# Deep transfer learning for alzheimer neurological disorder detection

**Abida Ashraf[1] · Saeeda Naz[1] · Syed Hamad Shirazi[2] · Imran Razzak[3] · Mukesh Parsad[4]**

**Abstract**

Alzheimer's disease is becoming common in the world with the time. It is an irreversible and progressive brain disorder that slowly destroys the memory and thinking skills and, eventually, the ability to perform the simplest tasks. It becomes severe before the notice-able symptoms appear and causes brain disorder which cannot be cured by any medicines and therapies, however its progression can be slow down through early diagnosis. In this paper, we employed different CNN based transfer learning methods for Alzheimer disease classification. We have applied different parameters, and achieved remarkable accuracy on benchmark ADNI dataset. We have tested 13 differnt flavours of different pre-trained CNN models using a fine-tuned approach of transfer learning across two different domain on ADNI dataset (94 AD, 138 MCI and 146 NC). Comparatively, DenseNet showed better performance by achieving a maximal average accuracy of % 99.05. Significant improve-ment in accuracy has been observed as compared to previously reported works in terms of specificity, sensitivity and accuracy. The source code of propose framework is publicly available.

**Keywords** Alzheimer disease · Deep transfer learning · ImageNet · ADNI dataset

## 1 Introduction

The brain is the most complex and integral part of the human body. The essential function of the brain is controlling the central nervous system. Important tasks of the human brain

https://github.com/imranrazzak/Neurological-Disorder

✉ Imran Razzak
  imran.razzak@ieee.org

[1] GGPGC No.1, Abbottabad, Pakistan

[2] Department of IT, Hazara University, Mansehra, Pakistan

[3] Deakin University, Geelong, Australia

[4] University of Technology, Sydney, Australia

include thinking, movement, coordination, creative visualization, learning, memory and emotional responses. As the brain is the controlling organ of the entire human body and any abnormal function of the brain can paralyze the intact body functionalities. Hippo-campus is an integral part of the brain and located deep in the temporal lobe that relates to memory, learning, emotions. Various stimuli can affect this and can result in psychiatric and neuro-logical disorders like epilepsy, Alzheimer, depression, or emotion [1]. Alzheimer's Disease (AD) is one of the brain abnormality, which causes problems with thinking, memory, and behaviour. Abnormal clumps (amyloid plaques), loss of connections neurons and tangled bundles of fibers (neurofibrillary, or tau, tangles) are the main features of Alzheimer's disease. This damage initially appears to take place in the hippocampus, thus, considered as a valid biomarker in patients with Alzheimer disease. With the further death of neurons, additional parts of the brain are affected. By the final stage, the damage is significant, and brain tissue has shrunk extensively. Unfortunately, there is no cure available for Alzheimer's disease because it is irreversible neuro-degenerative dementia. For the better prognosis of Alzheimer's disease, it is essential to develop reliable automatic detection system to detect and diagnose the disease at the earliest stage. The diagnosis of Alzheimer disease is mainly clinical but imaging studies [20] 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 conducted like a blood test, CT (Computed tomography), MRI (Magnetic resonance Imaging) and PET (Positron Emission Tomography) scans, EEG (Electroencephalography) and Genotyping. According to the recommendations of the American Academy of Neurology, structural imaging of the brain (either contrast or non-contrast) CT scan or MRI is most suitable evaluation method for the early detection of dementia [9]. Imaging studies help us to rule out curable causes of progressive cognitive deficit which includes normal pressure hydrocephalus, etc. MRI and CT scans show diffuse cortical and cerebral atrophy in patients with Alzheimer disease.

The emerging of new technologies and their impacts on every field of life, we have an opportunity to improve the diagnosis of various diseases in hospitals. By using computer-aided algorithms, we can enhance the medical diagnosis process in terms of efficiency and accuracy. Different conventional machine learning techniques were used by many researchers to classify AD using Neuro-imaging data like Support Vector Machine (SVM), Decision Tree (DT), or k-Nearest Neighbours (KNN), etc. Since 2006, deep learning methods have got more attention of researchers and have brought dramatic improvements and advancements in pattern recognition, image processing, automated driving, computer vision, and medical imaging applications. The number of deep learning methods (Convolutional Neural Network, Recurrent Neural Network, etc.) and transfer learning techniques (scratched layers, freeze weights or fine-tuned weights) [38] used to detect AD with less effort of radiologists automatically.

It is challenging to distinguish among the different stages of Alzheimer Disease. For this purpose, careful clinical assessment is required along with the keen observation of radiologist. This process is tedious and expensive. The motivation behind our work is to diagnose AD with the help of transfer learning, and deep CNN architectures and progression of the disease could be slow down. Automated detection of AD will also reduce the involvement of radiologist and cost. In this paper, we propose a deep neural (DNN) based Alzheimer disease prediction system to extract salient visual features from MRI. The main aim of this method is to overcome the limitations of traditional machine learning approaches. These features are used to discriminate among AD, MCI and NC and predict the required disease in the early stage.

The key contributions of this study are:

- Different architectures of the second generation and third generation of Deep Neural Network (DNN) are deployed to investigate the visual features using MRI images of ADNI dataset in characterizing AD, MCI and NC.
- Use of multiple data augmentation schemes for increasing and enhancing the input space using raw images for extraction of salient features.
- Comparison of thirteen architectures of deep neural networks like Spiking neural networks, DenseNet, MobileNet, Squeeznet, ResNet, VGG, GoogLeNet, etc. using multiple representations of input samples for improved prediction and classification rate.
- Spiking neural networks are biologically plausible and very energy efficient among the neural networks. In this work, we have achieved promising results, which are almost equivalent to the second generation neural networks in terms of accuracy.

The rest of the paper is structured as follow. Section 2 containing background of deep neural networks and their architectures, in Section 3 related work is discussed, Section 4 contains the details about the ADNI dataset used in this work. In Section 5 the details of the models we have applied are explained while Section 6 includes Experimental setup, results, discussion, Section 7 contains concluding remarks.

## 2 Deep neural networks

In the past few years deep learning models have outperformed the state of the art machine learning approaches due to its promising results in different domains. Deep learning approach allows multi-layered processing models to learn and characterize data with a high level of abstraction, imitating how the human brain comprehends and perceives multi-modal data [25]. Deep learning is a rich family of techniques encompassing a variety of neural networks, spiking neural networks, probabilistic models, supervised and unsupervised feature learning algorithms [27, 28]. Convolutional Neural Networks are the type of networks which perform well for object recognition and image classification. It is a feed-forward artificial neural network that learns image features through convolutional layers by its own. CNN had the ability to learn features without explicit labelling. With the passage of time, researchers developed different CNN models like AlexNet, LeNet, GoogLeNet, VGG (16 and 19), ResNet, SegNet, UNet and many more to facilitate the humans in a variety of fields [23]. In this article, we have discussed the 12 models from the second generation neural networks and one model from the third generation of neural networks. In the following section, all of these networks and their architectures are briefly explained.

### 2.1 Second generation neural network

The ANN models are presently in their second generation. The second-generation networks are made of interconnected arrays of distinct processing elements or neurons. These distinct elements are added and transform the input signals to output. Appropriate adjustments of synaptic weights transform input patterns into analogous output patterns. In second-generation neural networks, the flow occurs in both forward and backward directions, and we deal with continuous output values. The second generation artificial neural networks have played a vital role in the understanding, innovation and exploitation of latent functional resemblances between human and artificial information systems. These networks are

computationally built on mathematical tools enthused by the neurons. In the current era the evolution of deep learning has revolutionized the domain machine learning especially for big data [4] and computer vision. In a deep learning approach, the network (ANN) is trained in a supervised manner by using the back-propagation. For training of deep learning a huge amount of training, data is vital, but the subsequent classification accuracy is very impressive. In second-generation neural networks, the neurons are characterized by static, single and continuous-valued activation. The main advantage of second-generation neural networks (ANN) over the third generation neural networks (SNN) is that in terms of accuracy, 3rd generation networks are lag behind the ANNs.

### 2.1.1 AlexNet

AlexNet is a deep Convolutional network developed by Krizhevsky et al. [33] designed for image classification. In 2012 AlexNet won the most difficult ImageNet Challenge for the recognition of visual objects known as ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. In this challenge, AlexNet achieved the highest accuracy by beating all the existing Machine learning and computer vision techniques. AlexNet won the challenge by reducing the top 5 errors from 26% to 15%. AlexNet was trained on ImageNet dataset that contained 1.2 million images. The model was based on 650,000 neurons and contained 60 million parameters for the classification of 1000 classes. This model was trained on 2 GPUs for 5 to 6 days. The architecture of this model consists of the input layer, 5 convolutional layers, max polling layers, dropout layers and 3 fully connected layers and the final softmax layer. The input layer takes RGB image of size $227 \times 227 \times 3$ as input. In case of greyscale image, it has to be converted into RGB image, and it then will become an acceptable input. It has 5 convolutional layers having different feature maps and strides. This layer convolves the output received from the previous layer with different kernels and proceeds with different linear and non-linear activation functions such as sigmoid, hyperbolic tangent, softmax, identity and rectified linear functions. In fully connected layers, each neuron will be connected to all the numbers in the previous volume. Three Fully connected layers are used by Alex net named FC6, FC7 and FC8. FC6 and FC7 use 4096 neurons, while FC8 is used for classification. Softmax is the last layer of this model, and its purpose is to label the data and predict classes or categories in the testing phase.

### 2.1.2 GoogLeNet

GoogLeNet was developed by Szegedy et al. [35] won the ILSVRC in 2014 with the top 5 error rate of 6%. It consumes less power and memory as compared to AlexNet. It also uses 12 times less number of parameters than the AlexNet model because it uses the inception module. GoogleNet architecture consists of 22 layers. Initial layers are simple convolutional layers and using a combination of $1 \times 1$, $3 \times 3$ and $5 \times 5$ convolutional filters. RELU $1 \times 1$ convolution is used for the reduction of dimensionality, after that weight and height is increased. After convolution layer, there are multiple inception modules and is the basic building block of GoogLeNet. It uses nine inception modules sequentially stacked, with two max-pooling layers along the way to reduce the spatial dimensions. For the previous input, $1 \times 1$ Conv, $3 \times 3$ Conv and $5 \times 5$ and $3 \times 3$ max polling are done. GoogLeNet does not use fully connected layers; instead, it uses average pooling by averaging each feature map from $7 \times 7$ to $1 \times 1$. GoogleNet used a much lower number of network parameters, i.e. 7M as compared to AlexNet that used 60M and VGG-19 used 138 M network parameters.

### 2.1.3 VGG

Karen Simonyan and Andrew Zisserman [32] developed a network called VGG (Visual Geometry Group). It is the runner up of ILSVRC in 2014. VGG proved that depth of network is vital for better performance. VGG network contains 16 CONV/FC layers and performs $3 \times 3$ convolutions and $2 \times 2$ pooling from the beginning to the end and have a lot of filters. It also contains two fully connected layers each having 4096 nodes followed by a softmax layer. VGG network uses Kernel size of $3 \times 3$. VGG has three flavours of models, i.e. VGG-11, VGG-16 and VGG-19, they have 11, 16 and 19 layers respectively. However, they have different numbers of convolutional layers VGG-11 contains eight convolutional layers, VGG-16 has 13 convolutional layers and VGG-19 has 16 convolutional layers. Among these three models, VGG-19 is the most computationally expensive; it contains 138M network parameters.

### 2.1.4 ResNet

ResNet was developed by Kaiming He et al. [10] is the winner of ILSVRC 2015 with an error rate of 3.6%. Its special feature is to skip connection, and it uses batch normalization. Skip connection is also called gated recurrent units. ResNet does not use fully connected layers at the end. It was trained on 8 GPUs for two to three weeks. In this network, after first two layers the spatial size get compressed to an input size of $224 \times 224$ to $56 \times 56$. ResNet architecture contains 152 layers and uses residual blocks. Various ResNet models are developed with a different number of layers; RestNet 34, 50, 101, 152 and 1202. The most commonly used ResNet50 has 49 convolutional layers and one fully-connected layer. It has 25.5 M network parameters for the whole network 3.9M MACs.

### 2.1.5 DenseNet201

After the development of skip connections in convolutional network, more efficient and accurate results were achieved by Using the advantage of skip connections, another model called DensNet was proposed by Huang et al. [14] in 2017. In standard convolution Network, each layer connects to its subsequent layer and L layers having L connections. In DenseNetwork, each layer is connected to every other layer in a feed-forward manner, and L layers have $L(L + 1)/2$ connections. DenseNet architecture was tested on four benchmark object recognition tasks like CIFAR-10, CIFAR-100 and ImageNet and evaluated the state of the art performance with fewer parameters. Besides other advantages of strengthens features propagation encourages feature rescue, DenseNet provides an improved flow of information and gradient throughout the network. The main difference between DenseNet and ResNet is that the inputs are concatenated in DenseNet while summed in ResNet.

### 2.1.6 MobileNetV2

In 2018, Howard et al. [13] introduced another member of Neural Network with mobile devices for detection, classification and segmentation named MobileNet-V1. The main idea behind the characteristic to run the deep network on mobile devices is to provide user more experience allowing access at anywhere and any time with other advantages of security, privacy and efficiency. To power the next generation of mobile vision applications, Sandler et al. [30] introduced MobileNet-V2, which is the modified version of MobileNet-V1. It is faster than the previous version and requires 2x fewer parameters while retaining the same

accuracy. MobileNet-V2 uses two new features in its architecture, i.e. inverted residuals and linear bottleneck in addition to depthwise separable convolution as in MobileNet-V1. Bottleneck encapsulate low dimensional input to high dimensional representation. As the bottleneck contains all the necessary information, the authors use skip connection between bottlenecks.

### 2.1.7 SqueezeNet

Another deep Neural Network named SqueezeNet was introduced in 2016 by Iandola et al. [16]. The main focus behind the development of this network is to achieve high-level accuracy with 5x fewer parameters than other networks. Smaller architectures provide advantages of efficient training due to fewer parameters, less bandwidth because it requires less communication, and it is feasible to develop on FPGAs without limitation of bandwidth. Designing a model with fewer parameters while retaining equivalent accuracy, they used three strategies for the architecture of SqueezeNet. They introduced the 'fire module' basic building block of SqueezeNet architecture. A fire module consists of Squeeze convolution layer which has 1x1 filters, forward to expand layer which has both 1x1 and 3x3 convolution filters. The SqueezeNet architecture starts with standalone convolution followed by eight fire module and ends with the final convolution layer. The number of filters increases per fire module from start to end of the network. After the layer conv1, fire4, fire8 and con10, max polling is performed with a stride of 2. The model was evaluated with other models and concluded that SqueezeNet achieves high accuracy like AlexNet with 5x fewer parameters.

### 2.1.8 Inception V3

In the development of convolutional neural network, Inception module plays a significant role. To achieve better and efficient performance, most CNN classifiers goes deeper and deeper. On the other hand, Inception V1/GoogLeNet was proposed by [35] to reduce the number of parameters. Later, various Inception architecture was proposed. Inception V1 uses a combination of $1 \times 1$, $3 \times 3$ and $5 \times 5$ convolution layers and max-pooling layer, to operate multiple sizes input at the same level and concatenated outputs forwarded to next inception. The network goes wider instead of deeper because deeper networks are more likely to be overfitting. In Inception V2, Batch Normalization was introduced. RELU (rectified linear unit) layer was used as activation. The major strength is to reduce representational bottleneck. The third version of the deep learning inception module named Inception V3 was introduced by [36]. It was trained on 1000 categories of ImageNet dataset, was the first runner up in ILVRC and achieved 78.8% accuracy. Inception V3 introduces factorizing convolutions which reduce the number of computation without affecting the network efficiency. Factorizing convolutions uses 5x5 filters instead of 3x3. The model architecture comprised of symmetric and asymmetric building blocks, having convolutions, dropout, max pooling and fully connected layers.

### 2.1.9 Inception-ResNet-v2

In 2017, Szegedy et al. [34] introduced three new deep model architectures i.e. Inception-ResNet-v1, Inception-ResNet-v2 and Inception v4. Inception-ResNet-v1 uses the concept of inception module and residual network instead of filter concatenation. It performed the state of the art performance with the almost equal computational cost of inception v3. Inception-ResNet-v2 achieves the state of the art performance with less computational cost

as compared to inception-ResNet-v1. Inception v4 is the simplified version of inception v3 and uses more inception modules than inception v3, and no residual blocks were used. They achieved 3.08% top-5 error rate on the ImageNet classification challenge on the test set.

## 2.2 Third generation of neural network and spiking neural network

To classify neural networks based on their computational units, the neural network has three generations. The first generation of neural networks are called perceptrons or threshold gates. These models generate an only digital output. They are considered universal computational units for digital input and output. A variety of neural networks are developed in first-generation including Hopfield net, Boltzmann machine and multilayer perceptron. The second generation neural network is more complicated that uses complex algorithms, architectures and lot of activation functions. Second generation neural network consists of neurons that apply an activation function to a continuous sum of input values to get continuous output. Neural network of second-generation support learning algorithms that are based on backpropagation and are able to compute input and outputs with analog functions. The networks in the third generation are more likely to mimic the behaviour of biological neurons. Transfer of information in these neurons models the information transfer in biological neurons. Second generation neural networks are accumulating accolades on various tasks in the field of computer vision, i.e. pattern recognition, segmentation and classification. But there are certain challenges i.e. inefficiency and computational cost posed by these networks. The spiking neural networks models address these problems. Spiking Neural Networks (SNNs) considered a third-generation neural network.

