Machine Vision and Applications (2020) 31:20
https://doi.org/10.1007/s00138-020-01069-2

SPECIAL ISSUE PAPER

Deep convolutional neural networks with transfer learning for automated brain image classification

Taranjit Kaur[1] · **Tapan Kumar Gandhi**[1]

Received: 13 August 2019 / Revised: 26 November 2019 / Accepted: 18 February 2020 / Published online: 27 March 2020

Abstract

MR brain image categorization has been an active research domain from the last decade. Several techniques have been devised in the past for MR image categorization, starting from classical to the deep learning methods like convolutional neural networks (CNNs). Classical machine learning methods need handcrafted features to perform classification. The CNNs, on the other hand, perform classification by extracting image features directly from raw images via tuning the parameters of the convolutional and pooling layer. The features extracted by CNN strongly depend on the size of the training dataset. If the training dataset is small, CNN tends to overfit after several epochs. So, deep CNNs (DCNNs) with transfer learning have evolved. The prime objective of the present work is to explore the capability of different pre-trained DCNN models with transfer learning for pathological brain image classification. Various pre-trained DCNNs, namely Alexnet, Resnet50, GoogLeNet, VGG-16, Resnet101, VGG-19, Inceptionv3, and InceptionResNetV2, were used in the present study. The last few layers of these models were replaced to accommodate new image categories for our application. These models were extensively evaluated on data from Harvard, clinical, and benchmark Figshare repository. The dataset was then partitioned in the ratio 60:40 for training and testing. The validation on the test set reveals that the pre-trained Alexnet with transfer learning exhibited the best performance in less time compared to other proposed models. The proposed method is more generic as it does not need any handcrafted features and can achieve an accuracy value of 100%, 94%, and 95.92% for three datasets. Other performance measures used in the study include sensitivity, specificity, precision, false positive rate, error, F-score, Mathew correlation coefficient, and area under the curve. The results are compared with both the traditional machine learning methods and those using CNN.

Keywords Transfer learning · Overfit · Classification · MR images

1 Introduction

Magnetic resonance (MR) imaging is a popular noninvasive modality for the visualization of different abnormalities in the brain due to its good soft tissue contrast and accessibility of multispectral images. Using information from MR images, computer-aided diagnosis (CAD) systems have been developed to benefit doctors in rapid diagnosis. CAD systems can provide the diagnosis depending upon the specific attributes present in the medical images. Typically, these systems usually employ the steps of preprocessing, attribute extraction, selection, and classification for categorizing normal/abnormal brain MR images [1–7]. Numerous methods have been proposed in the literature that employs classical machine learning algorithms for the detection of abnormal brain images. These studies have proposed solutions based on K-nearest neighbor (KNN) [1, 8], support vector machine (SVM) [2, 5, 9, 10], Kohonen–Hopfield neural network (KHNN) [4], and artificial neural networks (ANNs) [6].

Sachdeva et al. [11] developed a CAD system for the automatic classification of the brain tumors using a content-based active contour model, intensity and texture features, genetic algorithm-based feature selection, and ANN-based classification. Lahmiri [9] compared three automatic diagnosis schemes for the categorization of the normal subjects from glioma patients using fractional-order Darwinian particle swarm optimization (DPSO), directional spectral distribution, multiscale analysis, and SVM. Bahadure et al. [8] devised CAD technology for the automatic classification of

✉ Tapan Kumar Gandhi
tgandhi@ee.iitd.ac.in

Taranjit Kaur
taran.rehal@yahoo.com; Taranjit.Kaur@ee.iitd.ac.in

[1] Department of Electrical Engineering, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110016, India

tumor MR images using an adaptive neuro-fuzzy inference system, KNN, and GA-based approach. The CAD consists of stages of preprocessing that involved contrast enhancement and skull stripping, segmentation, feature extraction, feature selection, and classification.

More recent techniques have been designed in the literature benefitting from advanced machine learning approaches. These approaches usually combine the machine learning algorithms with the optimization techniques like PSO [12], biogeography-based optimization (BBO) [13], artificial bee colony (ABC), BAT [14], and their hybridizations [15, 16]. Zhang et al. [12] developed a classification mechanism for MR brain image categorization into normal and abnormal classes using kernel SVM and PSO. Wang et al. [15] employed the PSO and ABC hybridized feed-forward neural network for the classification of normal and abnormal brains. Yang et al. [13] proposed a BBO optimized SVM for the categorization of MR images. Lu et al. [14] developed a mechanism for MR image categorization using the BAT optimized extreme machine learning algorithm.

Recently, deep learning algorithms like convolutional neural networks (CNNs) have become prevalent in the domain of machine learning. CNN can automatically extract the image features by parameter tuning of the convolutional ("conv") and the pooling ("pool") layers. The CNN models have been successfully applied to several biomedical applications [17–23]. Nevertheless, few studies have addressed pathological brain categorization using CNNs [24, 25]. Zuo et al. [22] proposed a deep learning mechanism for dermatology diagnostics. Abiwinanda et al. [24] proposed a CNN-based deep learning model for brain tumor classification. The model involved an architecture comprising of convolution, max pooling, flattening layers followed by a full connection from a single hidden layer. The authors obtained a training accuracy of 98.51% and validation accuracy of 84.19%. Afshar et al. [25] proposed brain tumor classification using an improved CNN architecture known as capsule network (CapsNet) [25]. CapsNet utilized the spatial relationship between the tumor and its surrounding tissues for the tumor classification. The proposed CapsNet architecture rendered an overall accuracy of 90.89%.

The classification capability of CNN is highly reliant on the size of the data used for the training. If the size of the dataset is small, the CNN model starts overfitting after several epochs. In this scenario, the special domain of deep learning, identified as transfer learning has been evolved and has proved to be effective in the field of medical diagnosis [18, 20, 26–30]. In transfer learning, weights of the "conv" layer from pre-trained models are used and only the last dense layers are trained with data from the newer classes.

Deniz et al. [26] employed a pre-trained VGG-16 and fine-tuned Alexnet model for breast cancer diagnosis. The features were extracted using the pre-trained models which were then classified using an SVM. Talo et al. [31] used the pre-trained ResNet-34 model for brain abnormality classification. The authors proposed modifications in the dense layers of the model, fine-tuning, and training the model with data augmentation. Deepak and Ameer [29] employed pre-trained GoogLeNet to extract features from MR images. The authors classified the extracted features using the softmax, SVM, and the KNN classifier. The authors attained a mean accuracy of 92.3%, 97.8%, and 98% using the standalone CNN, SVM, and KNN classifier. Swati et al. [30] proposed a block-wise fine-tuning mechanism using a pre-trained VGG19 for brain tumor classification. The authors achieved an overall accuracy of 94.82%.

The conventional and advance machine learning techniques discussed in the past suffer from the limitations that they need: (1) the prerequisite for the segregation of the region of interest (ROI) manually [1, 2, 7], (2) the requisite of large number of attribute extraction [1–6], and (3) these techniques are often abetted with the dimensionality reduction/feature selection techniques to decrease the extracted feature space dimensionality [1, 2, 5, 6], before providing a final prediction on the MR brain image label.

To overcome the above-mentioned shortcomings, DCNNs have evolved, but they start overfitting when the data size is small. Pre-trained DCNN models with transfer learning were developed to solve this issue of overfitting, and so far, only a few works have experimented them for pathological brain image classification task [29–31]. Also, the methodological review reveals that a comparative study investigating the capability of the pre-trained DCNN models with transfer learning on different versions of abnormal MR image dataset is missing.

In the present paper, we have compared the capability of different pre-trained DCNN models with transfer learning for automated MR brain image classification (normal vs abnormal, glioma vs meningioma vs pituitary). The prominent pre-trained DCNN models like Alexnet [32, 33], GoogLeNet [34], ResNet50 [35], ResNet101 [35], Vgg-16 [36], Vgg-19 [36], Inceptionv3 [37], and InceptionResNetV2 [38] have been used for efficient classification. The key contributions of the present study are as follows:

1. To explore the different architectures of deep convolutional networks like Alexnet, GoogLeNet, ResNet50, ResNet101, Vgg-16, Vgg-19, Inceptionv3, and InceptionResNetV2 and to deploy them for pathological MR brain classification using the concept of transfer learning.
2. A substantial improvement in performance is observed when transfer learned deep Alexnet model is applied to different datasets from Harvard repository, Clinical dataset, and multiclass tumor dataset from the Figshare in contrast to the existing state-of-the-art methods.

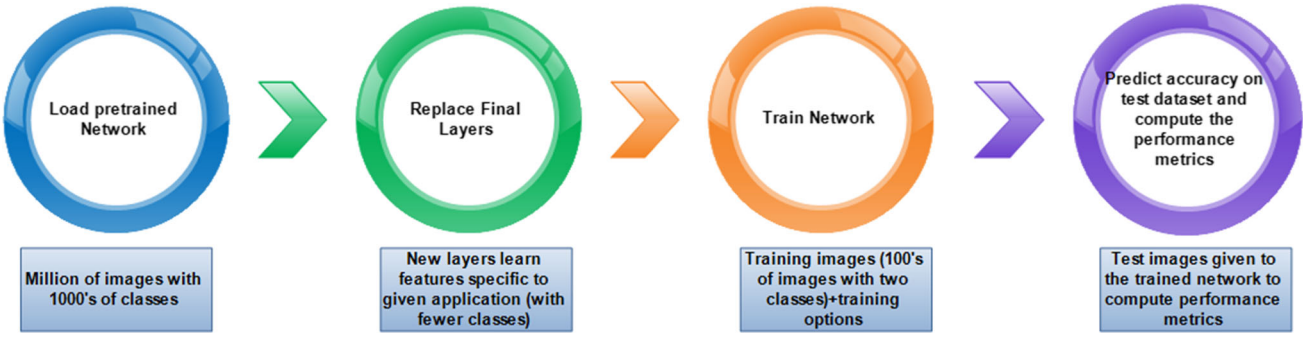

Fig. 1 A brief illustration of the mechanism for transfer learning

3. The proposed strategy is different from works by [26, 29], in which the pre-trained CNNs are used as off-the-shelf feature extractors and trained a separate classifier for classification.
4. The proposed strategy is different from the works by [30, 31], in which ResNet 34 and VGG-19 are used for classification using Harvard and Figshare data.

The paper is organized as follows: the concept of transfer learning in deep CNNs is described in Sect. 2; the methodology is given in Sect. 3; the database and performance measures are presented in Sect. 4; the results and discussion are given in Sect. 5; and lastly, the conclusion is given in the end.

2 Transfer learning in DCNN

Deep CNN, i.e., DCNNs usually require huge labeled image datasets to attain ceiling level of classification accuracy. Nevertheless, in several fields, the acquirement of such an image dataset is challenging and annotating it is expensive. In the event of such difficulties, the usage of "off-the-shelf" attributes of established DCNNs like Alexnet [32, 33], GoogLeNet [34], ResNet50 [35], ResNet101 [35], Vgg-16 [36], Vgg-19 [36], Inceptionv3 [37], and InceptionResNetV2 [38] pre-trained on huge labeled natural image database (like ImageNet) have proved to be beneficial for resolving cross-domain image categorization problems via transfer learning [20]. In DCNN, depictions imbibed at various network layers relate to several levels of abstractions existing in the input dataset. The filters at the initial layers excerpt color and edge info, while the filters at the latter layers encode texture and shape. The notion behind transfer learning is twofold: Firstly, it is cheaper, and secondly, it is effectual to deploy DCNN trained on "big data" image datasets and "transfer" their training ability to the newer image categorization situation than training a DCNN from scratch [18]. With suitable fine-tuning, pre-trained DCNNs has outperformed the DCNN trained from scratch for biomedical applications [20, 39]. Just for the illustration, the mechanism of transfer learning using any of the above-discussed models is briefly described in Fig. 1.

3 Methodology

The methodology for automated brain image classification using transfer learning is shown in Fig. 2. The structure shown in Fig. 2 consists of layers from the pre-trained model and few new layers. For the present work, for all the models only the last three layers were replaced to accommodate the new image categories. The step-by-step procedure is outlined below:

Step 1 Download the dataset of MR images from the repository. It comprises of the normal and abnormal MR images. Place them in the imgData/train directory.
Step 2 Formulate an image DataStore to read images. Labels all the images according to their folder name and include all the subfolders in the directory
Step 3 Resize the input images so that they are consistent with the size of the input layer of the pre-trained network model
Step 4 Partition the data into training and test sets; 60% of images per category are taken for training and 40% as a test dataset to test the network
Step 5 Review network architecture (replace the final layers/modify the layers as follows)

- For "alexnet," "Vgg-16," and "Vgg-19" alter just the last three layers of the pre-trained network with a set of layers ("fully connected layer," a "softmax layer," and a "classification output layer") to categorize the images into the respective classes
- For "GoogLeNet" change the last three layers of the network. The last three layers of the network, i.e., "loss3-classifier," "prob," and "output" are changed by a "fully connected layer," a "softmax layer," and a "classification output" layer. Link the last transferred layer

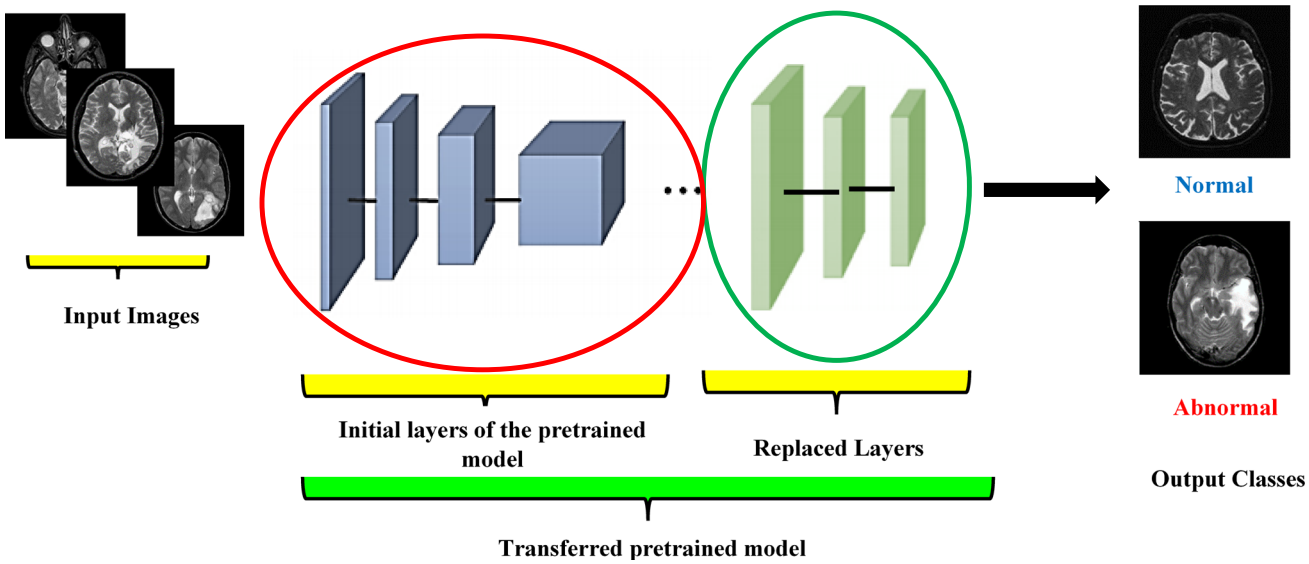

Fig. 2 Mechanism of transfer learning using pre-trained models

remaining in the network ("pool5-drop_7 × 7_s1") to the new layers

- For "ResNet50," replace last three layers, i.e., "fc1000," "fc1000_softmax," and "ClassificationLayer_fc1000" of the network with "fully connected layer," a "softmax layer," and a "classification output" layer. Connect the last transferred layer remaining in the network ("avg_pool") to the new layers
- For "ResNet101," replace the last three layers, i.e., "fc1000," "prob," and "ClassificationLayer_predictions" of the network with a "fully connected layer," a "softmax layer," and a "classification output" layer. Connect the last transferred layer remaining in the network ("pool5") to the new layers
- For "Inceptionv3" and "InceptionResNetV2" substitute the last three layers of the network, i.e., "predictions," "predictions_softmax," and "ClassificationLayer_predictions" with "fully connected layer," a "softmax layer," and a "classification output" layer. Connect the last transferred layer remaining in the network ("avg_pool") to the new layers

Step 6 Train the Network
Step 7 Test the new model on the testing dataset
Step 8 Report the performance metrics

4 Database and performance measures

4.1 Database

The experiments were done on the images taken from Harvard repository [40], the clinical dataset from Fortis Memorial Research Institute, Gurgaon, India [41, 42], and images from Figshare [43] repository. Images from the Harvard repository were divided into three versions as exiting in the literature [9, 44]. Table 1 details the three versions of the Harvard dataset that has been employed in the present paper. The first encompasses 50 images. The second contains 74 images. The third variant consists of 160 images.

Table 1 Particulars of the Harvard database

Dataset	Details
Version 1 [9]	50 T2-w: 20 abnormal affected with glioma tumor and 30 normal
Version 2 [44]	74 T2-w: 52 abnormal comprising of Glioma, herpes encephalitis, metastatic bronchogenic carcinoma, multiple sclerosis, and Alzheimer's, and 22 normal
Version 3 [44]	160 T2-w: 140 abnormal and 20 normal. The MR images in the abnormal category have glioma, meningioma, Alzheimer's disease, Pick's disease, sarcoma, Alzheimer's disease plus visual agnosia, and Huntington's disease

For the clinical dataset, images from 200 patients were used (age range 18–65 years) from the Fortis Memorial Research Institute, Gurgaon, India [41, 42]. The MRI scan was performed on a 3T MRI (Philips Ingenia, Best Netherland) using a 15-channel head coil. 3D fluid attenuating inversion recovery (FLAIR) was done using the scan parameters as follows: TR (ms) = 4500, TE (ms) = 279, NEX 1, and slice thickness 10 mm. From the data of 200 patients, 500 images were used out of which 250 were normal and 250 were abnormal, i.e., those affected by glioma tumors.

The opinion of the radiologist was considered as the gold standard. Moreover, the distinction between the normal and

abnormal categories was made by the radiologist having more than 25 years of experience in radiology. All the decisions were made by the consensus. The patient data were normalized using the T2.nii template on the SPM toolbox in MATLAB. A gap of about 10 mm between the slices was selected, thereby rendering 16 slices per patient.

Additionally, solely to contrast the performance of the deep transfer learned model on another challenging medical data, experimentations were also done on the dataset from Figshare [43]. It consists of 3064 brain MR images of 233 patients having glioma (1426), meningioma (708), and pituitary tumors (930). T1-CE images were available in axial, coronal, and sagittal views. The images were available in. mat format with size as 512×512. Images were resized to match the dimensions of the input layer of the pre-trained model. Since the images were grayscale, additional channels were created by replicating the pixel values three times. The sample MR images are shown in Fig. 3.

4.2 Performance measures

To validate the effectiveness of the pre-trained DCNN models with transfer learning, several performance measures are chosen. Sensitivity, specificity, and accuracy are the most commonly chosen measures. The classification accuracy is interrelated with sensitivity and specificity which utilize the terms: true positive (TP), true negative (TN), false negative (FN), and false positive (FP). TP is the number of abnormal images labeled as abnormal, TN is the number of normal images labeled as normal, FP is the number of the normal images labeled as abnormal, and FN is the number of abnormal images labeled as normal. Mathematically sensitivity and specificity are given as

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{1}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{2}$$

$$\text{Accuracy} = \frac{TN + TP}{TP + FP + TN + FN} \tag{3}$$

The above-stated measures are effective measures only in the case of the balanced dataset. Unbalanced datasets necessitate further validation using additional performance metrics. Additional indices were computed from the generated confusion matrix that includes the precision, false positive rate, error, F-score, Mathew correlation coefficient (MCC), and kappa index. Additionally, the area under the curve (AUC) was computed using the method proposed in [45].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{False Positive Rate} = \frac{FP}{TN + FP} \tag{5}$$

$$\text{Error} = \frac{FP + FN}{TP + TN + FN + FP} \tag{6}$$

F-score is the harmonic mean between precision and recall. Mathematically, it is given as

$$\text{F - score} = \frac{(1 + \beta^2(\text{Precision} \cdot \text{Recall})}{(\beta^2 \times \text{Precision} + \text{Recall})} \tag{7}$$

where β is usually 1.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{8}$$

$$\text{Kappa} = \frac{\text{Accuracy} - \text{randomAccuracy}}{1 - \text{randomAccuracy}} \tag{9}$$

$$\text{randomAccuracy} = \frac{(TN + FP) \times (TN + FN) + (FN + TP) \times (FP + TP)}{\text{Total} \times \text{Total}} \tag{10}$$

5 Experimental results

In this section, detailed information about the experimental setup and the obtained results is given. The experimental setup includes the info about training the model and the software platform used in the present work.

5.1 Experimental setup

To evaluate the performance of the proposed pre-trained models with transfer learning, a fixed partitioning scheme has been chosen. The datasets are partitioned into training and test sets using a ratio of 60:40. It means formerly, 60% of the dataset is employed for training and the rest 40% is employed for testing the transferred models. Additionally, in "Discussion" section when the comparison is made with the current state-of-the-art works we have followed the same settings existing in the literature employing both the fixed partition and the cross-validation schemes.

The pre-trained models with transfer learning were designed in MATLAB 2017b platform and executed using a system configuration of Intel Xeon CPU E3-1225, 3.3 GHz, and 8 GB RAM. The transferred models were trained using Stochastic gradient descent with momentum (SGDM). The mini-batch size was taken as 10, the maximum number of epochs was 20 and the learning rate was 0.0001. Lowering the learning rate surges the network training time and increasing the rate leads training to get stuck at a suboptimal result.

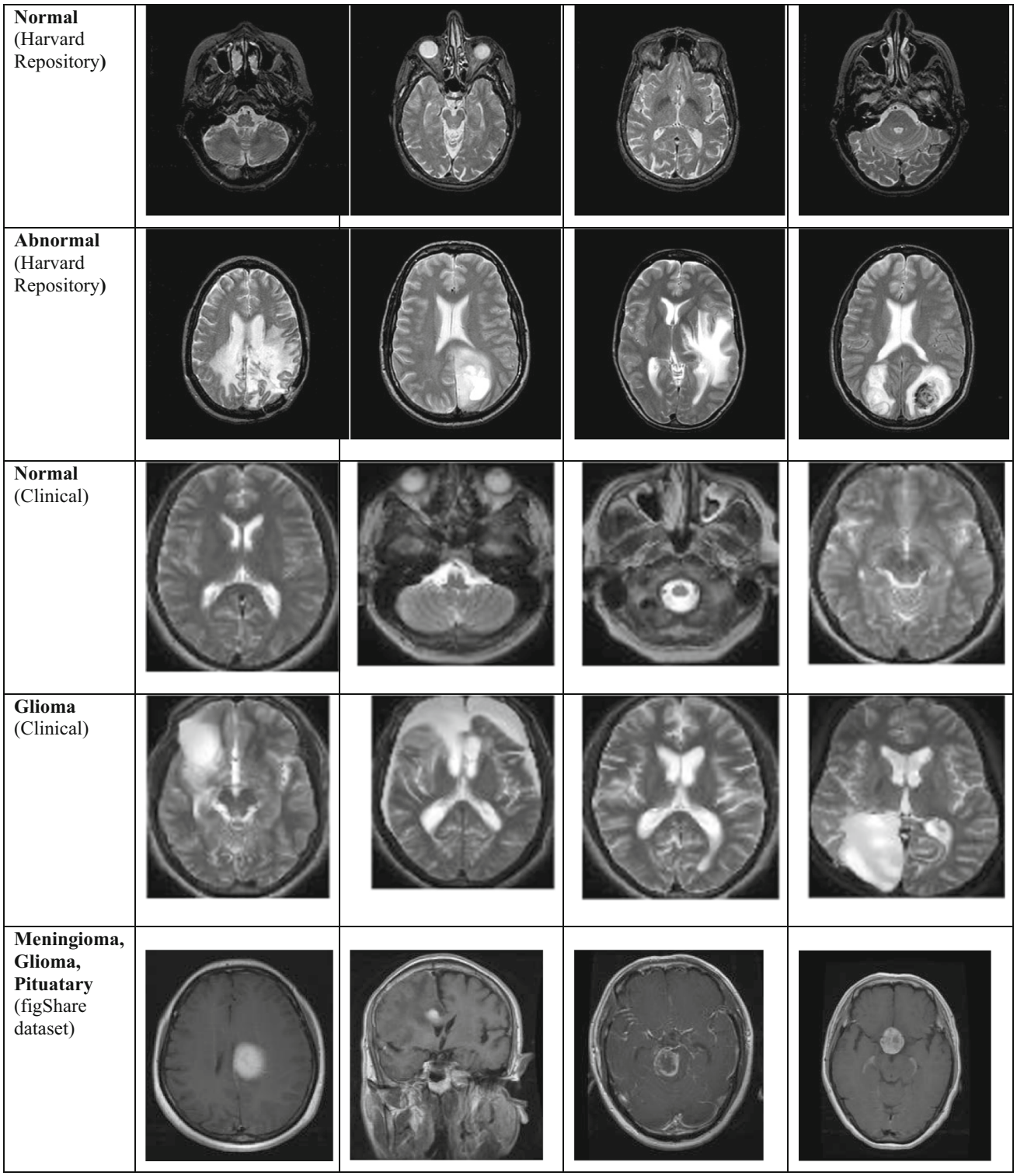

Fig. 3 Sample normal and abnormal brains from the Harvard repository, clinical dataset and Figshare dataset

5.2 Results

The classification performance of the proposed transferred models is evaluated on different datasets, and the results are summarized in the form of tables and graphical figures. Table 2 illustrates the performance comparison for all the pre-trained models using all the chosen performance metrics.

Table 2 Performance metrics over all the datasets

Pre-trained model	Dataset	Accuracy	Error	Sensitivity	Specificity	Precision	False positive rate	F-score	MCC	Cohen's kappa	AUC
Alexnet	Version 1	100	0	100	100	100	0	100	100	100	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	100	0	100	100	100	0	100	100	100	100
	Clinical	94	0.06	95	93	93.13	0.07	94.05	88.02	88	98.63
	Figshare	95.92	0.041	94.77	97.71	96.21	0.03	95.40	93.37	90.84	99.66
GoogLeNet	Version 1	85	0.15	87.5	83.33	77.78	0.167	82.35	69.75	69.39	97.92
	Version 2	96.67	0.033	100	88.89	95.45	0.111	97.67	92.11	91.8	99.47
	Version 3	98.44	0.016	100	87.5	98.25	0.125	99.12	92.72	92.45	100
	Clinical	88	0.12	88	88	88	0.12	88	76	76	94.76
	Figshare	88.73	0.112	85.39	93.78	89.09	0.06	86.31	81.36	74.65	98.47
ResNet50	Version 1	100	0	100	100	100	0	100	100	100	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	100	0	100	100	100	0	100	100	100	100
	Clinical	86	0.14	90	82	83.33	0.18	86.54	72.23	72	94.01
	Figshare	91.51	0.08	90.65	95.79	90.22	0.04	90.38	86.15	80.89	98.52
ResNet101	Version 1	95	0.05	87.5	100	100	0	93.33	89.87	89.36	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	100	0	100	100	100	0	100	100	100	100
	Clinical	87	0.13	97	77	80.83	0.23	88.18	75.53	74	95.86
	Figshare	92.16	0.08	90.69	95.86	91.58	0.04	91.07	87.12	82.37	98.78
VGG-16	Version 1	100	0	100	100	100	0	100	100	100	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	100	0	100	100	100	0	100	100	100	100
	Clinical	91.5	0.085	94	89	89.52	0.11	91.71	83.1	83	97.65
	Figshare	92.37	0.07	90.96	96.05	92.62	0.039	91.59	88.03	83.65	99.12
VGG-19	Version 1	100	0	100	100	100	0	100	100	100	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	100	0	100	100	100	0	100	100	100	100
	Clinical	87	0.13	77	97	96.25	0.03	85.56	75.53	74	96.71
	Figshare	91.27	0.087	91.92	95.98	89.91	0.040	90.622	86.52	80.35	98.71
Inceptionv3	Version 1	95	0.05	87.5	100	100	0	93.33	89.87	89.36	100
	Version 2	93.33	0.067	100	77.78	91.3	0.222	95.45	84.27	83.05	100
	Version 3	93.75	0.063	100	50	93.33	0.5	96.55	68.31	63.64	99.54
	Clinical	78	0.22	90	66	72.58	0.34	80.36	57.69	56	99.55
	Figshare	87.92	0.12	87.21	94.14	86.04	0.058	86.50	80.58	72.81	97.65
InceptionResNetV2	Version 1	100	0	100	100	100	0	100	100	100	100
	Version 2	100	0	100	100	100	0	100	100	100	100
	Version 3	93.75	0.063	100	50	93.33	0.5	96.55	68.31	63.64	99.55
	Clinical	84	0.16	83	85	84.69	0.15	83.84	68.01	68	93.14
	Figshare	89.95	0.100	87.59	94.64	89.40	0.053	88.21	83.36	77.40	98.40

The tabular findings indicate that the pre-trained Alexnet model with transfer learning achieves the finest results over all the three datasets followed by VGG-16. The attained values of accuracy, error, sensitivity, specificity, precision, false positive rate, F-score, MCC, Cohen's kappa, and AUC were 100, 0, 100, 100, 100, 0, 100, 100, 100, and 100 for three versions of the dataset from the Harvard repository. The performance evaluation on challenging clinical and Figshare datasets yielded the values equal to 94, 0.06, 95, 93, 93.13, 0.07, 94.05, 88.02, 88, and 98.63 and 95.92, 0.041, 94.77, 97.71, 96.20, 0.03, 95.40, 93.37, 90.83, and 99.66 which are better than the other pre-trained models. The value of performance metrics indicates the benefit of transfer learning in reducing overfitting and increasing the speed of the convergence.

The training progress curve and the ROC plot for the best performing model are shown in Figs. 4 and 5 for all datasets. From the training progress curve, it is apparent that the Alexnet model reaches to ceiling level of accuracy for all the datasets in merely 20 epochs. Moreover, the ROC

