REGULAR PAPER

Transfer learning using freeze features for Alzheimer neurological disorder detection using ADNI dataset

Saeeda Naz[1] · Abida Ashraf[1] · Ahmad Zaib[2,3]

Abstract

Machine learning and deep learning play a crucial role in identification of various diseases like neurological, skin, eyes, blood and cancers. The deep learning algorithms can be performed promising for prediction of Alzheimer's disease using MRI scans. Alzheimer disease becoming more common in the people (age 65 years or above). The disease becomes severe before the symptoms appear and causes brain disorder that cannot be cured by medicines and other therapies and treatments. So the early diagnosis is necessary to slow down its progression. Detection and prevention of Alzheimer disease is one of the active research area for the researchers nowadays. In this paper, we employed architectures of convoutional networks using freeze features extracted from source data set ImageNet for binary and ternary classification. All experiments were carried out using Alzheimer's disease national initiative (ADNI) data set consisting of MRI scans. The performance of proposed system demonstrates for classification of Alzheimer's disease versus mild cognitive impairment, normal controls versus mild cognitive impairment, and cognitive normal versus Alzheimer's disease. The results of proposed study show that VGG architecture outperforms the state-of-the-art techniques and number of architectures of conveNet (AlexNet, GoogLeNet, ResNet, DenseNet, Inceptionv3, InceptionResNet) in Alzheimer's disease detection, and achieves an identification test set accuracy of 99.27% (MCI/AD), 98.89% (AD/CN) and 97.06% (MCI/CN).

Keywords Alzheimer disease · Transfer learning ·

1 Introduction

Alzheimer disease (AD) is a neurological disorder and a most common form of dementia affecting a large population of elderly people in the world, e.g., AD is the sixth leading cause in the United State of America (USA) for deaths [1]. In this disease, the capability of mental health decreases and patient cannot lead normal life and survive without help, support and assistance of their family members. In 2006, Brookmeyer et al. [2] conducted a study showing that 26.6 million people are affected worldwide by Alzheimer disease.

A neuropathologist and German psychiatrist *Alois Alzheimer* identified and named the AD for the first time in 1906.

Alzheimer disease causes due to environmental and genetic factors, head injuries, and by chemicals. Common symptoms of AD are memory loss, cognition problems, recognition and communication problems, mood, and behavior disruption. Alzheimer disease leads death of brain cells that causes loss of memory, thinking and cognitive power. It has a progressive course and marked as pre-clinical period. The speed of progression of this disease can differ from patient to patient but it has a poor prognosis. It causes behavioral disturbance which impairs social functioning of patient. The typical onset and symptoms of this disease occurs after the age of 65 [3] but it may start before decades and the symptoms have not been observed before this age. Alzheimer disease is increasing day by day and report regarding AD concluded that Alzheimer's patients will become double in the next 20 years [2]. Alzheimer disease mainly affects the elderly people and leads to death if not detected and treated at the early stage. Therefore, it becomes necessary to detect

✉ Saeeda Naz
saeedanaz292@gmail.com

1 Computer Science Department, Government Girls Postgraduate College No. 1, Abbottabad, Pakistan

2 Benazir Bhutto Shaheed Women, Children District Head Quarter Hospital, Abbottabad, Pakistan

3 Women Medical College, Abbottabad, Pakistan

the disease in the earliest stage so that progression of disease could be slow down.

The pathophysiology of AD disease involves the alteration of structure of neurons. Neuron has internal structure called micro-tubules. These micro-tubules act like tracts which guide nutrients and molecules to the axons. A protein called TAU protein is used to stabilizes them. In AD patients, TAU protein is changed chemically and it tends to pair with other TAU proteins. This protein then collapses the neurons and it result in formation of Neurofibrillary tangles. It causes first malfunction and then later on death of cells. In addition of NFTs, it also includes senile plaques called beta amyloid plaques at the microscopic and cerebrocortical atrophy at the macroscopic level. Plaques are the dense insoluble protein deposits in the surrounding of the neurons. These consist of fragment of "beta amyloid protein" and formed by the lager protein named as "amyloid precursor protein". These fragments join together that results in formations of plaques and alteration of structure [4]. These disturb the transfer of information through the cells and plaque forms as nerve endings of cells decline, preventing messages from travelling from one cell to another cell.

The hippocampus and medial temporal lobe are the first part of brain that is to be involved and it can be detected on the MRI. Hippocampus is the structure located deep in the brain and involved in memories processing and connects to brain structure involved in thinking and decisions making. It is also characterized by loss of neurons and synapses in the cortex, temporal lobe parietal lobe and also involves typically cingulate gyrus. There is gross atrophy of brain and it is shown in Fig. 1.

The diagnosis of Alzheimer disease is mainly clinical. The doctors rely on clinical evaluations like cognitive testing, neurological examination and medical history, but imaging studies are conducted to rule out other possible causes of dementia which includes cerebrovascular disease, vitamin B12 deficiency, syphilis, and thyroid disease and many others. In addition to these, different tests are also conducted like blood test, CT, MRI (Magnetic resonance Imaging) and PET (Positron Emission Tomography) scans, EEG (Electroencephalography) and Genotyping. American Academy of Neurology recommends that structural imaging of the brain either contrasts or non contrast CT scan or MRI is the most appropriate initial evaluation in cases of patients of dementia. Imaging studies help us to rule out curable causes of progressive cognitive deficit which includes normal pressure hydrocephalus, etc. The MRI and CT scan shows diffuse cortical and cerebral atrophy in patients with Alzheimer disease. Atrophy of the hippocampus is important and considered a valid biomarker in patients with Alzheimer disease.

The rapid growth and advances of new technologies and recent development of deep learning have a great effect on our every field of life and have many applications like pattern
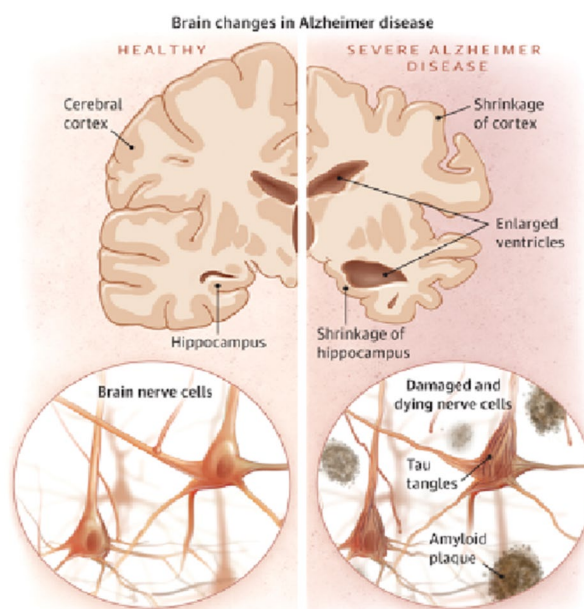

Fig. 1 Normal and Alzheimer brain [5]

recognition [6–8], image processing [9, 10], automated driving, computer vision and medical imaging [11, 12]. The diagnosis of different diseases in hospitals can be improved using computer aided algorithms in terms of efficiency and time. Many different machine learning techniques were used by different researchers to classify AD using Neuroimaging data. Different conventional methods like support vector machine (SVM) and different deep and transfer learning methods like convolutional neural network, recurrent neural network etc. were used by different researchers to automatically detect AD with less effort of radiologists.

As there is slightly difference between different stages of Alzheimer Disease and it is very difficult to identify which requires careful clinical Assessment and careful observation of radiologists which is very cost effective so motivation of our proposed study is to diagnose AD with the help of transfer learning and number of deep CNN architectures that the progression of disease could be slow down at early stage .

According to limitation of traditional models in feature extraction and classifications in different fields, we propose a deep convolutional neural network (CNN)-based Alzheimer disease prediction system to extract salient visual features from MRI scans. These features can be used to discriminate among Alzheimer Disease (AD), Mild Cognitive Impairment (MCI) and Normal Controls (CN) and predict AD in the early stage.

The key contributions of this study are:

– Different architectures of CNN are employed to examine and investigate that which architecture is able to extract

the necessary visual features using MRI images of in characterizing AD, MCI and CN.

- With assistance to transfer learning, multiple data augmentation schemes use for increasing and enhancing the input space using raw images for extraction of salient features.
- Comprehensive comparative study of CNNs using freeze features from source data set to transfer the knowledge for identification AD in MRI images.
- Investigating the impact of transfer leaning across two different domains using twelve architectures of CNN and providing experimental results of maximum number of CNNs in this study as a baseline for researchers in the field of AD detection.

The rest of the paper is structured as follow. Section 2 explains the closely related work using . Section 4.1 covers the details of . Sections 3 and 4 present the different CNN architecture and proposed study. In Sects. 5 and 6, the experimental setup, results and discussion are described and finally, the Sect. 7 summarised the proposed study.

2 Related work

Alzheimer disease (AD) is a common neurodegenerative disorder and is very necessary to diagnose it at very early stage, therefore, different models and methods introduced by the research community to diagnose Alzheimer's disease. in this section, we review different deep learning-based methods employed for AD identification.

In 2014, Suk et al. [13] applied multi model fusion using hierarchical feature representation for AD detection. The proposed system evaluated on selected subset of MRI and PET scans having total 398 subjects from dataset. Th subset consists of 93 subjects of AD, 104 of MCI and 101 of Normal Controls. They developed patch level feature learning model called Multi Model DBM (MM DBM) using tissue Densities of MRI patch and voxel intensities of PET patch then they used train Restricted Boltzmann Machine (RBM) as a preprocessor which can transform real values observations into binary form. The extracted features passed to the multilevel classifier to classify AD vs. CN and MCI vs. CN and achieved accuracy, sensitivity and specificity of 95.35%, 94.65%, 95.22% and 85.67%, 95.37%, 65.87%, respectively.

Saman et al. [14] implemented Le-Net for detection of AD from healthy controls using records of 28 AD patients and 15 healthy persons from . They reported the system accuracy of 96.85%. Mathew et al. [15] used a subset of dataset for experiments having 151 MRIs of patients including 71 normal controls and 87 AD patients. The PCA (Principal Component Analysis) and DWT (Discrete Wavelet Transform) applied for feature extraction and fed the features

to SVM (Support Vector Machine) for classification. They achieved accuracy of 84% and 91% for AD vs. CN and MCI vs. CN, respectively.

In 2016, Iftikhar et al. [16] presented an Ensemble classification approach for Alzheimer disease and MCI. Firstly, they deployed the cortical thickness and volumetric cortex-based features to detect AD. Secondly, they put forward the data to the ensembles for classification using total 180 subjects having 60 AD patients, 60 MCI and 60 normal controls and achieved accuracy of 91.66% sensitivity of 92% and specificity of 89% for AD vs. MCI. For vs. CN and CN vs. MCI. Asl et al. [17] proposed a deep 3D convolutional neural network (3D-CNN) to diagnose the AD. The experiments were carried out on MRI including 70 AD, 70 MCI and 70 CN data obtained from . Local features were extracted from 3D input image using CAE (Convolutional Auto Encoder). Model was trained on CAD Dementia data set which contains T1 weighted MRIs of AD, CN and CN. Skull stripping, spatially normalization were performed as pre-processing. Features extracted from CADDementia were used as biomarker to detect the AD on dataset using fine tune approach. Classification was performed using 10 fold cross validation technique and achieved 97.6% of accuracy for AD vs. CN classification.

In 2017, Ju et al. [18] used MRI and textual data (age, gender, and genetic information) to diagnose the Alzheimer using deep neural network. MRI images of 91 Mild Cognitive Impairment (MCI) and 79 Normal Controls were acquired from ADNI-2 data set along with their genetic information and used this genetic information to find the prevalence between MCI and age, gender and ApoE. They used Data Processing and Analysis of Brain Imaging (DPABI) for pre-processing. They Rf-MRI time series data and correlation coefficient data as input to the LDR, LR, SVM and autoencoder network and concluded that test accuracy increases for correlation coefficient data. The accuracy, sensitivity and specificity obtained by LDR are 67.72%, 65%, 66% and by LR are 71.38%, 77%, 62%. The accuracy, sensitivity and specificity obtained by SVM are 78.91%, 79%, 64%. The accuracy, sensitivity and specificity obtained by autoencoder are 86.47%, 92%, 81%. All these calculations are achieved using correlation coefficient data.

Farooq et al. [19] proposed deep learning GoogLeNet, ResNet-18 and ResNet-152 for multiclass classification of AD. Experiments were conducted using with four classes named AD, LMCI, MCI and CN having MRI's of 33,22,449 and 45, respectively. They obtained accuracy of 98.88%, 98.01% using ResNet-18 and 98.14% using ResNet-152. Backstorm et al. [20] employed 3D-ConvNet using MRIs of brain. In preprocessing step, they applied cortical reconstruction, trim edges, image resize and intensity normalization and extracted automated features. Data was gathered from dataset and conducted

experiments on 340 subjects which includes 1190 MRI scans of 199 AD patients (103 male, 96bfemale) and 141 normal controls (75 male, 66 female) to detect Alzheimer and Normal persons and obtained 98.78% accuracy on test data set when data is randomly portioned into 60%, 20% and 20% for training, testing and validation, respectively.

Kazemi and Houghten [21] used fMRI data from ADNI to classify different stages of Alzheimer disease. They gathered data of 197 subjects of five classes named as AD, CN, SMC, EMCI and LMCI including 107 females and 90 males. Brain extraction, slice timing correction, spatial smoothing, high pass filtering, spatial normalization and image conversion were used as pre-processing. A deep CNN classifier, AlexNet was used for classification with five-fold cross validation technique and split of 60% for training, 20% for testing and 20% for validation was used for each experiments. They achieved average accuracy of 97.63% and accuracy per class is 94.97%, 95.64%, 95.89%, 98.34% and 94.55% for AD, EMCI, LMCI, CN and SMC, respectively. In [22], Qiu et al. used methods of MMSE and CNN for identification of MRI images of MCI vs. CN with recognition rate up to 90.9%. Lin et al. [23] implemented CNN and achieved 88.79% classification rate for AD and CN. Payan et al. [24] applied CNN on 2265 MRI scans of . The proposed method evaluated for identification of AD vs. CN, AD vs. MCI, CN vs. MCI and AD vs. CN vs. MCI and reported accuracy up to 95.39%, 82.24%, 90.13% and 85.53%, respectively. Xia et al. [25] employed 3D CLSTM for extraction of deep salient features and recognition of 94.19% for AD using AD (198), CN 229 and MCI (408) from dataset.

In 2019, Ghahnavieh et al. [26] used transfer leaning to detect AD using MRI's from . MRI's of 132 subject for each AD and CN were used for conducting experiments. They used recurrent network with convolutional neural network to better understand the relationship between sequence of input images. They extracted features using convolutional neural network and then train recurrent neural network and improved accuracy. In 2021, Ashraf et al. [27] realised the fine-tuned features using different CNN architectures using and reported 99.05% recognition rate. In [28], Tissue segmentation was applied on each subject to extract the gray matter tissue extracted using tissue segmentation approach and then classy binary class using VGG family to achieve 98.73% recognition rate for AD vs. NC. Liu et al. [29] deployed AlexNet and GoogLeNet and got 91.40% and 93.02% with minimum power consumption.

Chen and Xia [30] classified AD vs. CN and MCI vs. AD using deep features and then applied sparse regression module.

3 Deep neural networks

Machine learning (ML) algorithms belong to the area of Artificial Intelligence (AI), in which computer learns relationship among data and made decisions without explicit knowledge and expertise. In the late 1990s, most of ML algorithms have been developed to surpass the human abilities in various fields of life especially in speech and vision but they did not achieve satisfactory performance. This challenging machine vision nature gives rise to a new class of neural networks (NN) which is inspired by biological structure of human brain. These networks are called Convolutional Neural Network (CNN). Deep Convolutional Neural Networks are a type of network that performs well for object recognition, classification, detection and segmentation and have shown state of art performance on various benchmarks. Deep CNN achieved its powerful learning ability with the use of multiple feature extraction strategies that can automatically learn discriminative features. Advancement in hardware processing units and availability of large amount of data has extended CNN's research and, therefore, very deep, different and interesting CNN architectures were developed by researchers. The distinct characteristics of CNN are its hierarchical structure. CNN structure composed of combination of Convolution Layer, Pooling Layer, Activation function, Normalization Layer and Fully-Connected Layer. Convolution layer also called CONV, consists of set of kernels computes the output by dividing the image into small blocks and convolving them with weights. Briefly, CONV layers extract features from input images. The operation of convolution is shown in Eq. (1)

$$F_k^l = \left(I_x, y \times K_l^k\right), \tag{1}$$

I_x, y represent input image, where x, y represent spatial locality, K_l^k represents l^th convolutional kernel of the kth layer. Pooling Layer also called POOL, performs down sampling operation by applying summation on the similar information of neighborhood of receptive field and outputs the dominant response from input regions.

$$Z_l = f_p\left(F_x^l, y\right), \tag{2}$$

Lth output feature map is represented by Z_l, input feature map is represented by F_x^l, y, whereas $f_p(.)$ represents the pooling operation. Pooling layer is used to extract combination of features and to reduce over fitting. For extracting invariant features different types of pooling operations like max, L_2, average and overlapping pooling are used. Element wise activation function is applied in normalization layer. Activation function helps in learning complicated patterns and serves as a decision maker. Different activation

functions like RELU, sigmoid, Thanh, max out and variants of RELU are used. The RELU and its invariant are most commonly used, because it helps to reducing gradient problem. Equation (3) defines activation function

$$T_l^k = f_a(F_l^k),\tag{3}$$

f_a is an activation function which is applied to out of a convolution function F_l^k that add non linearity and T_l^k represent transformed output for kth layer. In this layer, size of input volume remains same. Fully-Connected Layer is used at the end of network as a classifier to compute class scores. Recently, researchers developed different CNN architectures such as LeNet, AlexNet, GoogLeNet, VGG, ResNet, InceptionV3, Inception-ResNetV2, mobileNetV2, DenseNet and many more.

4 Proposed study design

In this section, the framework explains for computerized Alzheimer disease detection using freeze approach of transfer learning. Number of CNNs architectures are employed to detect the neuro-psychological deformations present in MRI images. We works on T1 weighted MRI brain scans from a benchmark dataset. we evaluate eleven pre-trained deep CNN models named as AlexNet, GoogLeNet, VGG-16/19, ResNet-18/50/101, MobileNetV2, InceptionV3, Inception-ResNet-V2 and DenseNet201. We extracted features from fully connected layers of CNNs. We conducted three scenarios, i.e., Scenarios 1: MCI vs. AD, Scenarios 2: AD vs. CN, and Scenarios 3: MCI vs. CN. Figure 2 shows the workflow of our proposed study consisting of three main steps: pre-processing and augmentation, feature extraction and classification. The details of these step are provided in subsection. Before these steps, the description of data set applied in our proposed study is very important
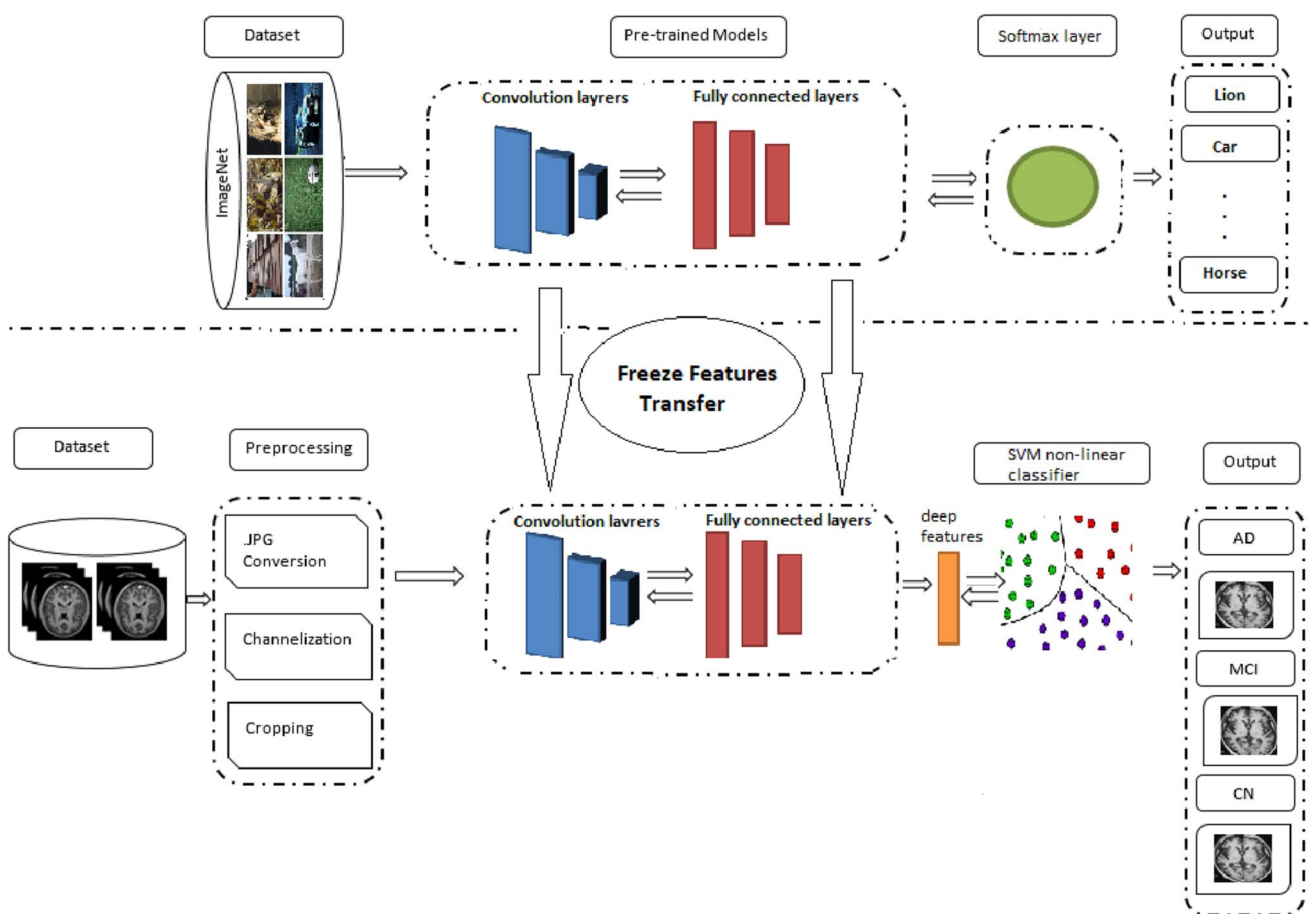

Fig. 2 Flow of proposed system using freeze features approach for AD identification

Table 1 Statistics of Alzheimer disease neuroimaging initiative data set

Subjects	ADNI-1	ADNI-Go	ADNI-2	ADNI-3
AD	200	–	150	133
LMCI	274	–	150	151
EMCI	–	200	150	–
CN	200	–	150	87
Total	474	200	550	291

Table 2 Demographic details and scans of subjects used in this study

Class	Subjects	Male/ Female	Total scans
AD	95	44/51	1425
MCI	138	66/72	1021
CN	146	87/59	1479

4.1 ADNI dataset

Data set is an important part in any field of pattern recognition and data mining. The benchmark data set for Alzheimer disease detection is ADNI (Alzheimer Disease Neuroimaging Initiative) [31]. All the experiments of proposed approach are carried out on datasets.

ADNI was launched in 2003 under the supervision of DR. Michael W. Weiner financed by public-private partnership with 27 million contributed by 20 companies and two foundations through the foundation to the National Institute of Health and 40 million for the National institute on Aging. ADNI is multisite, longitudinal study aimed to develop clinical, genetic, imaging and bio specimen biomarkers for the early diagnosis of Alzheimer disease (AD) [31]. It includes 1800 subjects with both males and females in three databases (ADNI, ADNI GO, ADNI 2 and ADNI 3). The statistics of ADNI is shown in Table 1.

In literature, a subset of used to train and validate different architectures of CNN classifiers. We also selected T1 weighted structural MRI data for 350 subjects, including 95 Alzheimer Disease (AD), 95 Cognitive Normal (CN) and 146 Mild cognitive Impairment (MCI). Multiple scans of each subjects were acquired at different time, so each subject

has different number of scans. In our data of subset, minimum scan is 3 and maximum scan is 15. Table 2 shows the number of participants in each class that we use in conducting experiments of proposed approach. Demographic details of the subjects are shown in Table 2.

4.2 Prepossessing

Preprocessing is an important step in image processing and data mining for enhancement, noise removal and smoothing of images [32–34]. In our study, we downloaded the raw MR images which are in DICOM (Digital Imaging and Communication in Medicine) format. As a first step of preprocessing, we converted images from DICOM to JPEG format. The converted JPEG images were in 1-channel and in different sizes. The convolutional neural networks (AlexNet, GoogLeNet, VGG, DenseNet and ResNet) requires 3-channel images data as input. Therefore, we convert all the data from 1- to 3-channel and re size. After channelization and re-size, we crop the data to remove the white spaces and to enhance the data (see Table 3).

For training convolutional neural network, having large data set is crucial. Data augmentation plays a significant role to significantly increase diversity of data for training the neural networks when originally data set is small and reducing the over fitting of data. In short, augmentation means randomly applying alteration on data set to increase the different variability of data in data set. For our experiments, we selected the data set of 379 patients that not enough for training CNNs and attaining better performance. To enhance and increase the variations of each samples of the data set, we implemented different augmentation techniques, i.e., flipping, rotation, illumination, and zooming. In ***flipping***, image information is mirrored in horizontal and vertical direction. In our data set, we are using MRI of brain. Brain structure containing symmetric features are ideally used for flipping. For each image I in a data set, we perform horizontal and vertical flipping. Flipping presents the network from favoring features that present in one side of the organ, while the disease might just has been present in other side.

In the *rotation*-based augmentation randomly rotates the image clockwise. we assign degree for rotation from 0 to 360. This way, the augmented images may better resemble

Table 3 Total number of images after data augmentation used in proposed study

Classes	Scans	Data augmentation techniques							
	Original	Rotation			Flipping		Zooming		Illumination
		90°	180°	270°	Horizontal	Vertical	Zoom in	Zoom out	
AD	1425	1425	1425	1425	1425	1425	1425	1425	1425
CN	1479	1479	1479	1479	1479	1479	1479	1479	1479
MCI	1021	1021	1021	1021	1021	1021	1021	1021	1021
Total	3925	37590							

patients that were scanned under a small angle. We use rotation of 90, 180 and 360 for our data set.

Zooming is a powerful augmentation that can make a network robust to (small) changes in object size. We perform both zoom in and zoom out data augmentation techniques. *Illumination* means to shine light on object and to make it clear to visible is another data augmentation technique used to enhance the data set. Before augmentation, total number of images of data set was 3925 and after data augmentation, 37,590 images were obtained. Table 4 shows the details of augmentation. In next section, the features extractions and classification will be discussed.

4.3 Features extraction and classification

As it is illustrated in Fig. 2 that the fixed and freeze feature-based kernels convolve with input vectors in convolution layers for evaluation of salient and distinct feature maps. The extracted deep features from different layers (conv, fc6, fc7, fc10000, predictions, or loss3-classifier) by applying different deep neural network architectures and then pass to support vector machine for classification and identification. In order to get generalised and reliable results, we employed 5 runs of each network and report the average accuracy of 5 runs.

The freeze weights of transfer learning is to cut away the first layers of the trained network and freeze their parameters, capturing generic image representations or "off the-shelf" features. We employed the freeze approach of transfer learning using eleven CNN architectures such as AlexNet, GoogLeNet, VGG-16, VGG19, ResNet18, ResNet50, ResNet101, MobileNetV2, InceptionV3, Inception-ResNet-V2, and DenseNet201 to obtain feature map utilizing only fully connected layers and then passing the feature map to SVM for predicting binary classification of three categories, i.e., AD vs. CN, AD vs. MCI, MCI vs. CN. These fully connected layers act as a fixed feature extractor, because the weights of all above layers are frozen or unchanged during re-training using target dataset.

5 Results analysis

This section details the results analysis of our proposed system for the Alzheimer's detection. For analysis and evaluation of the performance of proposed system, we have used and performed experiments on augmented data set. In this study, we selected 379 subjects from with three classes (AD, MCI, and CN). Each subject has number of MRI scans of th brain. The total scans of original data set is 3925 and then 39,750 scans after augmentation. We used accuracy to evaluate the predictions of AD using number of CNNs architectures using freeze features approach of transfer learning. The freeze and fixed features has superiority over scratched CNN and fine-tuned transfer learning approach in case of time complexity and memory. Freeze feature-based transfer learning need to freeze all convolution and fully connected layers, and only train last layer(s) on new data set. The data set is divided into 80% as training set, 10% as validation set and 10% as testing set.

Deep convolutional neural network was trained on augmented data set. The split of data set remains same for all experiments. As described earlier, fully connected layers of all deep CNN models were used except AlexNet, 5th convolutional layer of AlexNet named 'conv5' was also evaluated with two fully connected layers named 'fc6' and 'fc7'. Also two fully connected layers of VGG-16 and VGG-19 was evaluated for classification. features extracted from the FC1000 of ResNet18/50/101, Prediction of InceptionV3 and Inception-ResNet-V2, Logits of MobileNetV2 and FC1000 of Densenet201. The results obtained from deep neural

Table 4 Results of CNN architectures using Freeze Features Extracted by ImageNet Dataset (source data set) on Augmented dataset (target data set)

Models	Layers	MCI vs AD	AD vs CN	MCI vs CN
AlexNet	conv5	71.84	**91.38**	63.92
	FC6	99.22	82.61	94.27
	FC7	96.80	78.84	96.37
GoogLeNet	loss3-classifier	95.04	78.61	69.99
VGG16	FC6	99.14	**98.89**	**97.06**
VGG19	FC6	**99.27**	95.89	96.82
ResNet18	FC1000	94.88	73.07	80.61
ResNet50	FC1000	98.75	92.30	90.89
ResNet101	FC1000	98.53	87.07	89.08
MobileNetV2	Logits	96.95	89.95	86.58
InceptionV3	Predictions	94.37	89.99	80.63
Inception-ResNet-V2	Predictions	90.73	88.10	78.64
DenseNet201	FC1000	98.73	93.23	86.15

The bold one are highest accuracy

networks using freeze features on augmented data set are provided in Table 4.

To investigate the effectiveness of the proposed AD detection system, we conducted three studies for binary classification- based on different comparative classes and evaluate the classification accuracy achieved by each architecture of CNN, separately. In the first study of binary classification (MCI vs. AD classification), fully connected layers of all the networks obtained best accuracy. The highest accuracy of 99.26% achieved by FC6 of VGGNet and the lowest accuracy of 71.48% achieved by conv5 of AlexNet. In this case, high level features performed best for the prediction. In the second study of binary classification for prediction of AD and CN classes, the VGGNet16 achieved highest accuracy of 98.89% and the lowest accuracy is 78.61% achieved by loss3-classifier of GoogLeNet. Other networks achieved an average accuracy around 80%. In the third study of binary classification for MCI and CN classes, FC6 of VGGNet16 achieved highest accuracy of 97.06% and loss3-classifier of GoogLeNet achieved lowest accuracy of 69.99%. The results obtained using fixed feature approach on augmented data set is shown in Table 4.

6 Discussion and comparison

Figure 3 shows the comparison of results obtained using CNN-based freeze approach of three binary classification problems. The x-axis of graph shows networks and y-axis shows the accuracy of the training model. As shown in Fig. 3, that MCI vs. AD classification performed well as compare to AD vs. CN and MCI vs. CN-based classification.

The comparison of our work with the techniques and accuracy of other systems presented in the literature discussed in Table 5. Dataset is evaluated by available pre-trained CNNs architectures using freeze approach. The performance of proposed system is evaluated in terms of sensitivity, specificity and accuracy by models. We have also compared our class wise results with the results presented in literature (shown in Table 5). In competing approaches, Suk et al. [13] implemented Deep Boltzmann Machine for analysis of (AD/CN and MCI/CN) and they have achieved 95.35% and 85.6% results, respectively. Ju et al. [18] assessed the MCI/CN target classes and they have achieved the accuracy up to 86.47%. Iftikhar et al. [16] employed ensemble SVM to classify MCI/AD, MCI/CN and AD/CN and reported
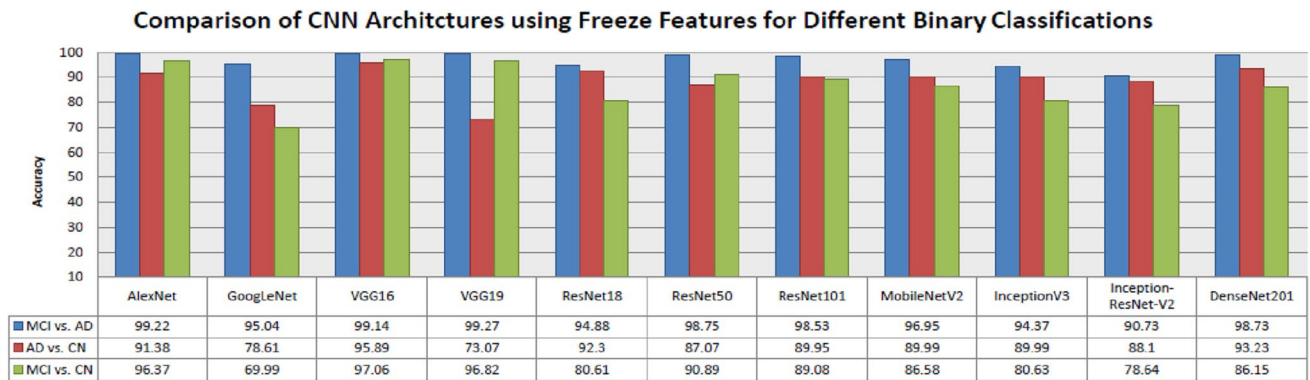

