of multi-center magnetic resonance imaging data. Stroke 50(7):1734–1741. https://doi.org/10.1161/STROKEAHA.119.025373

150. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629. https://doi.org/10.1007/s13244-018-0639-9

151. Yanase J, Triantaphyllou E (2019) A systematic survey of computer-aided diagnosis in medicine: past and present developments. Expert Syst Appl 138:112821. https://doi.org/10.1016/j.eswa.2019.112821

152. Yildirim M, Cinar A (2020a) Classification of Alzheimer's disease MRI images with CNN Based hybrid method. Journal homepage: http://iieta.Org/journals/isi, 25(4), 413-418. https://doi.org/10.18280/isi.250402.

153. Yildirim M, Cinar A (2020b) Classification of Alzheimer's disease MRI images with CNN Based hybrid method. J homepage: http://iieta.Org/journals/isi, 25(4), 413-418. https://doi.org/10.18280/isi.250402

154. Yin W, Kann K, Yu M, Schütze H (2017) Comparative study of CNN and RNN for natural language processing. arXiv preprint arXiv:1702.01923

155. Young T, Hazarika D, Poria S, Cambria E (2018) Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. IEEE Computational Intell Mag 13(3):55–75. https://doi.org/10.1109/mci.2018.2840738

156. Zhang L, Wang M, Liu M, Zhang D (2020a) A Survey on Deep Learning for Neuroimaging-Based Brain Disorder Analysis. Front Neurosci, 14. https://doi.org/10.3389/fnins.2020.00779

157. Zhang Y, Zhang H, Adeli E, Chen X, Liu M, Shen D (2020b) Multiview feature Learning with multiatlas-Based functional connectivity networks for MCI diagnosis. IEEE Trans Cybernetics PP:1–12. https://doi.org/10.1109/TCYB.2020.3016953

158. Zhao X, Zhao X-M (2020) Deep learning of brain magnetic resonance images: a brief review. Methods 192:131–140. https://doi.org/10.1016/j.ymeth.2020.09.007

159. Zotin A, Simonov K, Kapsargin F, Cherepanova T, Kruglyakov A, Cadena L (2018) Techniques for medical images processing using shearlet transform and color coding. In: Computer vision in control Systems-4. Springer, Cham, pp 223–259. https://doi.org/10.1007/978-3-319-67994-5_9
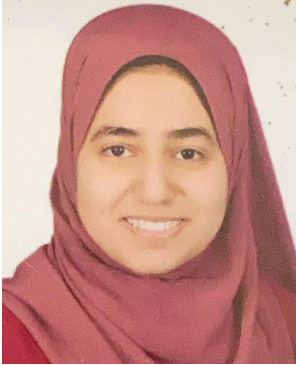
Doaa Ahmed Arafa , Teaching Assistant at the Department of computer engineering and control systems, Faulty of engineering, Mansoura University.

Hossam El-Din Moustafa , Associate Professor at the Department of Electronics and Communications Engineering, the founder and executive manager of Biomedical Engineering Program at the Faculty of Engineering, Mansoura University.

Amr M. T. Ali-Eldin , Professor in computer engineering and control systems department at the Faculty of Engineering, Mansoura University.

Hesham A. Ali , He is currently a Professor in computer engineering and control system and an Associate Professor in information system and computer engineering, University of Mansoura.

Affiliations

Doaa Ahmed Arafa [1] · **Hossam El-Din Moustafa** [2] · **Amr M. T. Ali-Eldin** [1] · **Hesham A. Ali** [1]

[1]   Computer Engineering and Control Systems Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt

[2]   Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt