Convolutional neural networks for multi-class brain disease detection using MRI images

Muhammed Talo[a], Ozal Yildirim[a*], Ulas Baran Baloglu[b], Galip Aydin[c], U Rajendra Acharya[d,e,f]

[a] Department of Computer Engineering, Munzur University, Tunceli, Turkey
[b] Department of Computer Science, University of Bristol, Bristol, U.K
[c] Department of Computer Engineering, Fırat University, Elazığ, Turkey
[d] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[e] Department of Biomedical Engineering, School of Science and Technology, Singapore School of Social Sciences, Singapore
[f] International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan.

Email Address: oyildirim@munzur.edu.tr

Abstract

The brain disorders may cause loss of some critical functions such as thinking, speech and movement. So, the early detection of brain diseases may help to get the timely best treatment. One of the common methods used to diagnose these disorders is the magnetic resonance imaging (MRI) technique. Manual diagnosis of brain abnormalities is time-consuming and difficult to perceive the minute changes in the MRI images especially in the early stages of abnormalities. Proper selection of the features and classifiers to obtain the highest performance is a challenging task. Hence, deep learning models have been widely used for medical image analysis over the past few years. In this study, we have employed the AlexNet, Vgg-16, ResNet-18, ResNet-34 and ResNet-50 pre-trained models to automatically classify MR images in to normal, cerebrovascular, neoplastic, degenerative, and inflammatory diseases classes. We have also compared their classification performance with pre-trained models which are the state-of-art architectures. We have obtained the best classification accuracy of 95.23% ± 0.6 with ResNet-50 model among the *five* pre-trained models. Our model is ready to be tested with huge MRI images of brain abnormalities. The outcome of model will also help the clinicians to validate their finding after manual reading of the MRI images.

Keywords: Brain disease, MRI classification, deep transfer learning, CNN, ResNet.

1. Introduction

The magnetic resonance imaging (MRI) is a popular technique widely used to investigate the abnormalities of human organs such as the brain (Legaz-Aparicio et al., 2017, Olson and Perry, 2013). Over time, it has become more preferable than other imaging technologies due to its harmless characteristics and producing high contract images (Akkus et al., 2017). The MR devices use powerful magnets and radiofrequency pulses instead of ionizing radiation. Furthermore, due

to recent developments, it is now possible to record functional imaging of organs by using functional MR (fMR) technology (Cheng et al., 2018; Logothetis et al., 2001; Michalopoulos and Bourbakis, 2015; Saignavongs et al., 2017; Zhou et al., 2016). Medical imaging technologies now provide an enormous amount of data to the researchers from healthcare area to diagnose the diseases fast and accurately.

Various signal processing-based approaches have been proposed for MR data classification problem (Chaplot et al., 2006; El-Dahshan et al., 2010; El-Dahshan and Bassiouni, 2018; Gudigar et al., 2019; Mohsen et al., 2017; Nayak et al., 2016, 2018, 2019, Wang et al., 2015, 2016, Zhang et al., 2011, 2015). The early studies usually analyzed a small set of brain images, and as a result, the outcome of these studies could not lead to the development of more generalized solutions which can work with extensive datasets (Gudigar et al., 2019). For instance, 2D discrete wavelet transform (DWT) and principal component analysis (PCA) methods were employed on a small set of brain images to achieve very high accuracy rates (El-Dahshan et al., 2010; Zhang et al., 2011). Furthermore, the combination of PCA with stationary wavelet transform (SWT) have improved the performance of the recognition systems so that researchers can perform their experiments with more number of images (Wang et al., 2015; Zhang et al., 2015). Similarly, independent component analysis (ICA) was used on fMR imaging data (Moritz et al., 2000). In these studies, usually, PCA or a substitute method was used for the feature reduction and various techniques are used to classify the images as normal and abnormal. (Chaplot et al., 2006) preferred support vector machine (SVM) based classification to obtain high accuracy scores using 52-images. The SVM is still a popular classifier on recent MR classification studies (Gudigar et al., 2019; Nayak et al., 2018; Wang et al., 2016). (Gudigar et al. 2019) have analyzed brain images by using various multiresolution techniques like discrete wavelet transform (DWT), curvelet and shearlet transforms. They have used particle swarm optimization (PSO) and SVM classifier. They have obtained the highest classification accuracy using shearlet transform based features. AdaBoost with random forest algorithm (Nayak et al., 2016) and SVM (Nayak et al., 2017) has been used to achieve very high recognition rates.

Deep learning is a popular research area in artificial intelligence field, which gives researchers an opportunity to develop accurate solutions to the complex problems involving big data. Many

successful studies have been carried out on the classification and segmentation of brain MRI data. (Akkus et al., 2017; Bernal et al., 2018). The convolution neural networks (CNN) are deep learning approaches, have been employed widely for segmentation of brain tumors (Lyksborg et al., 2015; Pereira et al., 2016; Zikic et al., 2014) and lesion segmentation (Carass et al., 2017; Maier et al., 2017). Similarly, deep learning approaches have been used effectively in brain MRI classification procedures (Lu et al., 2019; Mohsen et al., 2017; Talo et al., 2019, Sarraf and Tofighi, 2016; Wegmayr et al., 2018). (Sarraf and Tofighi, 2016) used CNN networks for Alzheimer's disease classification with MRI and fMRI data. (Mohsen et al., 2017) provided a performance of 96.9% using deep neural network (DNN) with 22 normal and 44 abnormal brain MR images.

The elimination of hand-crafted feature extraction phase is possible with the usage of deep learning methods. Previous studies in the literature had to deal with feature extraction, and this arises performance and accuracy drawback for the developed hybrid solutions. Moreover, another major problem encountered in medical image analysis is the lack of labeled data and absence of sufficient experts to label the available data. The use of transfer learning, a deep learning technique, is important as it can be a solution to both of these problems. Transfer learning method is beneficial as these pre-trained models already know how to classify the images rather than starting a model that does not know anything about the classification task. The MR image analysis consists of a massive amount of image data as opposed to a single image data, and it is foreseeable that, it has high computational complexity with its applications (McBee et al., 2018). When the processed data is huge as in the case of medical data, the solution is either to use a small subset of data as done in the previous studies or to use transfer learning method (Litjens et al., 2017). (Talo et al., 2019) used a deep learning based transfer learning approach to classify the given brain MR image into normal or abnormal classes with 100% using 5-fold validation strategy. (Lu et al., 2019) used AlexNet pre-trained model to classify brain images as normal and abnormal. (Jain et al., 2019) proposed a deep transfer learning based Alzheimer's disease diagnosis system. They used pre-trained Vgg-16 model for their work.

In this study, a model based on deep transfer learning is developed for the classification of brain abnormalities. As compared to the previous studies in the literature, we have almost doubled the number of images used (total of 1074 images) in this study. Another significant contribution of

this study is this is the first study involving five classes of brain classes: degenerative disease, inflammatory disease, cerebrovascular disease, neoplastic disease, and normal class.

2. Material and Methods

2.1. Brain Dataset

The MR brain images used in this study were downloaded from the Harvard Medical School database (http://www.med.harvard.edu/AANLIB/). We have used the whole dataset except some slides that do not contain any part of the brain in it, such as an image with end part of a skull. All brain images have equal size of 256×256 pixels and T2-weighted in the axial plane. There are 42 different cases in the database in five different categories and they consist of whole brain images from different subjects. The brain image consists of different number of slices. In Fig.1 different slices of sarcoma is shown. We have obtained a total of 1074 images/slices from this dataset.

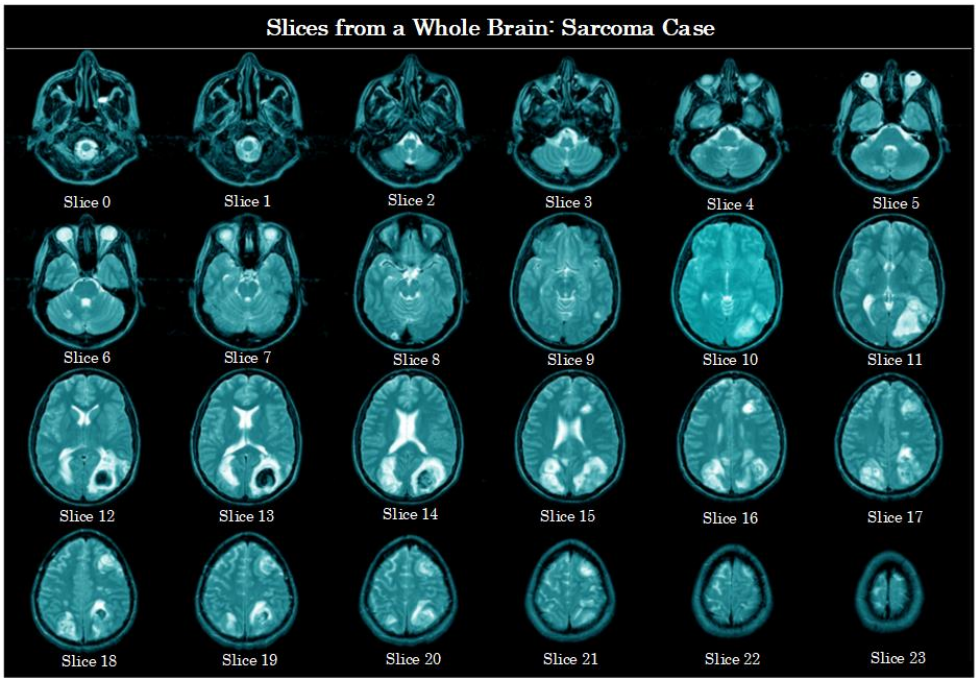

Figure 1. Sample slices of whole brain images of sarcoma class.

In this work, we have considered *five* classes for classification. The classes are normal class and *four* main disease categories: cerebrovascular (brain attack), degenerative, neoplastic (brain tumor), and inflammatory diseases. The cerebrovascular, neoplastic, degenerative, and inflammatory disease classes have 16, 8, 9, 7 subjects respectively and the normal class has *two* subjects. The cerebrovascular diseases contain multiple embolic infarction, diffusion, acute stroke (speech arrest), subacute stroke (hesitating speech), etc. types of pathologies. The neoplastic diseases include glioma, sarcoma, metastatic bronchogenic carcinoma, metastatic adenocarcinoma, and meningioma disease types. The degenerative diseases contain Alzheimer's disease, pick's disease, cerebral calcinosis, Huntington's disease, motor neuron disease, and visual agnosia diseases. Inflammatory pathologies consist of acquired immune deficiency syndrome dementia, herpes encephalitis, creutzfeld-jakob, etc. Sample MR images of five classes are shown in Fig.2.

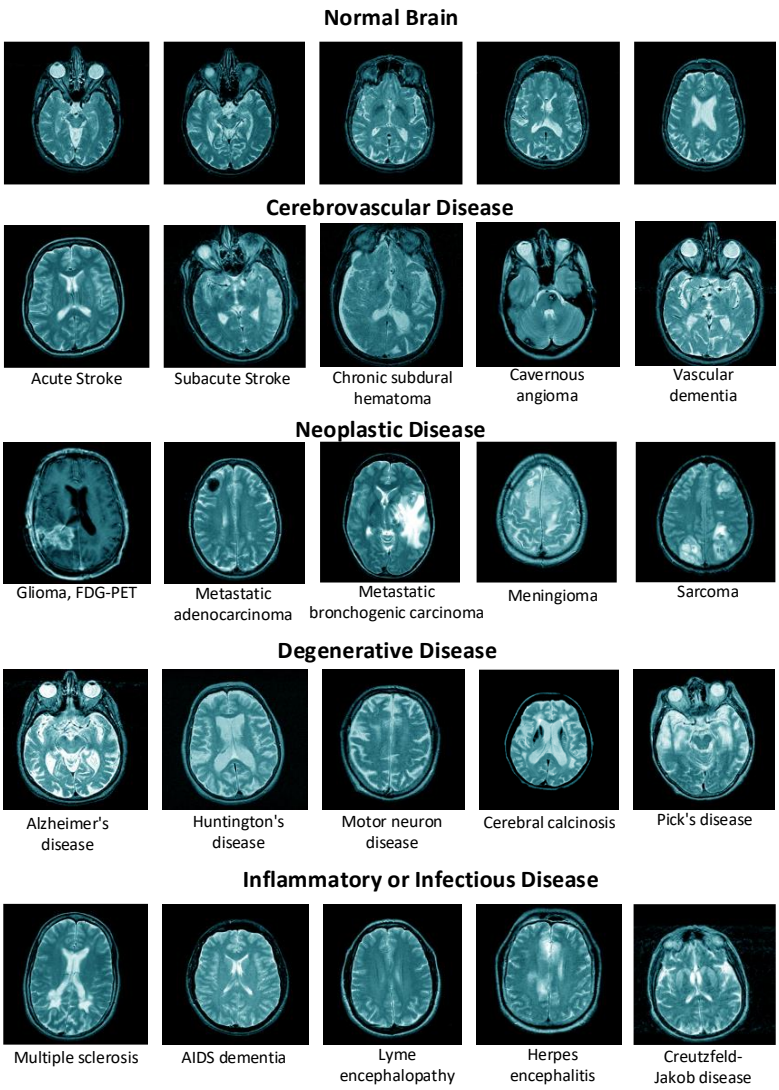

Figure 2. Sample brain MR images of five classes of this study.

2.2. Deep transfer learning

Deep learning models, especially CNNs, have shown tremendous success in classifying medical images (Lundervold et al., 2018). Training deep learning models from scratch usually requires a large amount of data and computation power. One of the main problems in the automated medical image analysis is the limited amount of labeled data. Hence, the transfer leaning technique is a solution to this problem (Talo et al., 2019; Yosinski et al., 2014). Transfer learning is the method by which the model uses the knowledge gained during the training of a relatively large dataset in a different but related problem. Many resources (GPUs) are required to train deep learning models

from scratch and the training process takes a long time. With transfer learning, we don't need to retrain the entire network from scratch for a different problem. Therefore, this technique eliminates the need for large amount of data and huge hardware requirement. In another word, the transfer learning method allows us to train deep learning models with relatively little data and it requires less computational power for training of new model. During the transfer learning process, only the classifier is trained in the new network, while the features learned from the large dataset are transferred. The transfer learning pipeline for the classification of five brain classes is given in Fig.3.

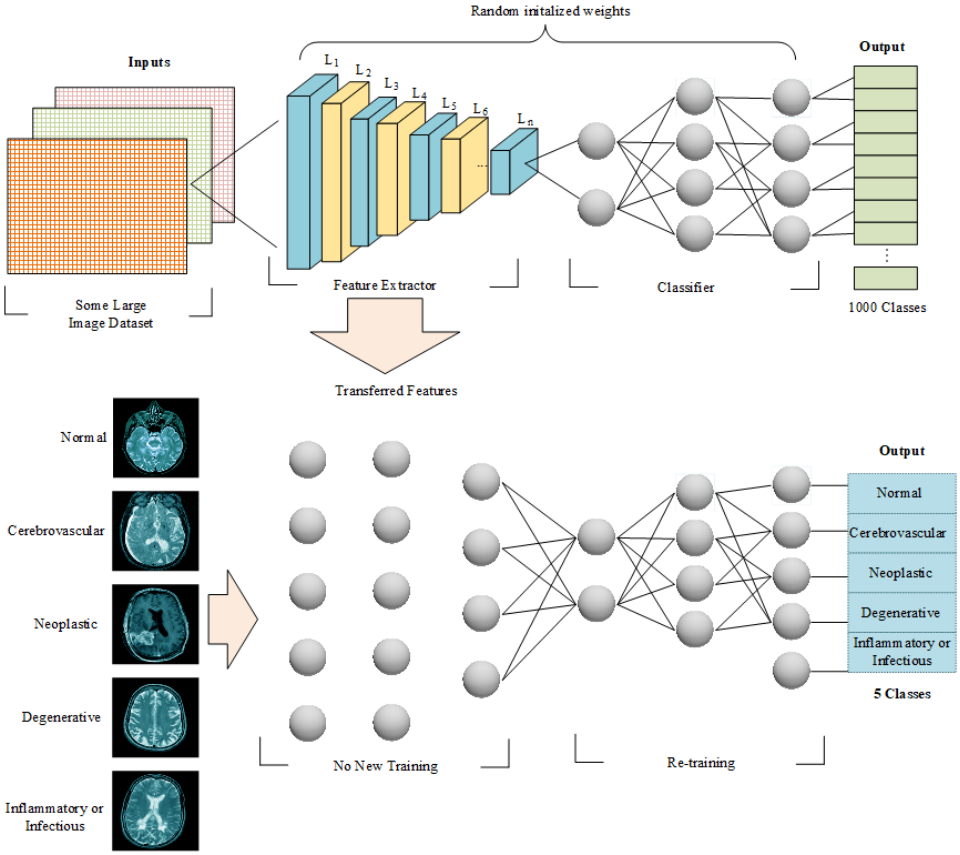

Figure 3. Schematic representation of the transfer learning technique employed for the classification of five brain classes.

In this study, we have employed the current popular deep learning architectures namely, AlexNet, VggNet, and ResNet to classify brain images into *five* different classes. Block diagrams of these pre-trained CNN models are shown in Fig.4.

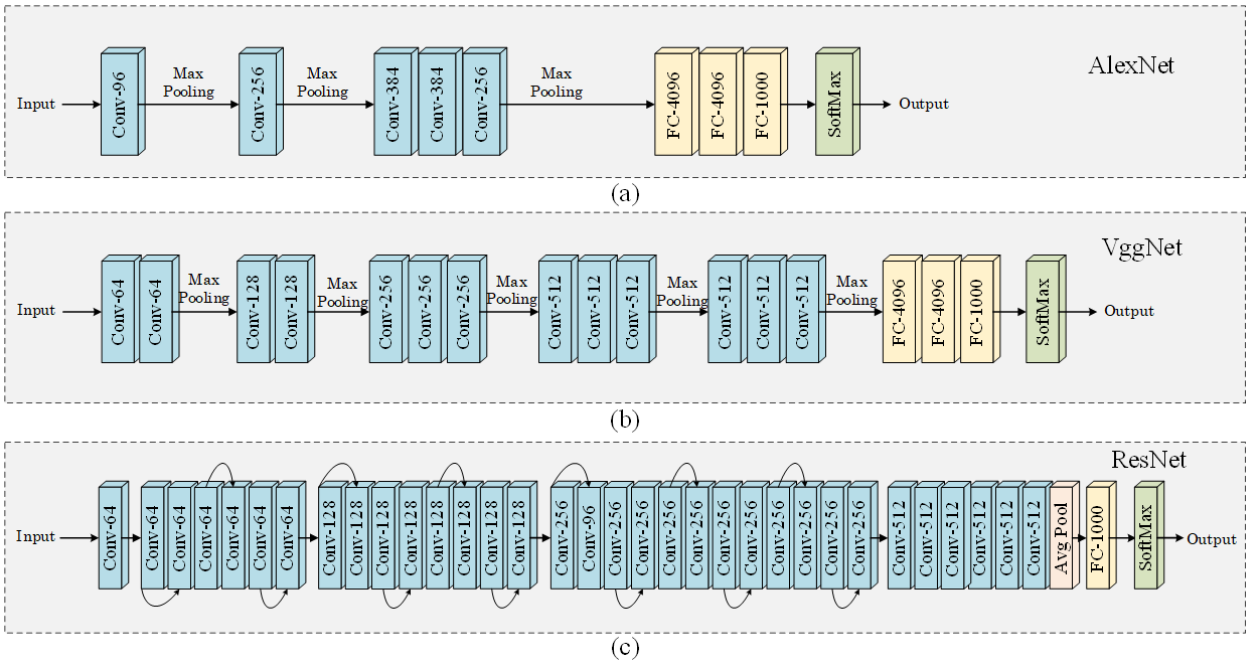

Figure 4. Block diagrams of pre-trained CNN architectures used in this work: (a) AlexNet, (b) Vgg-16, (c) ResNet

AlexNet developed by Krizhevsky, Sutskever, and Hinton (Krizhevsky et al., 2012) won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. The model dropped the top 5 error rates from 26% (second place) to 15.3% and outperformed all competitors in the competition. The sample AlexNet architecture is shown in Fig.4 (a). VggNet developed by (Simonyan and Zisserman, 2014) got the second place in the *ILSVRC-2014* challenge for the classification task with 7.3% error rate and got the first place for localization project. VggNet trained using ImageNet database has yielded improved classification accuracy as compared to AlexNet. The typical Vgg-16 architecture is given in Fig.4 (b). The Residual neural network (ResNet) was invented by (He et al., 2016) and got the first place for the classification task in the ILSVRC 2015 challenge with 3.57% error rate. ResNet architecture was able to train a neural network with 152 layers and yielded better classification performance than other best known architectures such as AlexNet and VggNet. Stacking many of residual blocks together forms the ResNet architecture and the schematic representation of the typical ResNet-34 is shown in Fig.4 (c).

3. Experiments and Results

We have compared the performance of AlexNet, Vgg-16, ResNet-18, ResNet-34 and ResNet-50 pre-trained networks in classifying the five classes of brain images. These architectures have performed well using the ImageNet database and hence can be used for medical image classification.

3.1 Experimental Setups

Convolutional neural networks (CNN) generally constructed by stacking convolutional, pooling and the fully connected layers (Krizhevsky et al., 2012; LeCun et al., 1998). Convolutional layers automatically extract hierarchical features from the images. The convolutional and poling layers are the convolutional base of CNN architecture. The classifier consists of couple of fully connected layers to classify the images based on extracted features. We have removed the last fully connected layer (classifier) of pre-trained network from the model and attached two fully connected layers to convolutional base. The last fully connected layer is connected to output *five* exits to classify brain MR images with a softmax activation function. Also, the dropout layers added between fully connected layers to prevent the overfitting. The dropout layers prevent the models from memorizing training data and generalize on the unseen data. By dropout technique, some of the neurons of the fully connected layers are ignored during training (Srivastava et al., 2014). Additionally, batch normalization layers are added to normalize layers of neural networks according to the input values. Batch normalization technique (Ioffe and Szegedy, 2015) allows the models to be trained more quickly in a stable manner by keeping all pixel values of the images in between 0 and 1. The details of the layers, output dimensions, and the parameter numbers of the proposed AlexNet model is given in Table 1.

Table 1. Details of layer types, output shapes and the number of parameters of the proposed AlexNet.

Layers	Output shape	Number of parameters
Convolutional base of AlexNet	(64, 256, 1, 1)	2,469,726

BatchNorm1d	(64, 512)	1,024
Dropout	(64, 512)	0
Fully connected	(64, 512)	262,626
ReLU	(64, 512)	0
BatchNorm1d	(64, 512)	1,024
Dropout	(64,512)	0
Fully connected	(64, 5)	2,565

We used 80% of the dataset for training and 20% for validation sets. We have employed 5-fold cross validation to compare the performance of each pre-trained model. Each pre-trained model was trained 5 times and the classification accuracy was calculated for each training. Then for all pre-trained models, the average of the classification performances is calculated.

We have used two steps to training the brain MR images and is given below:

(i) In stage-1 of training, we have only trained the attached few layers of the models for 15 epochs and set the learning rate hyper-parameter to the value of 1e-3. The initial training of the model runs fast because we only trained the last few attached layers and keep the rest of the layers frozen. So the only computation cost is required for the newly added layers.