Comparison of CNN Architctures using Freeze Features for Different Binary Classifications

	AlexNet	GoogLeNet	VGG16	VGG19	ResNet18	ResNet50	ResNet101	MobileNetV2	InceptionV3	Inception-ResNet-V2	DenseNet201
MCI vs. AD	99.22	95.04	99.14	99.27	94.88	98.75	98.53	96.95	94.37	90.73	98.73
AD vs. CN	91.38	78.61	95.89	73.07	92.3	87.07	89.95	89.99	89.99	88.1	93.23
MCI vs. CN	96.37	69.99	97.06	96.82	80.61	90.89	89.08	86.58	80.63	78.64	86.15

Fig. 3 Comparison of performance of CNN architectures using freeze features approach for binary classification of MCI vs. AD, AD vs. CN and MCI vs. CN

Table 5 Comparison of performance of proposed system with performance of other systems for AD detection on dataset

Authors	Model	Accuracy (%)		
		MCI/AD	MCI/CN	AD/CN
Ju et al. [18]	Stacked auto encoder	–	86.47	–
Qiu et al. [22]	MMSE + CNN	–	90.9	–
Lin et al. [23]	CNN	–	–	88.79
Kazemi and Houghten [21]	AlexNet	–	–	98.74
Suk et al. [13]	DBM	–	85.67	95.35
Xia et al. [25]	3D CLSTM	–	79.0%	94.19%
Iftikhar et al. [16]	Ensemble SVM	91.66	90.83	98.83
Proposed	**VGG16**	99.14	**97.06**	**98.89**
	VGG19	**99.27**	96.82	95.89

The bold one are highest accuracy

91.66%, 90.83% and 98.83%, respectively. In our proposed study, VGGNet has produced more promising and baseline results upto 99.27% for MCI/AD, 98.89% for MCI/AD and 97.06% for MCI/CN.

7 Conclusion

In this study, the proposed technique relies on various architectures of convolutional neural network and transfer learning using freeze weight. We present pre-trained models using freeze feature-based approach that extracts visual and salient features. This approach has quick training and less computational complexity as compare to fine-tuned and scratched CNN-based approaches. We have extracted extracted freeze features from different layers of different models and trained different networks based on their salient features and patterns from different layers. For evaluation of proposed systems, we have used dataset. The proposed approach have compared with the recent state-of-the-art techniques. The performance of proposed networks MCI vs. AD classes showed that VGG19-SVM using features by fc6 has highest accuracy up to 99.27%. Likewise, the proposed techniques realised the accuracy of 98.89% using VGG16-SVM for AD vs. CN classes. The 97.02% accuracy achieved by VGG16-SVM using freeze fc6 layers for MCI vs. CN. In future, we thoroughly explore fusion of different layers and convolutions networks to get a robust AD identification system.