### 2.2.1 Spiking neural networks

Spiking Neural Networks (SNNs) are artificial neural networks that mimic the biological behaviour of neural networks and have the potential to solve problems related to biological stimuli. SNNs operates using spikes and incorporate in real-time rather than continuous input values. SNNs are more powerful than networks in the second generation. Spiking neural networks (SNN) have attained significant rank in the research community due to high bio fidelity [40] and its sparse event driven nature. Spiking neural networks are more biologically plausible from their competitors the artificial neural networks (ANN). The main difference between ANN and SNN is that ANNs are usually trained by stochastic gradient descent, while SNN have trained with timing-dependent plasticity. Training of deep convolution neural networks is power and memory-intensive job while spiking neural networks are beneficial to minimize the power consumption. The communication or information flow in these neural networks take place by a sequence of spikes. Neurons in SNN can process substantial information using a small number of Pulse (Spikes). The SNNs are much faster than Rate coding neural networks as these uses Pulse coding for information processing. The SNNs [7] are much faster than Rate coding neural networks as these uses Pulse coding for information processing. The past few years have witnessed a significant shift in modelling and formulation of neural network that can replace ANN as new computing paradigm and the next generation of artificial neural networks. The spiking neural networks (SNNs) was propelled by the need for a better understanding of how the biological brain learns with noisy spiking activity to compute functions robustly, efficiently and reliably. In spiking neural network, information is communicated in a sparse, asynchronous and parallel fashion. The SNNs are compatible to neuromorphic hardware due to fast inference, event-driven

nature and low power consumption. Various methods have been developed for training spiking neural network that makes them competitive to the deep convolutional neural network. In this article, we employed a basic idea coined by [15] of differentiable approximation of spiking neurons during training and actual spiking neuron during testing. We used the Nengo library, which can perform this transformation automatically.

## 3 Related work

Alzheimer's disease (AD) is the main cause of memory loss and progressive brain atrophy. It is the most common neuro-degenerative and irreversible degenerative diseases in the world. Alzheimer's disease (AD) is the most widespread disease with an estimation of more than 30 million affected people around the world. Prediction and Diagnosis AD through clinical assessment is a complicated and challenging task. Cognitive tests and Clinical Dementia Rating are not enough for the early detection of AD. Histological examination at post-mortem biopsy can provide the final diagnosis of AD, which is infeasible for living patients. Therefore, imaging plays a vital role in the enhancement of AD diagnosis and prediction. The interest of research is rapidly growing in computer vision, and medical imaging that employs deep learning and machine learning approaches for the diagnosis of AD. In these machine learning techniques, features like tissue density, shape and pixel or voxel intensity are exploited for the training of a classifier for the classification of subjects like AD, MCI and CN.

Saman and Ghassem [31] have introduced a deep learning method to detect the AD from healthy controls. They collected records of 28 AD patients and 15 healthy persons from ADNI and performed preprocessing steps of skull stripping, motion correction, spatial smoothing, noise removal and registration. After preprocessing step the data is passed to Le-Net model, and 96.85% accuracy is achieved. Another method for the early detection of AD is proposed by Mathew et al. [22], in which ADNI dataset is used for experiments on 158 MRIs of patients including 71 healthy controls and 87 AD patients. They apply image normalization, image resize, and image cropping and image reorientation as preprocessing. While for feature extraction PCA (Principal Component Analysis) and DWT (Discrete Wavelet Transform) was used. They used SVM (Support Vector Machine) for classification and achieved an accuracy of 84% and 91% for Ad vs. CN and MCI vs.CN, respectively.

Asl et al. [12] proposed a deep 3D convolutional neural network (3D-CNN) to diagnose the AD. The experiments were carried out on MRI, which includes 70 AD, 70 MCI and 70 NC data obtained from ADNI. Local features were extracted from the 3D input image using CAE (Convolutional Auto Encoder). The model was trained on CAD-Dementia data set which contains T1 weighted MRIs of AD, CN and NC. Skull stripping, spatially normalization were performed as preprocessing. Features extracted from CAD-Dementia were used as bio-marker to detect the AD on ADNI dataset using a fine-tuned approach. The classification was performed using ten-fold cross-validation technique and achieved 97.6% of accuracy for AD vs.NC classification.

In [18] Ju et al. used MRI and textual data (age, gender, and genetic information) to diagnose the Alzheimer by applying 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. They 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 preprocessing [24, 26]. Rf-MRI time-series data and correlation coefficient

data was used 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% respectively. The accuracy, sensitivity and specificity obtained by autoencoder are 86.47%, 92%, 81% respectively. All these calculations are achieved using correlation coefficient data.

Farooq et al. [5] employed deep learning models i.e. GoogLeNet, ResNet-18 and ResNet-152 for multi-class classification of AD. Experiments were conducted using ADNI dataset with four classes named AD, LMCI, MCI and CN having MRI's of 33, 22, 449 and 45 respectively. By using GoogLeNet they obtained accuracy of 98.88%, 98.01% for ResNet-18 and 98.14% for ResNet-152. Backstorm et al. [2] described an efficient and simple approach for the detection of AD called three- dimensional convolutional neural network architecture (3D ConvNet) using MRIs of brain. They have performed pre-processing activities like cortical reconstruction, trim edges, image resize and intensity normalization and extracted automated features. Data was gathered from ADNI dataset and conducted experiments on 340 subjects which includes 1190 MRI scans of 199 AD patients (103 male, 96female) and 141 normal controls (75 male, 66 female) to detect Alzheimer and Normal persons and obtained 98.78% accuracy on test dataset when data is randomly portioned into 60%, 20% and 20% for training, testing and validation respectively. The authors in [6] proposed a one-class classification (OCC) technique that requires samples from single class for training. They have embedded the minimum variance information with OCC architecture. This approach enhanced the generalization capability of the classifier, and it also reduced the intra-class variance. They have conducted experiments on 18 benchmark datasets, and the proposed method has yielded more than 5% F1 score as compared to the existing state of the art approaches. The main advantage of the one-class classifier is that it is used where data samples from other classes are very few or not available.

A framework [21] consist of two deep learning models is presented. The first one is the multi-task deep CNN for hippocampus segmentation and Alzheimer disease classification. This multi-task deep CNN generates binary segmentation mask of the hippocampus. The feature learned by multi-task model is insufficient for the accurate classification of AD. Therefore, 3D patch hippocampus features are extracted based on centroid to cover the deficiency. The second model is 3D-DenseNet used for learning features for AD classification. They have performed classification based on three classes (AD,MCI,NC). The proposed method achieved a classification accuracy of 88.9% for AD/NC classification, which is higher than the voxel-wise features (86.1%) and ROI features (84.7%). Hippocampus is the area of the brain which is affected first of all. For the early detection of AD the shape and volume are measured by using structural MRI. However, these features contain limited information which can lead to erroneous segmentation. Kazemi and Houghten [19] used fMRI data from ADNI to classify different stages of Alzheimer disease. They gathered data of 197 subjects of five classes named as AD, NC, 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 preprocessing. 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 experiment. They achieved an 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, NC and SMC respectively. Tajbakhsh et al. [37] in their study "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?" find the answer that whether training

of CNN from scratch (full training) or fine-tuned CNN approach sufficient for medical image analysis problems. For this purpose, they conducted experiments using both strategies, i.e., CNN from scratch and CNN fine-tuned strategy to classify four different classes of medical imaging applications and they concluded that fine-tune approach on ImageNet dataset for the classification, detection and segmentation of medical images, it provides efficient performance than training from scratch. The fact is that to train a CNN from scratch requires a large labelled training dataset of medical imaging, which is very difficult to meet the requirements in the medical field. It also requires considerable expertise, more memory usage and computational power. On the other hand, pre-trained CNN on ImageNet dataset provides the promising results on computer vision applications and across various application from natural images to medical image analysis. After all the experiments, they demonstrated that results that pre-trained CNN with fine-tuned performs best and in worst scenario pre-trained CNN with fine-tuned performs same as training from scratch. Therefore, ImageNet dataset could be used not only for computer vision applications but also in medical image analysis. As due to the deep architecture of CNN, it provides significant results to extract discriminating features. In 2019, Ghahnavieh et al. [3] used transfer learning to detect AD using MRI's from ADNI dataset. MRI's of 132 subjects for each AD and NC were used for conducting experiments. They used recurrent neural network with convolutional neural network to better understand the relationship between a sequence of input images. They extracted features using convolutional neural network and then trained recurrent neural network and improved accuracy. The authors in [39] proposed a 3D-CNN, which involves dense connections to distinguish between AD and MCI using MRI images. The distinction between AD and MCI is beneficial in identifying various types of dementia disease for appropriate treatment. They have introduced dense connections to address the problem of limited training data. The advantage of dense connections is that it improved the information and gradients propagation which made the training easy. They have used a fusion approach to combine base classifiers. The ensemble-based model produced achieved 97.19%. accuracy. Summary of the related works discussed above is shown in Table 1.

## 4  Data set

In this study, we obtained the neuro-imaging data from Alzheimer Disease Neuroimaging Initiative (ADNI) database, which is publicly available at the website [29]. ADNI was launched in 2004 under the supervision of DR. Michael W. Weiner financed by a public-private partnership, with 27 million contributed by 20 companies and two foundations. The amount of 40 million for the National Institute on Aging is received through the foundation. ADNI is multisite, the longitudinal study aimed to develop clinical, imaging, genetic and biospecimen biomarkers for the early diagnosis of Alzheimer disease (AD) [29]. The primary goal of ADNI has been to test whether MRI (Magnetic Resonance Imaging), PET (Positron Emission Tomography), and other biomarkers, clinical and other neurological assessments can be combined to detect the Alzheimer disease at the early stage of MCI (Mild Cognitive Impairment). It includes 1800 subjects with both males and females in four databases (ADNI, ADNI GO, ADNI 2 and ADNI 3).

In this study, we acquired baseline, 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

**Table 1** Summarizes Alzheimer Disease Detection using ADNI dataset based work in the literature

| Reference | Preprocessing | Features | Dataset | Classifier | Accuracy |
|---|---|---|---|---|---|
| Saman et al. [31] | Motion correction, skull stripping, spatial smoothing, low level noise removal | Model based | MRI of 28 AD and 15 CN | LeNet | 96.86 (%) (AD/CN) |
| Mathew et al. [22] | Image reorientation, Cropping, Skull stripping, Image normalization | DWT, PCA, FDR | MRI's of 71 CN and 87 AD | SVM | 84 (%) (AD/CN) 91.7 (%) (MCI/CN) |
| Asl et al. [12] | Skull stripping, spatially normalization | Local features were extracted from 3D input image using Convolutional Auto Encoder | T1 weighted MRI scans of 70 AD, 70 MCI and 70 CN | 3D-CNN | 97.6 (%) (AD/NC) |
| Ju et al. [18] | Motion correction, Slice timing correction, skull stripping, Normalization, Smoothing | AAL Atlas is used to identify the region of interests and 90 regions of interest were selected | MRI of 91 MCI and 79 CN | LDA LR Stacked Auto Encoder | 67.72 (%) 71.38 (%) 78.91 (%) 86.47 (%) (MCI/CN) |
| Backstrom et al. [2] | Cortical reconstruction, Trim edges, image re size, intensity normalization | Model based | 3D MRI scans of 199 AD (103 M, 96 Fe) And 141 Normal Controls (178 M,162) | 3D ConvNet | 98.78 (%) (AD/CN) |

**Table 1** (continued)

| Reference | Preprocessing | Features | Dataset | Classifier | Accuracy |
|---|---|---|---|---|---|
| Kazemi and Houghten [19] | Brain extraction, slice timing correction, spatial smoothing, high pass filtering, spatial normalization | Model based | MRI of 55 CN, 25 SMC,46 EMCI, 39 LMCI, and 29 AD | AlexNet | 98.74 (%) (AD/CN) |
| Ghahnavieh et al. [19] | Intensity normalization, registration | Model based | 132 subjects for AD and CN | AlexNet + LSTM GoogLeNet + LSTM VGGNet-16 + LSTM VGGNet-19 + LSTM SqueezeNet + LSTM ResNet-18 + LSTM ResNet-50 + LSTM ResNet-101 + LSTM Inceptionv3 + LSTM | 87.50 (%) 84.38 (%) 84.38 (%) 90.62 (%) 81.25 (%) 87.50 (%) 78.12 (%) 81.25 (%) (AD/CN) |
| Liu et al. [21] | affine registration, hippocampal image patches | Model based | 97 subjects for AD, 233 MCI, 119 NC | Multi-task CNN DenseNet Combination | 86.6 (%) 87.2 (%) 92.5 (%) (AD/NC) |
| Wang et al. [39] | grad-warping, intensity correction | Model based | 221 subjects for AD, 297 MCI, 315 NC | D-DenseNet ensemble method | 94.77 (%) 97.19 (%) 92.5 (%) (AD/MCI) |

**Table 2** Demographic detail of dataset used in this paper

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

have different number of scans. In our data, minimum scan is 3 and maximum scan is 15. Demographic details of the subjects are shown in Table 2.

## 5 Proposed study

The detailed flow of the proposed study of Alzheimer Disease (AD) detection using deep neural networks is as shown in Fig. 1. The proposed study relies on data-augmentation and fine-tuned features extraction using twelve architectures of CNN (second generation of deep neural network) and a spiking neural network (third generation of deep neural network). In fine-tuned transfer learning approach, the final layers of pre-trained convolutional neural network are replaced to fine-tune CNNs on target classes of our target ADNI dataset and predict Alzheimer's disease.

This section contains the implementation of proposed methodology consisting of pre-processing, feature extraction, training and classification (as illustrated in Fig. 1). The prior step of proposed system is preprocessing and data-augmentation. The DICOM and one channel based images converted into jpeg and three channels and then prune the extra area in the surrounding. Different data augmentation techniques applied to enhance and present various representation of the input space for the classifier. The second step is to the used pre-trained weight of various CNNs and retrained on target dataset. The fine-tuned features
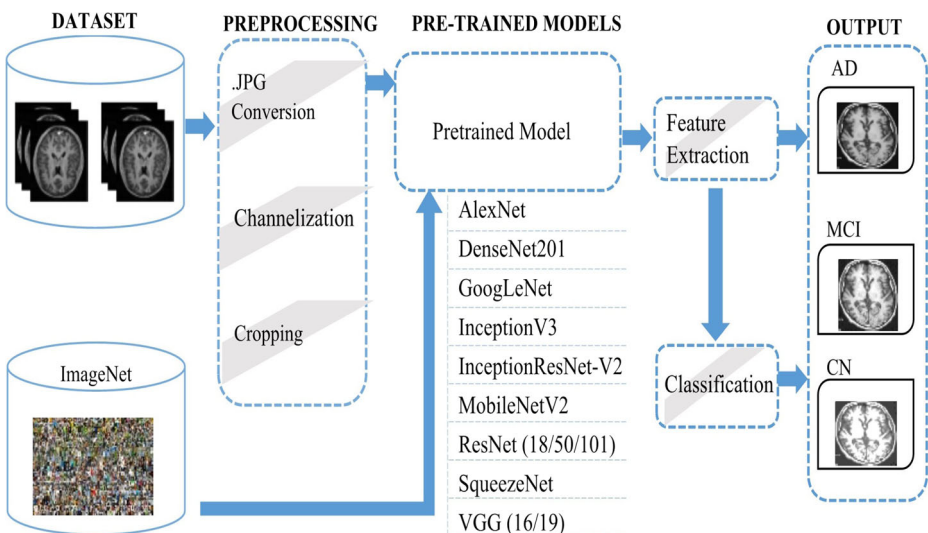


**Fig. 1** Flow diagram of proposed study for automatic Alzheimer detection

using different architectures of CNNs are fed to target classifier using ADNI dataset. The final step is the classification of target models, and an output layer presented target labels for Alzheimer disease.

## 5.1 Prepossessing

The preprocessing is like a data mining approach; it involves the transformation of raw data format into more interpret-able and suitable format for further processing. In this work, the fresh MRI images which are obtained 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. All the deep learning models (AlexNet, GoogLeNet, VGG, MobileNet, SqueezNet, ResNet and SNNs) require 3-channel images data as input. Therefore, we have converted all the data from 1-channel to 3-channel and we resized all the images according to the requirements of each model. After channelization and resizing, we cropped the data to remove the white spaces and to enhance the data. This approach is used to determine the extreme points in contours along with x,y coordinates. These techniques help target the region of interest.

DNN offers a significant boost in performance and can yield skilful models in case of more extensive training data. Data augmentation is the phenomenon of artificially increasing and enhancing the size of training data from the original dataset. This approach introduces variations in the images which improves the classification accuracy of a model. This makes the model more generalized and less prone to over-fitting.

For training deep neural network, having large dataset is crucial. Data augmentation plays a key role to significantly increase the diversity of data for training the neural networks when originally dataset is small. In this study, we collected the dataset for 350 patients which is not enough for training the deep neural network and attaining better performance. To enhance or increase the dataset, we used different augmentation techniques i.e. horizontal and vertical flipping, rotation of 90, 180 and 270, illumination, zoom in and zoom out. With these techniques, we can increase the size of the dataset. Before augmentation, a total number of images of the dataset were 3925, and after data augmentation, 37590 images were obtained.

## 5.2 Features extraction

Feature extraction is an important step in convolutional neural networks because prediction is based upon features learned by convolutional layers in deep CNN. In machine learning approaches, hand made or manually designed features were used specifically to the related problem. However, deep CNN's be able to learn features by themselves related to problems. For automatic extraction of features in deep CNN's, convolutional layers were used. Convolutional layers at the start of CNN's uses filters to detect low-level features like edges, blobs and colours. The layers at the end of the networks use filters to learn high level, global and more complicated features. Features that learned by convolutional layers simply are numbers which represent how certain pattern is located. In this study, convolution layers of CNN architectures (AlexNet,GoogLeNet, VGG (16/19), ResNet (18/50/101), MobileNet-v2, SqueezeNet, Inception-v3, Inception-ResNet-v2, DenseNet201 and third-generation Spiking Neural Networks (SNN)) be employed as feature extraction purpose, and fully connected layers be employed for classification using transfer learning approach. From the study of CNN architectures, it is found that the internal representation of CNN

layers is difficult to interpret. We can investigate the features by visualizing different convolutional layers of CNN. Here, using AlexNet, Fig. 2 shows the output of strongest activation channel for three classes (AD, MCI and CN) after Conv1, Conv2, Conv3, Conv4 Conv5, FC6, FC7, and FC8 layers of AlexNet.

## 5.3 Training and classification

In practice, it is notable that training of CNN with random initialization or from scratch is not easy because it requires a huge number of data samples. It is not possible, especially in medical image analysis, because data sets are usually small. Alternatively, pre-trained CNN on huge dataset like ImageNet has become a standard procedure. Transfer learning is the improvement of learning in a new task by transferring the knowledge learned from the existing similar task. The main advantage of transfer learning is rapid progress.

In this work, we have employed a fine-tuning approach of transfer learning on ADNI dataset. It is the enhancement of learning through the transfer of knowledge while performing a new task. We presented deep pre-trained convolutional neural networks for fine-tuning. A pre-trained model is a model that is trained on a large benchmark dataset, i.e., ImageNet. In the transfer learning approach, there are two datasets. One is called source dataset, and
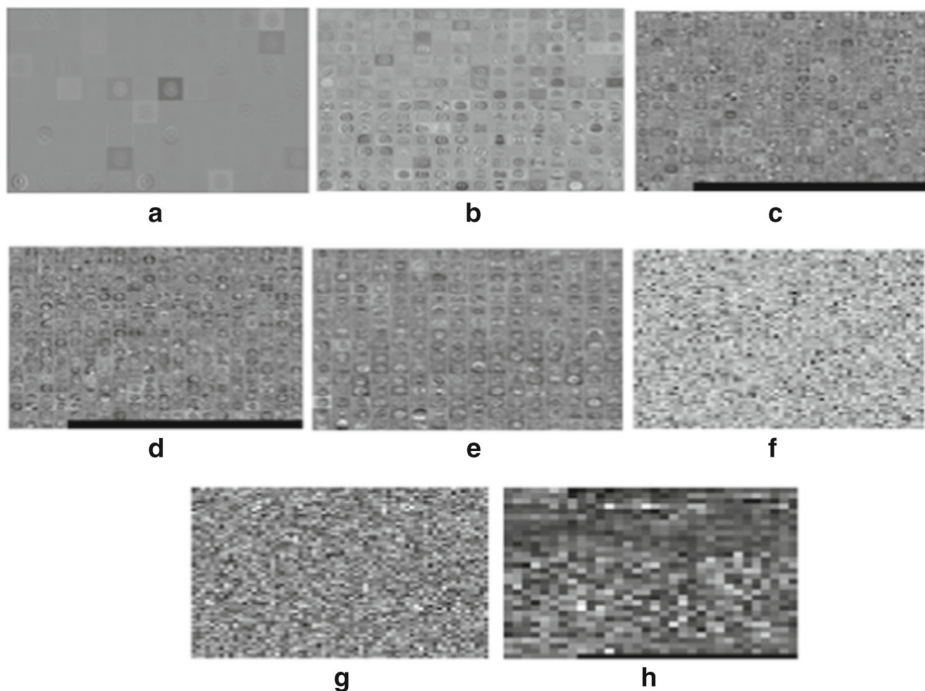


**Fig. 2** **a** Visualization of a sample of an AD patient using 96 filters in the first convolutional layer of AlexNet, each tile in image "i" the output of channel in convolution 1, A channel that is mostly grey does not participate mostly on the activation f input image. First convolutional layer of most of the CNN networks detects low-level features such as edges, blobs or colours. Going deeper in-network convolutional layers detect more complicated features **b** and **h** shows the visualization of Conv1, Conv2, Conv3, Conv4 Conv5, FC6, FC7, and FC8 layers of AlexNet. These layers show more complicated information than first convolution layer

the other is called target dataset. In our case, source dataset is ImageNet on which deep neural networks are already trained, and the target dataset is a subset of ADNI, that is described in the previous section. We passed the target dataset as input to the pertained models of deep CNN. The input is passed through different convolutional layers and these convolutional layers act as a feature extractor. The convolution layers transform the input image for feature extraction by using kernel or convolutional matrix. Automatic features were extracted, and weights of the layers were updated and stored in a new vector. These weights are then passed to the fully connected and softmax layers for the prediction of AD, CN and MCI.

To classify Alzheimer's Disease (AD), we presented pre-trained Convolutional Neural Networks, i.e., AlexNet, GoogLeNet, VGG-16/19, ResNet-18/50/101, MobileNet-V2, SqueezeNet, Inception-V3, Inception-ResNet-v2 and DenseNet201, and spiking neural network (SNN). All these models are pre-trained on ImageNet dataset (source dataset). Transfer learning fine-tuned approach is used to deploy these pre-trained networks to train the target dataset (ADNI dataset) to classify AD.

We resized all the input images to 229 x 229 pixels for inception-V3 and Inception-ResNet-v2 model, 227 x227 pixels for AlexNet and SqueezeNet model; and 224 x 224 pixels for rest of the models described above. In this study, for all experiments, the same training parameters were used. These parameters include a solver of stochastic gradient descent with momentum (sgdm) with a momentum of 0.9. The function of momentum is to make the model fast and stable by smoothing the weights in each iteration. A mini-batch size of 10 and the maximum number of epochs was 30 with an early stopping in order to meet the best network according to the validation set, 300 maximum validation frequency and with an initial learning rate of 3e-4. L2 regularization was employed to minimize the training loss. Same training, testing and validation set was used for all the experiments. All experiments were implemented using MATLAB except the SNN was implemented in Python. These experiments were performed on Windows 10, 64-bit operating system with 16 GB RAM and 8GB GTX 1070 GPU (Fig. 3).

Spiking Neural Network model is based on the proposal outlined by the winner [33] of ImageNet challenge (ILSVRC) 2012. It comprises of a relatively simple layout, i.e. 5 convolutional layers, average pooling layers, dropout layers and three fully connected lay-
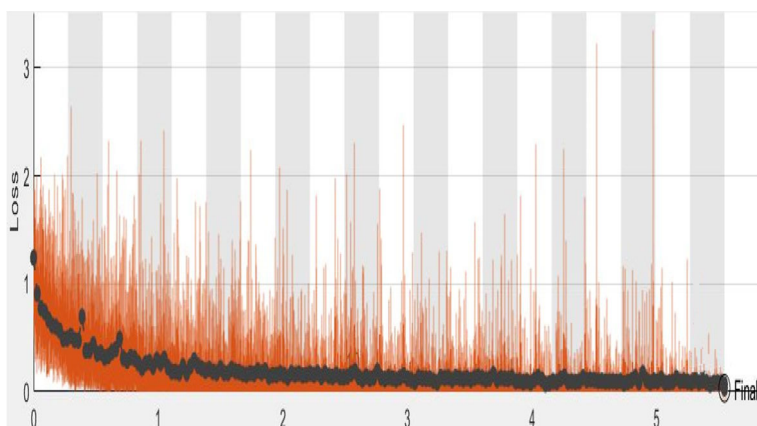


**Fig. 3** Training graph of Dense Network having least error

ers. In order to make this Artificial Neural network model (ANN) transferable to spiking neuron, few modification are required i.e. removal of local response normalization layer, changing maximum pooling layers with average pooling layers. Training with a distinguishable version of the leaky integrate-and-fire neuron. We performed a training approach for SNN termed as "constrain-then-train", that has constraints due to properties of a spiking neuron. After due training, we convert the constrained ANN into spiking Neural network and parameters of constrained ANN are utilized directly by SNN without any further scaling. We employed deep learning technique as used by i.e. SGD as optimization strategy and backpropagation for parameter gradient determination. The LIF neurons parameters $\tau_{RC}$, $\tau_{ref}$ and $\gamma$ have a significant effect on the training process. The key factor while choosing these parameters is that for initial weights and biases in the model and derivatives of non-linearity for inputs from ADNI-dataset should be around 1, which is helpful to minimize exploding gradient and vanishing gradient problems. We choose membrane time constant $\tau_{RC}$ =0.05 for making a smooth transition from normalized LIF neurons to spiking LIF during testing, $\tau_{ref}$ =0.001 which helps to reduce the effects of saturation on neurons behaviour and $\gamma$ =0.02 which reduces the distance between Spiking LIF and Normalized LIF curve. Training with a moderate amount of stochastic variation aka noise have a significant effect on spiking neurons and reduces the error in transitioning to SNN because of noise impact on the neurons output variability that SNN encounters when filtering spiking train. We added Gaussian noise distribution to approximately $\sigma$ =10 . We experimented with $\sigma$ =0 and $\sigma$ =20 but best results obtained with moderate amount of noise i.e $\sigma$ =10.

We have analyzed different pre-trained CNN architectures for the AD classification. The details of input size to model, number of layers and total parameters are shown in Table 3. In our work, the fine tuning process play a key role for increasing the performance of models. We have observed that fine tuning of the models is a very crucial step to enhance the accuracy of AD classification within acceptable times.

**Table 3** Details of pre-trained CNN architectures employed on Source dataset (ImageNet) and Target dataset (ADNI)

| Network | Model | Input size | Depth (Layers) | Parameters (million) |
|---|---|---|---|---|
| Network-1 | AlexNet | $227 \times 227$ | 8 | 62.3 |
| Network-2 | GoogLeNet | $224 \times 224$ | 22 | 6.7977 |
| Network-3 | VGGNet-16 | $224 \times 224$ | 16 | 138 |
| Network-4 | VGGNet-19 | $224 \times 224$ | 19 | 144 |
| Network-5 | MobileNetv2 | $224 \times 224$ | 54 | 3.4 |
| Network-6 | SqueezeNet | $227 \times 227$ | 18 | 5 |
| Network-7 | ResNet-18 | $224 \times 224$ | 18 | 11 |
| Network-8 | ResNet-50 | $224 \times 224$ | 50 | 25.6 |
| Network-9 | ResNet-101 | $224 \times 224$ | 202 | 44.5 |
| Network-10 | Inception V3 | $299 \times 299$ | 42 | 27 |
| Network-11 | Inception-ResNet-v2 | $299 \times 299$ | 164 | 54 |
| Network-12 | DensNet201 | $224 \times 224$ | 201 | 20 |
| Network-13 | $3^{rd}$ Generation SNN | $224 \times 224$ | 5 | 62.3 |

# 6 Results and discussion

Most of the available works on AD detection have been focused on one or two architectures of CNN using different parameters and different samples of ADNI dataset. The quantitative comparison of the different architectures of CNN becomes a difficult task by various authors in different works. In this study, we have exploited 13 architectures of deep neural networks using the same parameters and the same number of samples of ADNI dataset to detect AD through a fine-tuned approach of transfer learning. Five runs of experiments(each model) were performed on original dataset and on an augmented dataset. The different representations of samples were obtained after applying diverse augmentation techniques to increase the size of the dataset. Recently, CNNs are quite effective in different computer vision tasks, although it is very difficult to train CNN from scratch. Transfer learning fine-tuned approach requires a large amount of dataset for the best prediction. Due to this reason augmentation was done and the dataset was increased to 30740 images. After augmenting the dataset, again deep convolutional neural network was trained using the same split and other parameters as well. In this study, we just considered 379 subjects from ADNI dataset with three classes (AD, MCI, and CN). For all the experiments, the dataset is divided into 80%, 10% and 10% as training, validation and testing set respectively. We first test the performance of all pre-trained networks on the original dataset. The performance of all the models was computed in terms of sensitivity, specificity and accuracy. The results obtained from deep neural networks on the original dataset are provided in Table 4 that shows the test accuracy, validation accuracy, sensitivity, specificity and time taken for each model in training. The highest accuracy achieved by the model on the original dataset is DenseNet and VGG-19. Both the networks have achieved test accuracy of 94.13% and 93.37% respectively. In terms of efficiency, VGG-19 was very efficient as compared to the DenseNet201. VGG-19 has several features which differentiate it from the AlexNet, i.e. it uses few numbers of parameters and very small receptive fields. VGG-19 is efficient with given receptive field and small size kernel.

The results obtained from deep neural networks on the augmented dataset are provided in Table 5 that shows improvement in accuracy for all models for the augmented dataset. DenseNet201 achieved the highest average accuracy of 99.05% due to its dense architecture. In the end, further evaluation of accuracy for each class is also calculated both for original and augmented data which is shown in Table 6. Highest accuracy for each class is also achieved by DenseNet201, i.e., 98.93%, 99.02% and 99.24% for AD, CN and MCI respectively.

All the fine-tuned networks are tested on the ADNI dataset to forecast results and benchmark its performance. Fine-tuning is performed to identify its impact on classification results. We used the pre-trained models because the size of our dataset was small as compared to ImageNet dataset. The architectures of these pre-trained models are shown in table 3. Initially, we evaluated the ADNI data with AlexNet fine-tuning, then we have used the pre-trained SqueezNet, we have changed the final layer of SqueezNet to predict the AD class. The squeezNet is employed with 18 layers and 5million parameters. However, the training time of SqueezNet in our case is pretty high on the augmented data. Inception version2,3 and Inception-ResNet-v2 suffered from over-fitting that is why their performance is degraded. ResNet 18,50 and 101 were also applied for the AD classification in terms of efficiency and Accuracy ResNet-18 performed better. Overall ResNet stood behind in terms of efficiency and Accuracy from DenseNet.

**Table 4** Evaluation of classification performance of CNN based Uni-model on original ADNI dataset using fine tuned features

| Network | Model | Training Time | Test Accuracy (%) | Validation Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---------|-------|---------------|-------------------|-------------------------|-----------------|-----------------|
| Network-1 | AlexNet | 214 min | 76.53 | 77.35 | 77.77 | 88.67 |
| Network-2 | GoogLeNet | 365 min | 80.61 | 85.50 | 90.08 | 79.77 |
| Network-3 | VGGNet-16 | 154 min | 83.42 | 89.82 | 83.88 | 91.94 |
| Network-4 | VGGNet-19 | 196 min | 93.37 | 90.08 | 93.35 | 96.72 |
| Network-5 | MobileNetv2 | 1357 min | 90.05 | 87.28 | 89.17 | 94.19 |
| Network-6 | SqueezeNet | 185 min | 86.51 | 81.63 | 86.65 | 93.28 |
| Network-7 | ResNet-18 | 440 min | 89.06 | 88.78 | 89.00 | 94.60 |
| Network-8 | ResNet-50 | 1094 min | 84.69 | 79.13 | 83.80 | 92.16 |
| Network-9 | ResNet-101 | 1378 min | 78.57 | 80.41 | 78.34 | 89.21 |
| Network-10 | Inception V3 | 133 min | 79.39 | 74.49 | 72.84 | 87.21 |
| Network-11 | Inception-ResNet-v2 | 689 min | 84.18 | 83.46 | 82.70 | 91.72 |
| Network-12 | DensNet201 | 632 min | 94.13 | 91.36 | 93.07 | 97.00 |
| Network-13 | $3^{rd}$ Generation SNN | 120 min | 75.07 | 74 | 76.06 | 87.75 |

**Table 5** Evaluation of classification performance of CNN based Uni-model on Augmented ADNI dataset using fine tuned features

| Network | Model | Training time | Test accuracy (%) | Validation accuracy (%) | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Network-1 | AlexNet | 563 min | 96.98 | 97.49 | 96.99 | 98.48 |
| Network-2 | GoogLeNet | 910 min | 91.14 | 92.56 | 90.70 | 95.48 |
| Network-3 | VGGNet-16 | 1408 min | 95.38 | 95.90 | 94.99 | 97.61 |
| Network-4 | VGGNet-19 | 1484 min | 96.42 | 96.39 | 96.49 | 98.22 |
| Network-5 | MobileNetv2 | 802 min | 96.71 | 96.80 | 96.63 | 98.36 |
| Network-6 | SqueezeNet | 1479 min | 88.19 | 88.87 | 87.49 | 93.91 |
| Network-7 | ResNet-18 | 778 min | 96.60 | 96.13 | 96.33 | 98.25 |
| Network-8 | ResNet-50 | 826 min | 94.09 | 94.26 | 93.94 | 97.03 |
| Network-9 | ResNet-101 | 2480 min | 94.98 | 94.23 | 94.78 | 97.47 |
| Network-10 | Inception V3 | 917 min | 94.78 | 95.13 | 94.52 | 97.33 |
| Network-11 | Inception-ResNet-v2 | 975 min | 90.60 | 90.65 | 89.96 | 95.20 |
| Network-12 | DensNet201 | 4805 min | **99.05** | 98.79 | 99.06 | 99.53 |
| Network-13 | $3^{rd}$ Generation SNN | 186 min | 91.84 | 90.12 | 89.82 | 95.62 |