(ii) In stage-2 process, we fine-tune both convolutional base and fully connected layers by unfreezing all layers. In this work, fine-tuning term is used to unfreeze all the layers of pre-trained models. In other words, initially, only the attached layers of pre-trained models are trained for 15 epochs, then the whole network is unfreezed to re-train all layers. At this stage, we have retrained the models for 15 more epochs by setting learning rate values in the range of [3e-6, 3e-4]. Since the training of the previous layers is mostly completed, there is no point in training all layers of a model at the same rate. Therefore, we have set the learning rate gradually during training. The learning rate of 3e-6 (low) is chosen for the first few layers and 3e-4 (high) for end layers. The learning rate for the middle layers is chosen in between 3e-4 and 3e-6 (moderate). An illustration of the whole training process is shown in Fig.5.

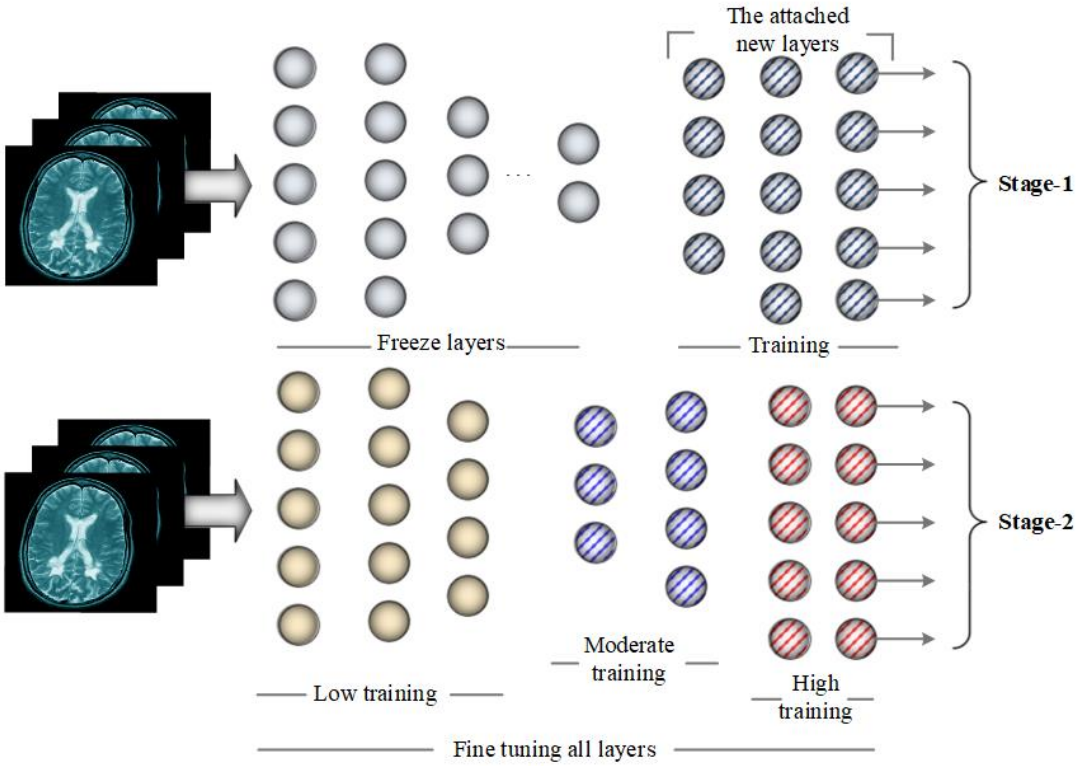

Figure 5. The training process of pre-trained model to classify brain MR images. Low, moderate and high values indicate the magnitude of the learning rate.

We have used PyTorch-v1 (Ketkar, 2017) deep learning framework with Python language to implement all pre-trained models. The RMSProp, stochastic gradient descent (SGD) optimizer, is used for parameter updates during training. The dropout rate after the first and the second fully connected layers are chosen to be 0.25 and 0.50, respectively. Hyper-parameter values of batch normalization were adjusted to 1e-5 for epsilon and 0.1 for momentum. All the models trained on NVIDIA GeForce GTX 1080 TI graphics card with 11 GB memory.

3.2 Results

During stage-1 training, only the attached layers of AlexNet, Vgg-16, ResNet-18, ResNet-34, and ResNet-50 pre-trained CNN models are activated to classify MR brain images for 15 epochs. In the Stage-2 training, every single layer of the model is activated for another 15 epochs. For each

pre-trained model, the loss and accuracy graphs for validation sets are shown in Fig.6 for a single fold.

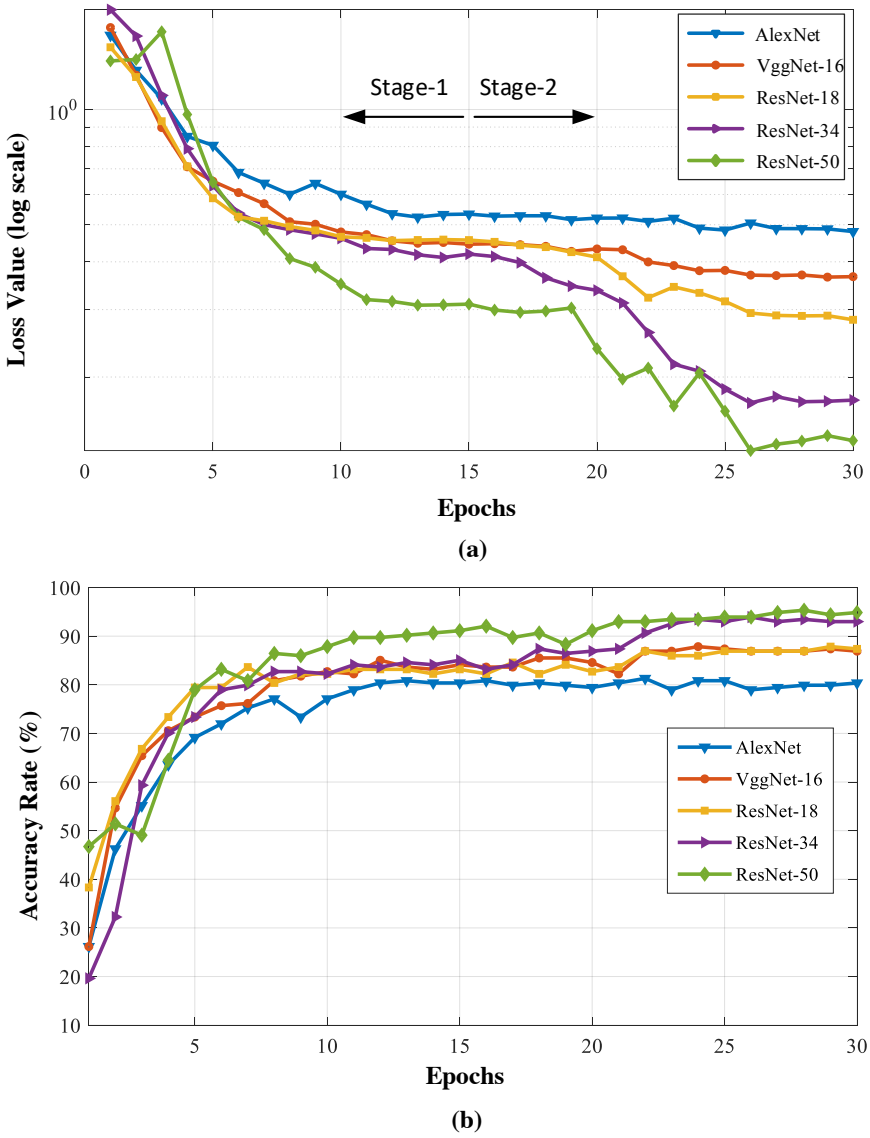

Figure 6. The performance of AlexNet, Vgg-16, ResNet-18, Resnet-34, and Resnet-50 architectures for fold-1: (a) validation losses, and (b) validation accuracies.

It can be seen from Fig.5 that, the best classification accuracy is achieved with the ResNet-50 model and the lowest performance was obtained with the AlexNet pre-trained model. Resnet-50 is modern architecture and has deeper layers as compared to AlexNet. Stage-2 process shows better classification performance as compared to Stage-1 due to the contribution of each layer. While the

ResNet-50 model showed the best improvement in Stage-2 step, this success was not observed in AlexNet architecture at the same stage. Additionally, the loss values of the models in Stage-1 are constant after the 10th epoch, while there is an obvious decline after 20th epoch in Stage-2. The most important reason for this fine-tuning process in Stage-2 is because the models unfreeze during this step and are trained moderately. This apparent decline in loss values is more pronounced in ResNet architectures. The classification accuracies of pre-trained models for 5 folds and the average time cost are shown in Table 2.

Table 2. Classification results obtained using different pre-trained models with average training time.

Models	**Average training time**	**Accuracy rate (%)**					
		Fold-1	**Fold-2**	**Fold-3**	**Fold-4**	**Fold-5**	**Overall**
AlexNet	133 s	80.37	76.64	84.11	81.78	77.57	80.09 ±3
Vgg-16	249 s	86.92	82.71	89.25	86.45	85.98	86.26 ±2.3
ResNet18	131 s	87.38	85.51	90.65	89.25	84.58	87.47 ±2.5
ResNet34	136 s	92.99	86.92	91.59	92.52	88.32	90.46 ±2.6
ResNet50	248 s	94.86	95.33	96.26	94.39	95.33	95.23 ±0.6

It can be seen from Table 2 that, the lowest classification performance (80.11% ± 3) is obtained with AlexNet model and the highest performance was obtained for the ResNet-50 model with 95.23% ± 0.6 overall accuracy. In terms of time cost, Vgg16 and ResNet-50 models have the highest train time. The main reason for this is ResNet-50 model is the largest deep learning structure among other pre-trained models (AlexNet, Vgg-16, ResNet-18, and ResNet-34). Since the ResNet-50 model allows the information flow through network with residual connections, that is, the gradient value does not diminish through back propagation, and the deepest structures have the best classification performance.

Fig.7 shows the confusion matrixes obtained using the validation data set obtained for the fold-3 of AlexNet, Vgg-16, ResNet-18, and ResNet-50 pre-trained models.

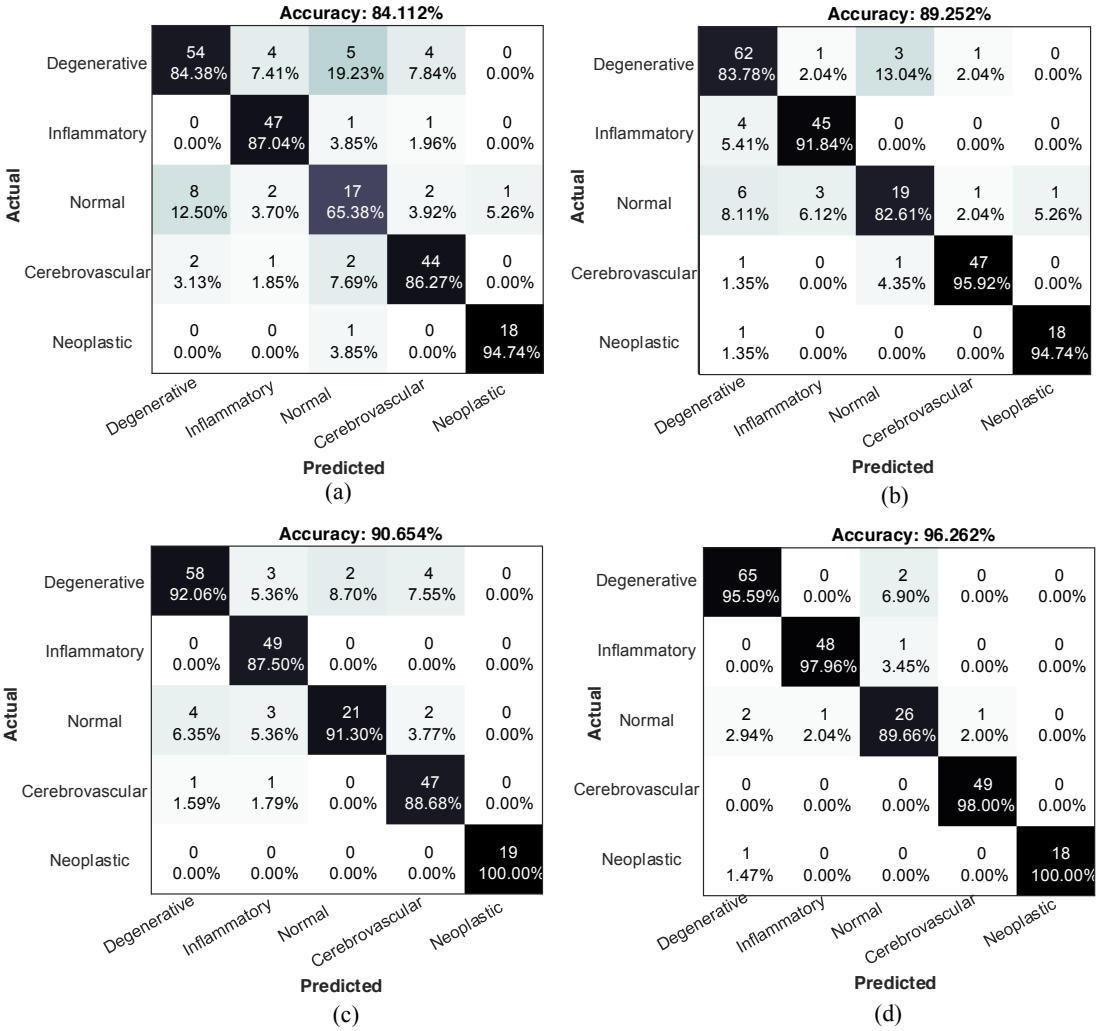

Figure 7. The confusion matrixes obtained using pre-train models with Fold-3 datasets: (a) AlexNet, (b) Vgg-16, (c) ResNet-18, and (d) ResNet-50. The percentages shown in cells of confusion matrices are column-wise percentages of TP/total number of predicted and FP/total number of predicted for each class.

The neoplastic disease group has the highest classification accuracy using MR images. The images with the lowest performance are normal brain images. While AlexNet, categorized only 17 normal brain images, the Vgg-16 model classified 19 images, ResNet-18 classified 21 images, and ResNet-50 classified 26 normal brain images out of 30 normal images correctly. The lowest classification accuracy of 84.1% and the highest classification accuracy of 96.26% are obtained using AlexNet and ResNet-50 models respectively for this fold.

For the detailed performance evaluation of the pre-trained models' validation, the commonly used precision, sensitivity, specificity, accuracy, and F1-Score values are calculated. The mean and

standard deviation values of performance parameters for AlexNet, Vgg-16, and ResNet-50 pre-trained models are given in Table 3.

Table 3. The average (mean ± standard deviation) values of the models for the 5-folds using the validation set.

	Classes	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score
AlexNet	Degenerative	84.76 ±2.9	81.88 ±5	76.72 ±2.7	89.68 ±3.2	79.18 ±3.4
	Inflammatory	90.68 ±3	71.73 ±10.1	95.59 ±3.5	89.29 ±3.3	81.64 ±7.5
	Normal	89.27 ±1.6	77.19 ±6.7	51.21 ±5.3	96.91 ±1.3	61.23 ±3.4
	Cerebrovascular	92.84 ±1	85.94 ±4.2	85.29 ±4.4	95.42 ±1.3	85.48 ±1.8
	Neoplastic	98.27 ±0.6	87.65 ±4.3	96.88 ±2.8	98.45 ±0.6	91.96 ±2.2
Vgg-16	Degenerative	89.86 ±1.8	86.28 ±4.4	86.89 ±5.4	91.83 ±2.7	86.40 ±2.1
	Inflammatory	94.96 ±0.9	85.33 ±4.6	92.81 ±4.2	95.65 ±1.1	88.78 ±2.2
	Normal	92.49 ±1.8	84.96 ±2.5	64.61 ±6.9	97.86 ±0.3	73.29 ±5.3
	Cerebrovascular	94.10 ±2.6	85.18 ±7.1	91.02 ±5.8	95.08 ±2.7	87.84 ±5.1
	Neoplastic	99.14 ±0.2	93.32 ±2.9	97.84 ±2.9	99.29 ±0.2	95.48 ±1.6
ResNet-50	Degenerative	96.14 ±1	94.37 ±1.2	94.11 ±3.1	97.16 ±0.7	94.21 ±1.6
	Inflammatory	97.89 ±0.9	91.16 ±4.4	99.59 ±0.9	97.47 ±1.3	95.13 ±2
	Normal	97.71 ±0.8	97.31 ±4.5	85.82 ±5	99.56 ±0.7	91.08 ±3.2
	Cerebrovascular	99.03 ±0.3	97.77 ±0.7	98.52 ±0.8	99.21 ±0.3	98.15 ±0.6
	Neoplastic	99.51	100	94.50 ±0.1	100	97.17

It can be seen from Table 3 that the lowest sensitivity values have emerged for normal brain image class. The sensitivity values of the normal class are 51.21% ± 5.3, 64.61% ± 6.9 and 85.82% ± 5 for AlexNet, Vgg-16, and ResNet-50, respectively. Using ResNet-50, the performance values of all *five* classes have increased significantly.

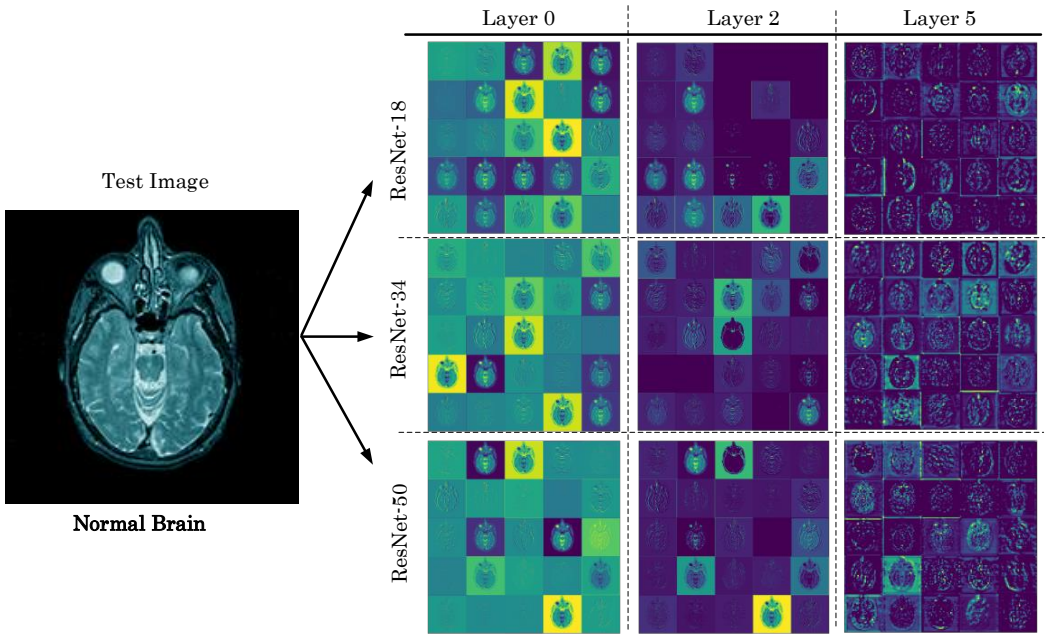

Figure 8. Sample image results of some layer activations on input test images obtained using pre-trained ResNet models: a) activations of normal brain test input image, and b) activations of brain test input image with neoplastic disease.

It can be seen from Fig. 8 that, in the initial layer, activations retain most of the image information present in brain image. The first layer activations serve as edge detectors. As the layer depth increases, the extracted properties become more abstract. Activations in the higher layers carry less information about specific data seen in the input image but carry more information about their individual classes.

Further studies on subject wise classification:

In this section, we have first performed subject-wise classification using a single Alzheimer subject belonging to degenerative class derived from Whole Brain Atlas dataset and then downloaded the Open Access Series of Imaging Studies (OASIS) dataset (Marcus et al., 2007) to examine the performance of the proposed deep model on various Alzheimer subjects. There are *four* Alzheimer subjects in the Whole Brain Atlas dataset. We have used *three* subjects for training and one for testing. The Alzheimer subject in the test set has 21 brain images. Resnet-50 pre-trained model has been retrained without seeing the test subject. The trained model has classified 14 out of 21 test images correctly and hence obtained the model performance of 66.67% for the single Alzheimer subject. The performance of the pre-trained model was not at the desired level as compared to the overall performance of the model due to limited number of Alzheimer subjects in the training set. Therefore, we increased the number of Alzheimer subjects by downloading from the OASIS dataset. We have acquired 3200 Alzheimer brain images from OASIS database. 2560 of these images were added to the training set and 640 of them were added to the test set. After this addition, the model was retrained and tested on 661 (640 from OASIS+21 from Whole Brain Atlas) images which had not seen before. The performance graphs of ResNet-50 pre-trained model are given in Fig. 9.

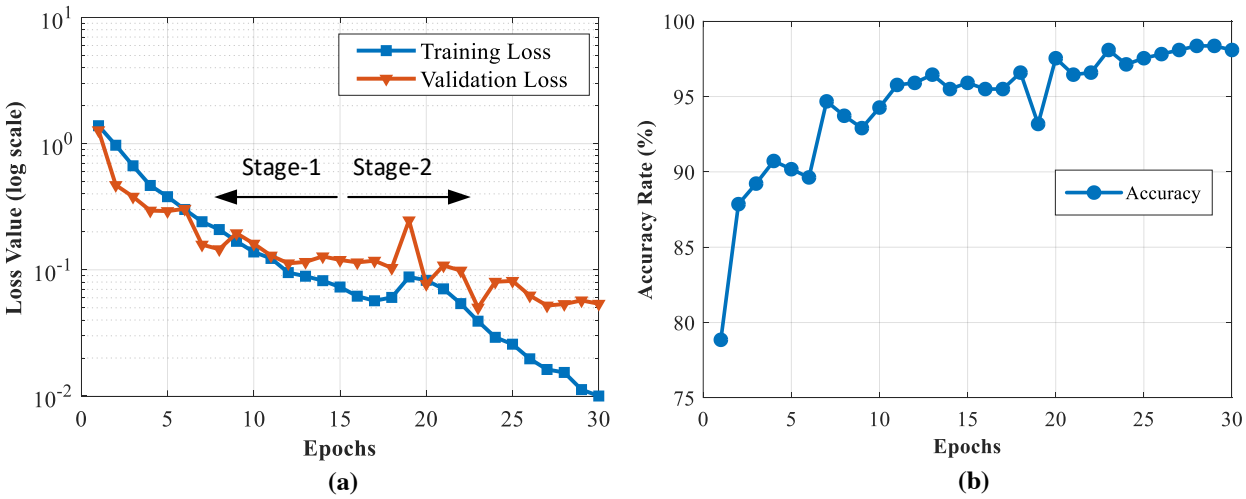

Figure 9. The performance graphs of ResNet-50 pre-trained model on OASIS added brain dataset, a) loss graphs b) accuracy graphs.

Resnet-50 model correctly classified 649 out of 661 Alzheimer test images (Degenerative class). This shows that the classification performance of ResNet-50 model for the Alzheimer images is 98.18%. The classification performance of ResNet-50 pre-trained model is significantly improved due to increased brain MR images.

4. Discussion

Various studies have been conducted using different feature extraction/selection and classification algorithms to classify brain MRI data. Table 4 presents the state-of-the-art studies conducted using the same brain MRI data. We have used all available data to classify MR images into five classes. The previous studies have been conducted using various number of MR images from the same dataset (the whole brain atlas) to classify MR images into various categories. In this study, we have used more data than the existing studies. The entire brain atlas dataset is not suitable for patient-level classification due to less number of patient images. However, in this study like rest of the studies, we have used the same dataset and implemented image-level classification.

Table 4. Comparison of the state-of-art studies conducted using the same dataset.