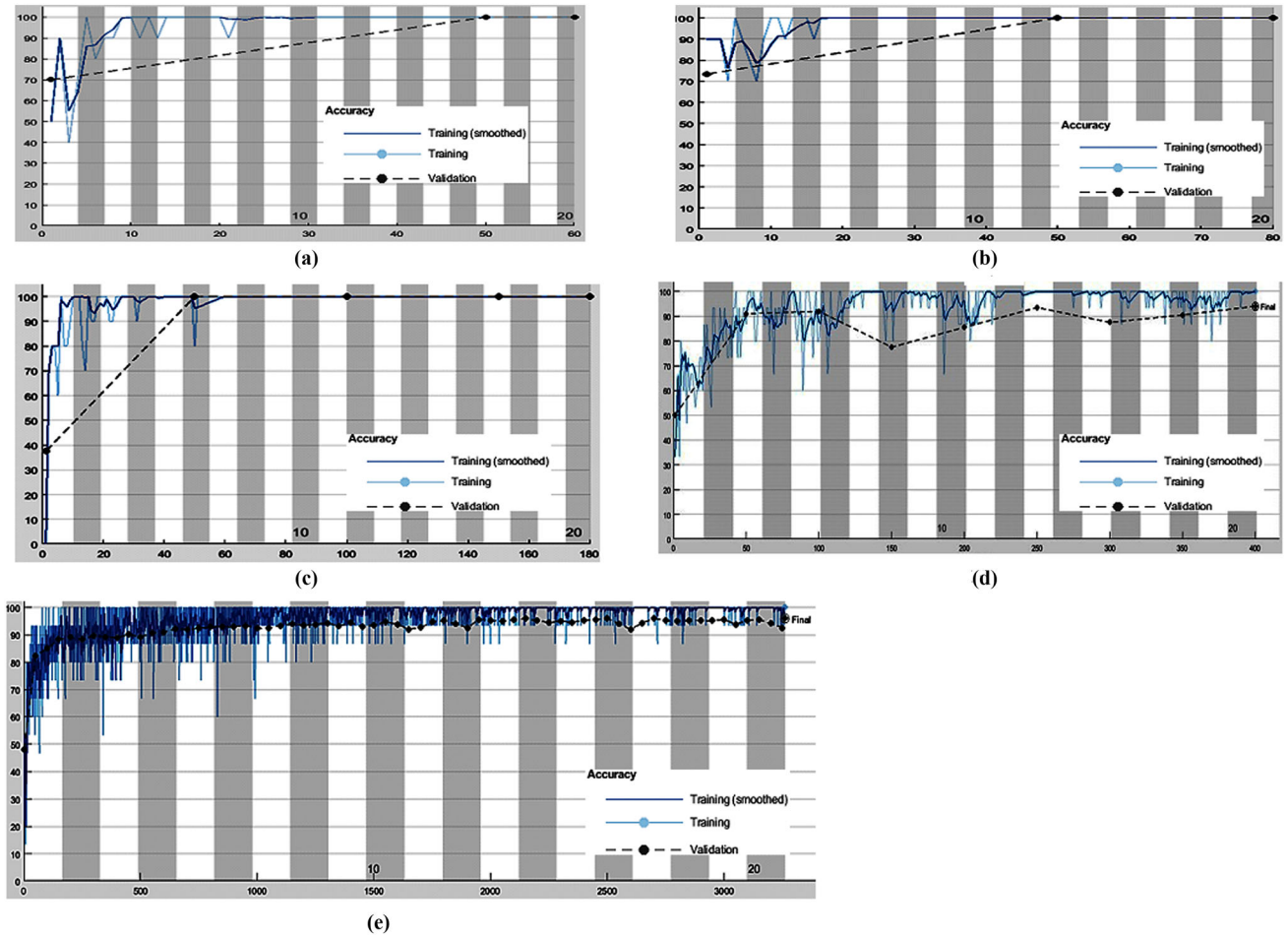

Fig. 4 Training progress curves (Accuracy (%) vs Iterations) for transferred Alexnet model for **a** version 1, **b** version 2, **c** version 3, **d** clinical dataset, **e** Figshare dataset

plots with a value equal to 100, 100, 100, 98.63, and 99.66 supplement the findings that Alexnet is capable of correctly classifying the samples from the datasets with high true fraction value and low false positive rate.

5.3 Discussion

The first subsection of discussion illustrates the comparison of the best performing transferred CNN models with other existing state-of-the-art works on brain tumor categorization using version 1 of the dataset from the Harvard repository and the challenging clinical dataset.

The second subsection compares the performance of the best performing transferred CNN models with other recent existing schemes that have reported classification results on versions 2 and 3 [40]. For versions 2 and 3, the comparison has been made in terms of the performance metrics reported in the literature, i.e., the no. of features, classification accuracy, computation time, and data partitioning scheme.

The third subsection contrasts the performance of the best performing transferred CNN model with other recent existing schemes that have reported classification results on challenging multiclass tumor dataset from Figshare.

5.4 Comparison of the transfer learning-based pre-trained DCNN models with recent works on normal vs tumor classification

The performance measures reported in Table 2 of "Results" section indicate that transferred CNN models, i.e., Alexnet [32, 33], ResNet50 [35], Vgg-16 [36], Vgg-19 [36], and InceptionResNetV2 [38], have accomplished 100% Se, Sp, and Acc on version 1 of the dataset dealing with glioma characterization. For the challenging clinical dataset, transferred Alexnet [32, 33] proved to be finest by attaining Se, Sp, and Acc values of 95%, 93%, and 94%.

The comparison of the best performing transferred model with the existing works dealing with glioma image categorization [1–4, 6, 7, 9, 46] is given in Table 3. From the tabular

Table 3 Comparison of the transfer learning-based pre-trained DCNN models with recent works on tumor classification

References			Number of images/segmentation approach	Feature extraction approach/selection method/classification technique with partition scheme
Performance				
Se	*Sp*	*Acc*		
[1]			26 Glioblastoma multiform/manual contouring of region of contrast enhancement.	Diffusion, perfusion, and spectroscopic features/genetic algorithm/KNN/leave one out method (LOOM)
78	79	–		
[2]			102 brain tumors/manually traced 4 regions	Intensity, shape, and Gabor texture features/SVM + recursive feature elimination [47, 48]/SVM/LOOM
Metastases versus gliomas				
87	79	85		
HG versus LG				
85	96	88		
[3]			25 solitary metastases and 38 glioblastomas/semiautomatic	Diffusion tensor image features/logistic regression
92	100	–		
[4]			65 metastases, 71 meningioma, 63 gliomas, and 51 astrocytoma images/no ROI delineation	DWT coefficients/NA/KHNN/Fixed Partitioning
–	–	93.60		
[5]			101 glioma tumors with different grades/automatic fuzzy clustering technique (FCM) [49]	First-order statistics/PCA/SVM/tenfold cross-validation
89	84	85		
[6]			428 images from 55 patients/semiautomatic content-based active contour model (CBAC) [50]	First- and second-order statistics, Laplacian of Gaussian, directional gabor features, rotation invariant circular gabor features, and rotation invariant local binary patterns/PCA/artificial neural network/tenfold cross-validation (CV)
–	–	91		
[7]			27 LG and 68 HG gliomas/manual	Texture attributes at various anatomical scales followed by statistical analysis
93	81	–		
[9]			30 healthy, 20 gliomas/Classical PSO	DSD + MSA/NA/SVM/LOOM
100	94.78	98.01		
[9]			30 healthy, 20 gliomas/darwinian PSO (DPSO)	DSD + MSA/NA/SVM/LOOM
99.50	96.70	98.38		
[9]			30 healthy, 20 gliomas/fractional-order DPSO	DSD + MSA/NA/SVM/LOOM
100	97.95	99.18		
[8]			67 H, 134 Ab/Watershed, FCM, DCT, and BWT	First- and second-order statistics and tumor area/GA 11 features/ANFIS, KNN, and GA-based classification/NA
92.36	91.42	92.03		
[51]			450 T2-w tumorous images/no ROI delineation	Gray-level difference method/modified GA (9 features)/BPNN//fixed partition
96	98	98.1		
[10]			120 HG, 80 LG (BRATS 2012)/adaptive thresholding	Geometric features/NA/ensemble of SVM, KNN, and NB/tenfold CV
LG/HG 93/98.20	LG/HG 98.9/93.67	LG/HG 94.40/94		
Alexnet [32, 33]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
100	100	100		
ResNet50 [35]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
100	100	100		