**Table 6** Experimental results of fine-tune based different CNN architectures

| Architecture | Original dataset | | | | Augmented dataset | | | |
|---|---|---|---|---|---|---|---|---|
| | AD(%) | CN(%) | MCI(%) | Overall accuracy (%) | AD(%) | CN(%) | MCI(%) | Overall accuracy(%) |
| AlexNet | 93.66 | 53.37 | 86.27 | 77.35 | 94.66 | 96.32 | 100 | 96.98 |
| GoogLeNet | 80.28 | 86.48 | 72.54 | 80.61 | 89.14 | 95.57 | 87.38 | 91.14 |
| VGGNet-16 | 80.98 | 82.43 | 88.23 | 83.42 | 94.69 | 98.48 | 91.80 | 95.38 |
| VGGNet-19 | 93.66 | 93.24 | 93.13 | 93.37 | 97.76 | 95.01 | 96.69 | 96.42 |
| MobileNetv2 | 95.77 | 87.84 | 74.51 | 90.05 | 97.78 | 96.47 | 95.65 | 96.71 |
| SqueezNet | 91.61 | 81.08 | 87.28 | 86.51 | 87.52 | 92.56 | 82.37 | 88.09 |
| ResNet-18 | 91.61 | 87.16 | 88.24 | 89.06 | 97.12 | 97.97 | 93.91 | 96.60 |
| ResNet-50 | 89.44 | 86.49 | 75.49 | 84.69 | 94.00 | 95.12 | 92.71 | 94.09 |
| ResNet-101 | 85.21 | 74.32 | 75.49 | 78.57 | 96.46 | 95.27 | 92.60 | 94.98 |
| Inception V3 | 87.32 | 74.32 | 56.86 | 74.49 | 95.81 | 95.79 | 91.95 | 94.78 |
| Inception-ResNet-v2 | 88.81 | 83.78 | 75.49 | 84.18 | 95.48 | 91.36 | 83.03 | 90.60 |
| DensNet-201 | 98.59 | 97.30 | 83.33 | 94.13 | 98.93 | 99.02 | 99.24 | **99.05** |
| $3^{rd}$ Generation SNN | 74.65 | 77.03 | 73.53 | 75.07 | 92.25 | 91.89 | 91.18 | 91.84 |

Training and validation performance of the DenseNet201 is shown in Fig. 4. DenseNet accomplished well because it is the logical extension of ResNet, DenseNet uses concatenating outputs instead of using the summation from prior layers. It is built on the idea of ResNet, but it does not add the activation generated by one layer to the other layer. DenseNet simply concatenates the activation produced. It also preserves some kind of global state instead of preserving all the information. Due to this reason, it has a lower number of parameters as compared to ResNet. In DenseNet, the current layer is connected with all its previous layers. DenseNet performs better because it has some advantages like strengthening features propagation, feature reuse, vanishing gradient problems and decreasing the number of parameters.

For the training of SNNs we convert the constrained ANN into spiking Neural network and parameters of constrained ANN are utilized directly by the SNN without any further scaling. SNN was trained by using 3500 labeled images of ADNI dataset and remaining 3500 reserved for testing. We run SNN model for 400 epochs and achieved an accuracy of 92% which are the state of the art results for SNN on ADNI dataset.

For effective treatment of AD the accurate and effective diagnosis of Alzheimer's disease is vital. The early detection of AD can play a key role in effective patient care and therapeutic development. In this work, we have reviewed different deep learning approaches based on neuro-imaging data for the investigative classification of Alzheimer's disease AD. We have analyzed different articles representing different approaches. Few of them used hybrid approaches that combined deep learning with machine learning approaches as a classifier while the rest has used only deep learning approaches. Among these techniques, few hybrid techniques have yielded up to 98% accuracy for AD classification. The deep learning approaches produced up to 96% classification accuracy along with 84.2% accuracy for MCI conversion prediction. Comparison of our results with the state of the art results on ADNI dataset in literature in terms of accuracy is shown in Table 7. We have split up our dataset into train and tests sets. For training of the models, 80% of data is allocated, while the remaining 20% data is allocated for testing and validation. The performance of five runs of DenseNet has shown in Fig. 4. DenseNet model has achieved 99.05% accuracy in the second run, which is quite promising and state of the art results on ADNI dataset.
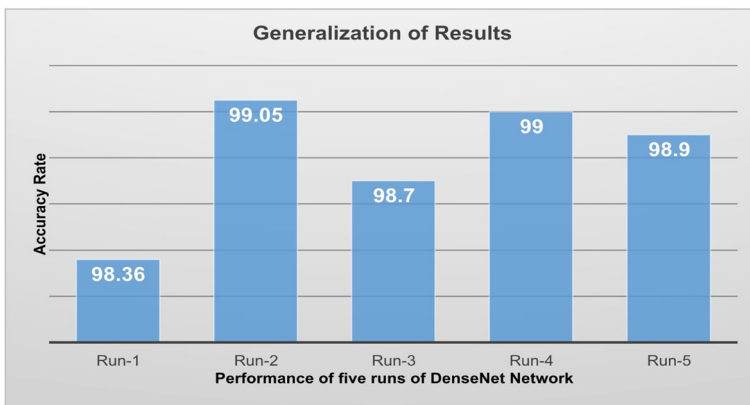


**Fig. 4** Generalization of results using leave-one-out cross validation

The results of all the models applied are presented in Tables 4 and 5. In Table 4, the results of all models are comparatively low because we have used non-augmented dataset. We have performed scaling, rotation, flipping, zooming and elimination in preprocessing phase to increase input space and to improve the classification accuracy. The preprocessing steps and data-augmentation is the crucial steps to improve the quality of images. We performed data augmentation to represent a more comprehensive set of possible data points, which has reduced the distance between training and validation sets. Rotation based augmentation is performed by rotating the image right or left on an axis between 1-359 degrees. Image translation is performed by shifting image up, down, right or left in order to avoid positional bias in the data. In medical imaging the labelling of data set requires an expert review which is time-consuming and an expensive task. That is why transfer learning is popular in medical image analysis.

We have conducted validation experiments by using leave-one-out-cross validation method. In this method, k is kept equal to N, where N separates times. In this approach function, more approximate is trained on the whole data up to a specific point and prediction is made up to that point. Average error was computed for the evaluation of the model. The performance of all the models is evaluated by calculating the sensitivity, specificity and accuracy shown in Table 4 and Table 5. The overall highest accuracy achieved in this work by using DenseNet model is 99.05% so far, which is the state of the art result for the AD classification. Another remarkable achievement of our work is that we have used biologically plausible spiking neural networks for the first time for AD classification and its results are promising and very close to the 2nd generation neural network models used in this work. SNNs exhibits favourable characteristics like low energy consumption, efficient inference and event-driven information processing. These features make SNN interesting candidates for the efficient implementation of DNN. In this work the Accuracy of SNN is more promising because the main challenge to SNN is that they do not reach the same level of accuracy as their machine learning counterparts. For SNN we have used the conversion approach in which DNNs is converted to SNN by adapting weights and parameters of SNN. This is a systematic way to convert conventionally trained CNNs to SNN. By using this approach, we can exploit the full toolkit of deep learning [17]. Despite lower accuracy SNN are much efficient and consume lower energy.

The comparison of our work with the modalities, techniques and accuracy present in literature is discussed in Table 7. ADNI dataset is used by all the approaches mentioned in the table. The performance is evaluated in terms of sensitivity, specificity and accuracy. We have run 13 different models shown in Table 6. In terms of Accuracy DenseNet outperformed all the other approaches and the approaches present in the literature. With 3way classification DenseNet so far has produced the best results, i.e. 99.05% over all accuracy. Another advantage of data augmentation is that it also alleviates the over-fitting problem. In terms of training time, spiking neural networks was fast due to its nature of lightweight. Although SNN's are efficient in terms of training time from ANN but SNN lag far behind ANN's in terms of accuracy. This gap between the accuracies of ANN and SNN is decreasing and can disappear in some tasks. In our work the gap in Accuracy of SNN and other networks is significantly closer. Class-wise comparison of all models is shown in Table 6. We have also compared our class wise results with the results present in the literature shown in Table 7. In competing approaches Ammarah et al. have performed analysis on four classes (AD, MCI, LMCI,NC) they have achieved 98.88% results, Hosseni et al. have used the three classes (AD,MCI,NC) they have achieved the accuracy up to 94.8%. Arfa et al. [5], Gupta et al. [8] and Hossein et al. [11] achieved 98/97/97, 96/74/88, 100/80/47 sensitivity, for (AD/MC/NC)