Study	Number of Used Images (slices)	Num. of Classes	Method and Classifier	Accuracy (%)
(Chaplot et al., 2006)	• Normal: 6 • Abnormal: 46	2	DWT+SOM+SVM	98
(El-Dahshan et al., 2010)	• Normal: 10 • Abnormal: 60	2	DWT + PCA + KNN	98
(Zhang et al., 2011)	• Normal: 18 • Abnormal:48	2	WT+PCA+BPNN	100
(Nayak et al., 2016)	• Normal:35 • Abnormal:220	2	DWT +PCA + ADBRF	99.53
(Gilanie et al., 2018)	• Normal:14 • Abnormal: 87	2	Gabor Texture + SVM	100
(El-Dahshan and Bassiouni, 2018)	• Normal: 30 • Abnormal:60	2	SWT+PCA+KLDA+ KNN	100
(Hooda and Verma, 2019)	• Normal: 24 • Abnormal:56	2	LBP+MCM	98.75
(Gudigar et al., 2019-a)	• Normal: 83 • Abnormal: 529	2	Shearlet transform + Texture + PSOSVM	97.38
(Talo et al. 2019)	• Normal:27 • Abnormal: 513	2	Deep Transfer Learning (ResNet-34)	100
(Lu et al., 2019)	• Normal: 114 • Abnormal: 177	2	Deep Transfer Learning (AlexNet)	100
(Gudigar et al., 2019-b)	• Normal: 83 • Degenerative: 132 • Inflammatory:114 • Cerebrovasc.: 187 • Neoplastic: 96	5	VMD+SVMN	98.20
The Proposed	• **Normal: 82** • **Degenerative: 209** • **Inflammatory:172** • **Cerebrovasc.: 373** • **Neoplastic: 238**	**5**	**Deep Transfer Learning (ResNet-50)**	**95.23 ±0.6**

It can be noted from Table 4 that, DWT (Chaplot et al., 2006; El-Dahshan et al., 2010; Nayak et al., 2016) , Gabor texture (Gilanie et al., 2018), local binary pattern (LBP) (Hooda and Verma, 2019), stationary WT (SWT) (El-Dahshan and Bassiouni, 2018), kernel linear discriminator analysis KLDA) (El-Dahshan and Bassiouni, 2018), shearlet transform (Gudigar et al., 2019-a) methods and variational mode decomposition (VMD) (Gudigar et al., 2019-b) were used for feature extraction

from images. The PCA (El-Dahshan et al., 2010; El-Dahshan and Bassiouni, 2018; Nayak et al., 2016; Zhang et al., 2011) method was applied for the reduction of features dimensions. In the final stage for classification SVM (Chaplot et al., 2006; Gilanie et al., 2018; Gudigar et al., 2019-a; Gudigar et al., 2019-b), minimal complexity machine (MCM) (Hooda and Verma, 2019), k-neighbor neighbor (KNN) (El-Dahshan et al., 2010; El-Dahshan and Bassiouni, 2018), back-propagation neural networks (BPNN) (Zhang et al., 2011) and adaboost with random forests (ADBRF) (Nayak et al., 2016) classifiers have been used. On the other hand, recently all of these processes are carried out with an end-to-end structure using deep learning architectures using the same data set (Lu et al., 2019; Talo et al., 2019). In these studies, popular models such as AlexNet and ResNet, which were previously trained on large image data, were used. The main reason to use these pre-trained architectures is that these models were developed using small number of brain MRI data as can be seen from Table 4. In this study, we have used a transfer learning-based approach to classify MRI data in to normal and *four* different disease classes. The main advantages of our work are summarized as follows:

(I) Generally, most of the studies are focused on the classification of normal and abnormal MRI. In this study, normal and *four* main brain diseases (degenerative, inflammatory, cerebrovascular and neoplastic) are classified. This is the first study involving *five* classes of brain MRI images with deep learning method.

(II) A total of 1074 MRI data are used, i.e., all the available images from the Harvard Medical School database are used. Hence, we have used the maximum number of images (Table 4).

(III) Our method is completely automated and does not involve any feature extraction, selection or classification steps.

(IV) The performance of popular deep learning models: AlexNet, Vgg-16, ResNet-18, ResNet-34 and ResNet-50 are compared.

Subject-wise experimental studies have shown that the proposed deep model has yielded five class classification accuracy of 66.67% using a single Alzheimer subject. The performance of the proposed model is not satisfactory. The main reason for this is the database used contains few subjects in each class. Therefore, we have used Alzheimer disease in the Degenerative class for the subject-wise classification. We have also shown that, the performance of ResNet-50 model

significantly improved using additional Alzheimer images from OASIS database. Based on this result, we can say that increasing the images of different cases increases the performance of the proposed model. We could not extend the subject-wise experiments in the study because the images in the databases were not fully compatible. The absence of a large set of subjects in each class limited the study.

The main disadvantage of this study is fewer number of available annotated data. Therefore, we have used transfer learning technique to overcome limited number of data. It has been proposed to consider the functional strength of the brain disease cohort to model the network for classification or clustering to obtain more meaningful results (Zhang et al., 2017; Zhang et al., 2018). Similarly, the use of approaches such as simultaneous constrained matrix and tensor factorization can help to extract unique properties from image (Zhou et al., 2016). We can implement such approaches which will highlight individual characteristics in inter-subject studies when more data is available.

5. Conclusion

In this study, we employed the AlexNet, Vgg-16, Resnet-18, Resnet-34, and ResNet-50 deep learning pre-trained models to classify MR images into normal, cerebrovascular, neoplastic, degenerative and inflammatory categories. These models are tested using Harvard Medical School MR brain images database using all available 1074 images. Our results show that ResNet-50 model obtained the highest accuracy of 95.23% ± 0.6 among the five models. The AlexNet has attained the lowest performance in classifying images into *five* categories.

Conflict of interest

The authors declare no conflict of interest.

6. References

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D.L., Erickson, B.J., 2017. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. J. Digit. Imaging. https://doi.org/10.1007/s10278-017-9983-4

Bernal, J., Kushibar, K., Asfaw, D.S., Valverde, S., Oliver, A., Martí, R., Lladó, X., 2018. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. Artif. Intell. Med. https://doi.org/10.1016/j.artmed.2018.08.008

Carass, A., et al., 2017. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. Neuroimage. https://doi.org/10.1016/j.neuroimage.2016.12.064

Chaplot, S., Patnaik, L.M., Jagannathan, N.R., 2006. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomed. Signal Process. Control. https://doi.org/10.1016/j.bspc.2006.05.002

Cheng, L., Zhu, Y., Sun, J., Deng, L., He, N., Yang, Y., Ling, H., Ayaz, H., Fu, Y., Tong, S., 2018. Principal States of Dynamic Functional Connectivity Reveal the Link Between Resting-State and Task-State Brain: An fMRI Study. Int. J. Neural Syst. https://doi.org/10.1142/S0129065718500028

El-Dahshan, E.S.A., Bassiouni, M.M., 2018. Computational intelligence techniques for human brain MRI classification. Int. J. Imaging Syst. Technol. https://doi.org/10.1002/ima.22265

El-Dahshan, E.S.A., Hosny, T., Salem, A.B.M., 2010. Hybrid intelligent techniques for MRI brain images classification. Digit. Signal Process. A Rev. J. https://doi.org/10.1016/j.dsp.2009.07.002

Gilanie, G., Bajwa, U.I., Waraich, M.M., Habib, Z., Ullah, H., Nasir, M., 2018. Classification of normal and abnormal brain MRI slices using Gabor texture and support vector machines. Signal, Image Video Process. https://doi.org/10.1007/s11760-017-1182-8

Gudigar, A., Raghavendra, U., San, T.R., Ciaccio, E.J., Acharya, U.R., 2019-a. Application of multiresolution analysis for automated detection of brain abnormality using MR images: A comparative study. Futur. Gener. Comput. Syst.https://doi.org/10.1016/j.future.2018.08.008

Gudigar, A., Raghavendra, U., Ciaccio, E. J., Arunkumar, N., Abdulhay, E., Acharya, U.R. 2019-b. Automated categorization of multi-class brain abnormalities using decomposition techniques with MRI images: A comparative study. IEEE Access, 7, 28498-28509.