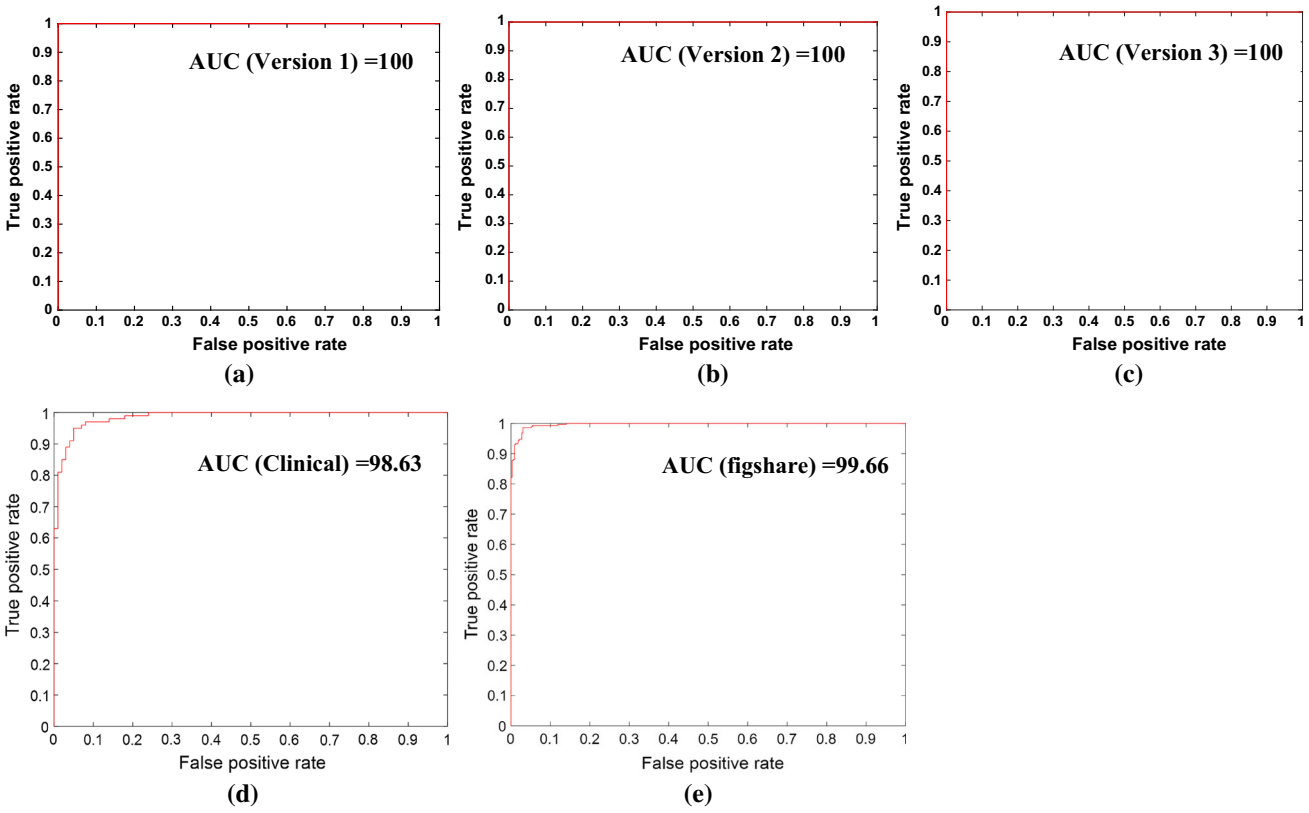

Fig. 5 ROC plots for transferred Alexnet model for **a** version 1, **b** version 2, **c** version 3, **d** clinical dataset, and **e** Figshare dataset

findings, it is seen that the transferred DCNN models are capable of distinguishing glioma images from the normal brains in contrast to the other presented approaches.

As the works reported in [1–4, 6, 7, 46], also dealt with the categorization of glioma, the objective assessment is hard to make as the results have been attained on their private self-collected data. Still, for the approach in [9] as the dataset was accessible from [40], the best performing pre-trained DCNN models with transfer learning have also been validated on this version of the dataset. For this version of the dataset, i.e., version 1, the models with transfer learning yielded specificity and accuracy equal to 100% that is higher than the mean value equal to 99.18% and 97.95% as indicated in the works by [9].

For the clinical dataset, transferred Alexnet yielded better value of performance measures in comparison to the works by [41, 42] that reported an accuracy value of 85.63%.

Additionally, on applying a tenfold cross-validation partitioning scheme to version 1 and the clinical dataset, the transferred Alexnet reached an accuracy of 100% and 90% better than the existing state-of-the-art works [5, 6, 41].

Furthermore, the transferred DCNN models used for the classification offer several benefits in contrast to existing tumor classification works; firstly, it offers ceiling level of accuracy for the database reported in [9] and the challenging clinical dataset. Second, no prerequisite for the manual delineation of the tumorous region is there, as existing in the works reported by [1–3, 7]. Third, it eliminates the need for the extraction of any features [5, 6, 41, 42]. It also evades the necessity of using any dimensionality reduction or feature selection technique as existing in the works by [1, 2, 6, 41, 42, 46] .