**Table 7** Class wise Comparison of results on ADNI dataset

| Author | Tehnique | Modality | Classification | Accuracy (%) |
|---|---|---|---|---|
| Backstrom et al. | 3D ConvNet | MRI | 2way (AD/CN) | 98.41 |
| Tong et al. | NGF | MRI+PET +CSF | 3way (AD/MCI/NC) | 60.2 |
| Suk et al. | DW-S MTL | MRI+PET +CSF | 3way (AD/MCI/NC) | 62.93 |
| Gupta et al. | SAE | MRI | 3way (AD/MCI/NC) | 85 |
| Hosseni et al | DSA-3DCNN | MRI | 3way (AD/MCI/NC) | 94.8 |
| Lui et al. | SAE-Zeromask | MRI+PET | 4way (AD/cMCI/ neMCI/NC) | 53.8 |
| Ammarah Farooq et al. | GoogleNet | MRI | 4way (AD/MCI/ LMCI/NC) | 98.88 |
| Ammarah Farooq et al. | ResNet-152 | MRI | 4way (AD/MCI/ LMCI/NC) | 98.14 |
| Kazemi et al. | AlexNet | MRI | 4way (AD/SMC/EMCI /LMCI/ CN) | 97.63 |
| Proposed | DenseNet | MRI | 3way (AD/MCI/CN) | 99.05 |

respectively. In our work DenseNet has produced more promising and consistent results for all the three classes the sensitivity, specificity and accuracy remained above 99%. DenseNet performs well because it follows a simple connectivity rule which naturally incorporates the properties of identity mappings, diversified depth and deep supervision. Due to these factors, it permits feature reuse which results in better learning.

## 7 Conclusion

In this article, we have used different second generation and third generation neural networks for the classification of AD. Multiple data augmentation schemes were used to improve the salient feature extraction from MRI images. We fine-tuned 13 deep learning models initially trained on ImageNet dataset. Even though our dataset is small, but due to transfer learning and effective augmentation, our results were quite impressive. We have compared our results with the state-of-the-art techniques, among which DenseNet achieved the best results. Experimental results revealed that on augmented images, DenseNet attained the classification accuracy of 99.05% for all the three classes. We also convert the constrained ANN into spiking neural network (SNN), and parameters of constrained ANN are utilized directly by the SNN without any further scaling. SNN achieved a very promising classification accuracy of 92%. As SNN are more biologically plausible and energy-efficient, but they are far behind from the second generation neural networks in terms of accuracy. The results of SNN on ADNI dataset are more closer to the second generation neural networks and state of the art results for AD classification. In future work, these results can further be improved by ensemble-based methods.

## References

1. Anand KS, Dhikav V (2012) Hippocampus in health and disease: an overview. Ann Indian Acad Neurol 15(4):239
2. Bäckström K, Nazari M, Gu IYH, Jakola AS (2018) An efficient 3d deep convolutional network for alzheimer's disease diagnosis using mr images. In: 2018 IEEE 15th International symposium on biomedical imaging (ISBI 2018). IEEE, pp 149–153

3. Ebrahimi-Ghahnavieh A, Luo S, Chiong R (2019) 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, pp 133–138

4. Emmert-Streib F, Yang Z, Feng H, Tripathi S, Dehmer M (2020) An introductory review of deep learning for prediction models with big data. Front Artif Intell 3:4. https://www.frontiersin.org/article/10.3389/frai.2020.00004, https://doi.org/10.3389/frai.2020.00004

5. Farooq A, Anwar S, Awais M, Rehman S (2017) 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, pp 1–6

6. Gautam C, Mishra PK, Tiwari A, Richhariya B, Pandey HM, Wang S et al (2020) Minimum variance-embedded deep kernel regularized least squares method for one-class classification and its applications to biomedical data. Neur Netw 123:191–216. http://www.sciencedirect.com/science/article/pii/S0893608019303934. https://doi.org/10.1016/j.neunet.2019.12.001

7. Gerstner W, Kempter R, van Hemmen JL, Wagner H (1999) Pulsed neural networks. chap. Hebbian learning of pulse timing in the barn owl auditory system. Cambridge, MIT Press. ISBN 0-626-13350-4; pp 353–377. http://dl.acm.org/citation.cfm?id=296533.296554

8. Gupta A, Ayhan MS, Maida AS (2013) Natural image bases to represent neuroimaging data. In: Proceedings of the 30th international conference on international conference on machine learning - volume 28; ICML'13. JMLR.org, pp III–987–III–994. http://dl.acm.org/citation.cfm?id=3042817.3043047

9. Harper L, Barkhof F, Scheltens P, Schott JM, Fox NC (2014) An algorithmic approach to structural imaging in dementia. J Neurol Neurosurg Psych 85(6):692–698. https://doi.org/10.1136/jnnp-2013-306285

10. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE Conference on computer vision and pattern recognition (CVPR). pp 770–778

11. Hosseini-Asl E, Gimel'farb G, El-Baz A (2016) Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network, arXiv:1607.00556

12. Hosseini-Asl E, Keynton R, El-Baz A (2016) Alzheimer's disease diagnostics by adaptation of 3d convolutional network. In: 2016 IEEE International conference on image processing (ICIP). IEEE, pp 126–130

13. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T et al (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv:170404861

14. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708

15. Hunsberger E, Eliasmith C (2015) Spiking deep networks with LIF neurons, arXiv:abs/1510.08829

16. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) Squeezenet: alexnet-level accuracy with 50x fewer parameters and < 0.5mb model size. arXiv:1602:07360

17. Iyer LR, Chua Y, Li H (2018) Is neuromorphic mnist neuromorphic? Analyzing the discriminative power of neuromorphic datasets in the time domain, arXiv:1807.01013

18. Ju R, Hu C, Zhou P, Li Q (2019) Early diagnosis of alzheimer's disease based on resting-state brain networks and deep learning. IEEE/ACM Trans Comput Biol Bioinforma (TCBB) 16(1):244–257

19. Kazemi Y, Houghten S (2018) 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, pp 1–8

20. Khan RU, Tanveer M, Pachori RB, (ADNI) ADNI (2020) A novel method for the classification of alzheimer's disease from normal controls using magnetic resonance imaging. Expert Syst, 1–22

21. Liu M, Li F, Yan H, Wang K, Ma Y, Shen L et al (2020) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease. Neuro Image 208:116459. http://www.sciencedirect.com/science/article/pii/S105381191931050X, https://doi.org/10.1016/j.neuroimage.2019.116459

22. Mathew J, Mekkayil L, Ramasangu H, Karthikeyan BR, Manjunath AG (2016) Robust algorithm for early detection of alzheimer's disease using multiple feature extractions. In: 2016 IEEE Annual India conference (INDICON). IEEE, pp 1–6

23. Naseer A, Rani M, Naz S, Razzak MI, Imran M, Xu G (2020) Refining parkinson's neurological disorder identification through deep transfer learning. Neural Comput Applic 32(3):839–854

24. Razzak MI, Imran M, Xu G (2018) Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks. IEEE J Biomed Health Inform 23(5):1911–1919

25. Razzak MI, Imran M, Xu G (2020) Big data analytics for preventive medicine. Neural Comput Applic 32(9):4417–4451

26. Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. In: Classification in BioApps. Springer, pp 323–350

27. Rehman A, Naz S, Razzak I (2020) Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities. arXiv:200409010
28. Rehman A, Naz S, Razzak MI, Akram F, Imran M (2020) A deep learning-based framework for automatic brain tumors classification using transfer learning. Circ Syst Signal Process 39(2):757–775
29. Sandeep C, Kumar AS, Susanth M (2017) The online datasets used to classify the different stages for the early diagnosis of alzheimer's disease (ad). Int J Eng Adv Technol 6(4):38–45
30. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
31. Sarraf S, Tofighi G et al (2016) Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri. arXiv:070441
32. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:14091556
33. Sutskever I, Hinton GE, Krizhevsky A (2012) Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 1097–1105
34. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4 inception-resnet and the impact of residual connections on learning. In: Thirty-First AAAI conference on artificial intelligence
35. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D et al (2015) Going deeper with convolutions. In: The IEEE conference on computer vision and pattern recognition (CVPR)
36. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
37. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB et al (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imag 35(5):1299–1312
38. Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M et al (2020) Machine learning techniques for the diagnosis of alzheimer's disease: a review. ACM Trans Multimed Comput Commun Appl, 16 (1). https://doi.org/10.1145/3344998
39. Wang S, Wang H, Shen Y, Wang X (2018) Automatic recognition of mild cognitive impairment and alzheimers disease using ensemble based 3d densely connected convolutional networks. In: 2018 17th IEEE International conference on machine learning and applications (ICMLA), pp 517–523. https://doi.org/10.1109/ICMLA.2018.00083
40. Yaqoob M, Wróbel B (2017) Very small spiking neural networks evolved to recognize a pattern in a continuous input stream. In: 2017 IEEE Symposium series on computational intelligence (SSCI), p 1–8. https://doi.org/10.1109/SSCI.2017.8285420