Harvard Medical School Data, The whole brain atlas, http://www.med.harvard.edu/AANLIB/.

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. https://doi.org/10.1109/CVPR.2016.90

Hooda, H., & Verma, O. P., 2019. Classification of magnetic resonance brain images using local binary pattern as input to minimal complexity machine, In Computing, Communication and Signal Processing (pp. 883-893). Springer, Singapore

Ioffe, S., and Christian, S., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167

Jain, R., Jain, N., Aggarwal, A., & Hemanth, D. J., 2019. Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images, Cogn. Syst. Res. https://doi.org/10.1016/j.cogsys.2018.12.015.

Ketkar, N., 2017. Introduction to PyTorch, in: Deep Learning with Python. https://doi.org/10.1007/978-1-4842-2766-4_12

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. 1 ImageNet Classification with Deep Convolutional Neural Networks. Adv. Neural Inf. Process. Syst. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2323. https://doi.org/10.1109/5.726791

Legaz-Aparicio, A.-G., Verdú-Monedero, R., Larrey-Ruiz, J., Morales-Sánchez, J., López-Mir, F., Naranjo, V., Bernabéu, Á., 2017. Efficient Variational Approach to Multimodal Registration of Anatomical and Functional Intra-Patient Tumorous Brain Data. Int. J. Neural Syst. https://doi.org/10.1142/S0129065717500149

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Med. Image Anal. https://doi.org/10.1016/j.media.2017.07.005

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature. https://doi.org/10.1038/35084005

Lu, S., Lu, Z., Zhang, Y.D., 2019. Pathological brain detection based on AlexNet and transfer learning. J. Comput. Sci. https://doi.org/10.1016/j.jocs.2018.11.008

Lundervold, A. S., & Lundervold, A., 2018. An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik, https://doi.org/10.1016/j.zemedi.2018.11.002

Lyksborg, M., Puonti, O., Agn, M., Larsen, R., 2015. An ensemble of 2D convolutional neural networks for tumor segmentation, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-19665-7_17

Maier, O., et al, 2017. ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. https://doi.org/10.1016/j.media.2016.07.009

Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., Buckner, R. L., 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. https://doi.org/10.1162/jocn.2007.19.9.1498

McBee, M.P., Awan, O.A., Colucci, A.T., Ghobadi, C.W., Kadom, N., Kansagra, A.P., Tridandapani, S., Auffermann, W.F., 2018. Deep Learning in Radiology. Acad. Radiol. https://doi.org/10.1016/j.acra.2018.02.018

Michalopoulos, K., Bourbakis, N., 2015. Combining EEG Microstates with fMRI Structural Features for Modeling Brain Activity. Int. J. Neural Syst. https://doi.org/10.1111/imcb.12212

Mohsen, H., El-Dahshan, E.-S.A., El-Horbaty, E.-S.M., Salem, A.-B.M., 2017. Classification using Deep Learning Neural Networks for Brain Tumors. Futur. Comput. Informatics J. https://doi.org/10.1016/j.fcij.2017.12.001

Moritz, C.H., Haughton, V.M., Cordes, D., Quigley, M., Meyerand, M.E., 2000. Whole-brain functional MR imaging activation from a finger-tapping task examined with independent component analysis. Am. J. Neuroradiol.

Nayak, D. R., Dash, R., Majhi, B., Acharya, U. R. 2019. Application of fast curvelet Tsallis entropy and kernel random vector functional link network for automated detection of multiclass brain abnormalities. Comput. Med. Imag. Grap., 77, 101656.

Nayak, D.R., Dash, R., Majhi, B., 2018. Pathological brain detection using curvelet features and least squares SVM. Multimed. Tools Appl. https://doi.org/10.1007/s11042-016-4171-y

Nayak, D.R., Dash, R., Majhi, B., 2017. Stationary Wavelet Transform and AdaBoost with SVM Based Pathological Brain Detection in MRI Scanning. CNS Neurol. Disord. Drug Targets. https://doi.org/10.1111/j.1365-2966.2006.10753.x

Nayak, D.R., Dash, R., Majhi, B., 2016. Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. Neurocomputing. https://doi.org/10.1016/j.neucom.2015.11.034

Olson, L.D., Perry, M.S., 2013. Localization of epileptic foci using multimodality neuroimaging. Int. J. Neural Syst. https://doi.org/10.1142/S012906571230001X

Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2016.2538465

Saignavongs, M., Ciumas, C., Petton, M., Bouet, R., Boulogne, S., Rheims, S., Carmichael, D.W., Lachaux, J.-P., Ryvlin, P., 2017. Neural activity elicited by a cognitive task can be detected in single-trials with simultaneous intracerebral EEG-fMRI recordings. Int. J. Neural Syst. https://doi.org/10.1142/S0129065717500010

Sarraf, S., & Tofighi, G., 2016. DeepAD: Alzheimer′ s disease classification via deep convolutional neural networks using MRI and fMRI. bioRxiv, 070441

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. https://doi.org/10.1214/12-AOS1000

Talo, M., Baloglu, U.B., Yildirim, O., Acharya, U.R., 2019. Application of deep transfer learning for automated brain abnormality classification using MR images. Cogn. Syst. Res., https://doi.org/10.1016/j.cogsys.2018.12.007

Wang, S., Lu, S., Dong, Z., Yang, J., Yang, M., Zhang, Y., 2016. Dual-Tree Complex Wavelet Transform and Twin Support Vector Machine for Pathological Brain Detection. Appl. Sci. https://doi.org/10.3390/app6060169

Wang, S., Zhang, Y., Dong, Z., Du, S., Ji, G., Yan, J., Yang, J., Wang, Q., Feng, C., Phillips, P., 2015. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. Int. J. Imaging Syst. Technol. https://doi.org/10.1002/ima.22132

Wegmayr, V., Aitharaju, S., Buhmann, J., 2018. Classification of brain MRI with big data and deep 3D convolutional neural networks, in: Medical Imaging 2018: Computer-Aided Diagnosis. https://doi.org/10.1117/12.2293719

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H., 2014. How transferable are features in deep neural networks?, In Advances in neural information processing systems (pp. 3320-3328)

Zhang, Y., Dong, Z., Liu, A., Wang, S., Ji, G., Zhang, Z., Yang, J., 2015. Magnetic Resonance Brain Image Classification via Stationary Wavelet Transform and Generalized Eigenvalue Proximal Support Vector Machine. J. Med. Imaging Heal. Informatics. https://doi.org/10.1166/jmihi.2015.1542

Zhang, Y., Dong, Z., Wu, L., Wang, S., 2011. A hybrid method for MRI brain image classification. Expert Syst. Appl. https://doi.org/10.1007/s40592-016-0061-3

Zhang, Y., Zhang, H., Chen, X., Lee, S. W., & Shen, D. (2017). Hybrid high-order functional connectivity networks using resting-state functional MRI for mild cognitive impairment diagnosis. Scientific reports, 7(1), 6530.

Zhang, Y., Zhang, H., Chen, X., Liu, M., Zhu, X., Lee, S. W., & Shen, D., 2019. Strength and similarity guided group-level brain functional network construction for MCI diagnosis. Pattern Recognition, 88, 421-430.

Zhou, X.X., Zhang, Y., Ji, G., Yang, J., Dong, Z., Wang, S., Zhang, G., Phillips, P., 2016. Detection of abnormal MR brains based on wavelet entropy and feature selection. IEEJ Trans. Electr. Electron. Eng. https://doi.org/10.1002/tee.22226

Zhou, G., Zhao, Q., Zhang, Y., Adalı, T., Xie, S., & Cichocki, A. (2016). Linked component analysis from matrices to high-order tensors: Applications to biomedical data. Proceedings of the IEEE, 104(2), 310-331.

Zikic, D., Ioannou, Y., Brown, M., Criminisi, A., 2014. Segmentation of Brain Tumor Tissues with Convolutional Neural Networks, in: Proceedings MICCAI-BRATS.