References

1. James, B.D., Leurgans, S.E., Hebert, L.E., Scherr, P.A., Yaffe, K., Bennett, D.A.: Contribution of Alzheimer disease to mortality in the united states. Neurology **82**(12), 1045–1050 (2014)
2. Brookmeyer, R., Johnson, E., Ziegler-Graham, K., Arrighi, H.M.: Forecasting the global burden of Alzheimer's disease. Alzheimer's Dement. **3**(3), 186–191 (2007)
3. Thakare P., Pawar V.: Alzheimer disease detection and tracking of Alzheimer patient. In: 2016 International Conference on Inventive Computation Technologies (ICICT); vol. 1. IEEE; 2016, p. 1–4
4. Abdalla, B., Yassin, M., Abir, M., Bisharat, B., Armaly, Z.: Traditional and modern medicine harmonizing the two approaches in the treatment of neurodegeneration (Alzheimer's disease - ad). In: Saad, M., de Medeiros, R. (eds.) Complementary Therapies for the Contemporary Healthcare, chap. 10. IntechOpen, Rijeka (2012). https://doi.org/10.5772/48558
5. Jin, J.: Alzheimer disease. JAMA **313**(14), 1488–1488 (2015)
6. Zahoor, S., Naz, S., Khan, N.H., Razzak, M.I.: Deep optical character recognition: a case of Pashto language. J. Electron. Imaging **29**(2), 023002 (2020)
7. Naz, S., Khan, N.H., Zahoor, S., Razzak, M.I.: Deep OCR for Arabic script-based language like Pastho. Expert Syst. **37**(5), e12565 (2020)
8. Rehman, A., Naz, S., Razzak, M.I., Hameed, I.A.: Automatic visual features for writer identification: a deep learning approach. IEEE Access **7**, 17149–17157 (2019)
9. Naz, S., Umar, A.I., Ahmad, R., Ahmed, S.B., Shirazi, S.H., et al.: Urdu Nasta'liq text recognition system based on multi-dimensional recurrent neural network and statistical features. Neural Comput. Appl. **26**(8), 219–231 (2015)
10. Naz, S., Umar, A.I., Ahmed, R.A.S.B., Siddiqi, I., Razzak, M.I.: Offline cursive Nastaliq script recognition using multidimensional recurrent neural networks with statistical features. Neurocomputing **177**, 228–241 (2016)
11. Naseer, A., Rani, M., Naz, S., Razzak, M.I., Imran, M., Xu, G.: Refining Parkinson's neurological disorder identification through deep transfer learning. Neural Comput. Appl. **32**(3), 839–854 (2020)
12. Rehman, A., Naz, S., Razzak, M.I., Akram, F., Imran, M.: A deep learning-based framework for automatic brain tumors classification using transfer learning. Circ. Syst. Signal Process. **39**(2), 757–775 (2020)
13. Suk, H.I., Lee, S.W., Shen, D., Initiative, A.D.N., et al.: Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. NeuroImage **101**, 569–582 (2014)
14. Sarraf S., Tofighi G., et al. Deepad: Alzheimer s disease classification via deep convolutional neural networks using mri and fmri. *BioRxiv* 070441 (2016)
15. Mathew, J., Mekkayil, L., Ramasangu, H., Karthikeyan, B.R., Manjunath, A.G.: Robust algorithm for early detection of alzheimer's disease using multiple feature extractions. In: IEEE Annual India Conference (INDICON). IEEE 2016, 1–6 (2016)
16. Iftikhar M.A., Idris A.: An ensemble classification approach for automated diagnosis of Alzheimer's disease and mild cognitive impairment. In: 2016 International Conference on Open Source Systems & Technologies (ICOSST). IEEE; p. 78–83 (2016)
17. Hosseini-Asl E., Keynton R., El-Baz A.: Alzheimer's disease diagnostics by adaptation of 3d convolutional network. In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE; p. 126–130 (2016)
18. Ju, R., Hu, C., Zhou, P., Li, Q.: Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. IEEE/ACM Trans. Comput. Biol. Bioinform. (TCBB) **16**(1), 244–257 (2019)
19. Farooq A., Anwar S., Awais M., Rehman S.: A deep cnn based multi-class classification of Alzheimer's disease using MRI. In: 2017 IEEE International Conference on Imaging systems and techniques (IST). IEEE; p. 1–6 (2017)
20. Bäckström, K., Nazari, M., Gu, I.Y.H., Jakola, A.S.: An efficient 3d deep convolutional network for Alzheimer's disease diagnosis using MR images. In: IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE 2018, 149–153 (2018)
21. Kazemi Y., Houghten S.: A deep learning pipeline to classify different stages of Alzheimer's disease from FMRI data. In: 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE; 2018, p. 1–8 (2018)
22. Qiu, S., Chang, G.H., Panagia, M., Gopal, D.M., Au, R., Kolachalama, V.B.: Fusion of deep learning models of MRI scans, mini-mental state examination, and logical memory test enhances diagnosis of mild cognitive impairment. Alzheimer's Dement. Diagn. Assess. Dis. Monit. **10**, 737–749 (2018)
23. Lin, W., Tong, T., Gao, Q., Guo, D., Du, X., Yang, Y., et al.: Convolutional neural networks-based MRI image analysis for the Alzheimer' disease prediction from mild cognitive impairment. Front. Neurosci. **12**, 777 (2018)
24. Payan A., Montana G.: Predicting Alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. arXiv preprint arXiv:150202506 (2015)
25. Xia, Z., Yue, G., Xu, Y., Feng, C., Yang, M., Wang, T.: A novel end-to-end hybrid network for Alzheimer' disease detection using

3d CNN and 3d CLSTM. In: IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE 2020, 1–4 (2020)

26. Ebrahimi-Ghahnavieh A., Luo S., Chiong R.: Transfer learning for Alzheimer's disease detection on MRI images. In: 2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT). IEEE; 2019, p. 133–138 (2019)

27. Ashraf, A., Naz, S., Shirazi, S.H., Razzak, I., Parsad, M.: Deep transfer learning for Alzheimer neurological disorder detection. Multimed. Tools Appl., 1–26 (2021)

28. Mehmood, A., Ahmad, A.S., Maqsood, M., Yaqub, M., et al.: A transfer learning approach for early diagnosis of Alzheimer' disease on MRI images. Neuroscience. **460**, 43–52 (2021)

29. Liu, J., Li, M., Luo, Y., Yang, S., Li, W., Bi, Y.: Alzheimer's disease detection using depthwise separable convolutional neural networks. Comput. Methods Programs Biomed. **203**, 106032 (2021)

30. Chen, Y., Xia, Y.: Iterative sparse and deep learning for accurate diagnosis of Alzheimer's disease. Pattern Recognit., 107944 (2021)

31. Sandeep, C., Kumar, A.S., Susanth, M.: The online datasets used to classify the different stages for the early diagnosis of Alzheimer' disease (ad). Int. J. Eng. Adv. Technol. **6**(4), 38–45 (2017)

32. Yan, C., Li, Z., Zhang, Y., Liu, Y., Ji, X., Zhang, Y.: Depth image denoising using nuclear norm and learning graph model. ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) **16**(4), 1–17 (2020)

33. Yan, C., Shao, B., Zhao, H., Ning, R., Zhang, Y., Xu, F.: 3d room layout estimation from a single RGB image. IEEE Trans. Multimed. **22**(11), 3014–3024 (2020b)

34. Yan, C., Gong, B., Wei, Y., Gao, Y.: Deep multi-view enhancement hashing for image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. **43** (2020)