5.5 Performance comparison of transfer learning-based pre-trained DCNN models with the recent literature for versions 2 and 3 of the dataset

Table 2 shows that the transferred DCNN models Alexnet [32, 33], ResNet50 [35], ResNet101 [35], Vgg-16 [36], Vgg-19 [36], and InceptionResNetV2 [38] have achieved 100% accuracy for versions 2 and 3 of the dataset from the Harvard repository. The comparative analysis with existing works on this dataset is given in Table 4. The comparative analysis has been restricted to the measures reported in the literature, i.e., classification accuracy, number of features, and computation time. The measures reported in Table 4 show that transferred DCNN models reach to a higher level of accuracy than the methods devised in works by [52–55] for which the value was 97%; 98.00%, 97.33%; 97.00, 96.98 and 99.85, 99.62.

Table 3 continued

References			Number of images/segmentation approach	Feature extraction approach/selection method/classification technique with partition scheme
Performance				
Se	*Sp*	*Acc*		
ResNet101 [35]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
87.50	100	95		
Vgg_16 [36]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
100	100	100		
Vgg_19 [36]			30 healthy, 20 gliomas/No Segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
100	100	100		
Inceptionv3 [37]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
87.50	100	95		
InceptionResNetV2 [38]			30 healthy, 20 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
100	100	100		
Alexnet [32, 33]			250 normal, 250 gliomas/no segmentation	No feature extraction/NA/cross-entropy-based/fixed partition
94	92	93		

–Findings not stated; Acc, accuracy; Sp, specificity; Se, sensitivity; LG, low-grade; HG, high-grade; DSD, directional spectral distribution; MSA, multiscale analysis; DCT, discrete cosine transform; BWT, Berkeley wavelet transform; BPNN, back-propagation neural network; NA, not applicable

Moreover, the transferred Alexnet model reaches this ceiling level of accuracy in the minimum amount of time in contrast to other transferred DCNN models. Table 4 shows that some of the approaches [15, 56–59] have achieved 100% accuracy on version 2, but their accuracy drops when switching to larger datasets. Though few of the methods in Table 4 achieved 100% accuracy for both versions 2 and 3 like those designed in the works by [16, 44, 55, 60], they use 13, 16, 9, and 6 features. The prime advantage of the pre-trained DCNN models with transfer learning in comparison with the reported works in Table 4 is that it needs no feature extraction mechanism and no intermediate feature selection phase. Additionally, in contrast to works in [31], the Alexnet model with transfer learning reaches to 100% recognition rate by merely using 25 layers.

Additionally, the performance of the transferred Alexnet model is also investigated under the tenfold cross-validation partitioning scheme. The transferred Alexnet model rendered 100% recognition, but the computation time (including both training and testing) was large. It was 811.0056 s and 1838.412 s using a fixed partition scheme, and the same surges to 1121.319 s and 2917.112 s.

5.6 Performance comparison of transfer learning-based pre-trained DCNN models with the recent literature for challenging multiclass Figshare dataset

The performance of the transferred DCNN models is also investigated on the challenging multiclass tumor dataset from the Figshare [43] having meningiomas, glioma, and pituitary tumors. The detailed results in terms of all the performance measures are presented in Table 2 of "Results" section. The comparative results of the best performing transferred model (Pre-trained Alexnet model with transfer learning) in contrast to existing state-of-the-art works are given in Tables 5 and 6. The comparative analysis in Table 5 shows that the pre-trained Alexnet model with transfer learning achieved an average value of 96.98, 97.25, 96.87, 96.916; 91.37, 97.71, 92.44, 91.85; and 98.17, 98.97, 97.65, 97.90 for recall, specificity, precision, and F1-score for glioma, meningioma, and pituitary tumors. The average measures are better than the values of 95.97, 93.79, 93.26, 94.52; 89.98, 96.42, 87.97, 88.88; and 96.81, 93.93, 87.34, 91.80 reported in the works by [30].

Also, in contrast to the existing works that have reported the performance in terms of overall accuracy, the best performing transferred model achieves the higher value of accuracy under the fixed partition and the cross-validation scenario as provided in Table 6. It achieved an accuracy

Table 4 Performance comparison of transfer learning-based pre-trained DCNN models with the recent literature for versions 2 and 3 of the dataset

References	Methods	No. of features	Accuracy (%)		Computation time (s)	Partition scheme
			Version 2	Version 3		
[52]	Discrete wavelet transform (DWT) + SVM	4716	98.00	97.33	–	Fixed partition
[12]	DWT + PCA + PSO optimized kernel SVM	1024	–	97.78	–	Fivefold cross-validation (CV)
[13]	DWT + BBO optimized SVM	10	–	97.78	–	5 × 5-fold CV
[14]	DWT + BAT optimized extreme learning machine	7	–	98.33	–	10 × 10-fold CV
[53]	DWT, PCA and feed-forward back-propagation ANN (FP-ANN)	7	97.00	96.98	–	Fixed partition
[53]	DWT, PCA, and KNN	7	98.00	97.54	–	Fixed partition
[56]	DWT, PCA, and back-propagation neural network (BPNN) with scaled conjugate gradient	19	100.00	98.29	0.0451*	Fixed partition
[57]	DWT + PCA + kernel SVM (KSVM) + Gaussian radial basis (GRB)	19	100.00	99.38	0.0448*	5 CV
[58]	DWT + wavelet entropy (WE) + spider web plots (SWP) + probabilistic neural network (PNN)	3	100.00	99.88	0.036*	Fixed partition
[60]	Ripplet transform (RT) + PCA + LS-SVM	9	100.00	100	0.042*	Stratified CV: 5 × 6-fold Stratified CV: 5 × 5-fold
[54]	DWT, entropy, and SVM	8	97	–	60*	LOOM
[54]	EMD, entropy, and SVM	5	99	–	60*	LOOM
[59]	Feedback pulse-coupled neural network (FPCNN) + DWT + PCA + FP-ANN	7	100.00	98.88	–	Fixed partition
[15]	Stationary wavelet transform (SWT) + PCA + ABC-PSO integrated algorithm (IABAP) feed-forward neural network (FNN)	7	100.00	99.44	219.077# 0.016*	Stratified CV: 10 × 6-fold Stratified CV: 10 × 5-fold
[15]	SWT + PCA + ABC-PSO + FNN	7	100.00	99.75		Stratified CV: 10 × 6-fold Stratified CV: 10 × 5-fold
[16]	WE + hybridization of BBO and PSO (HBP)-FNN	6	100.00	100	208.2510# 0.053*	Sixfold CV Fivefold CV
[55]	Discrete wavelet packet transform (DWPT) + Shannon entropy (SE) + generalized eigenvalue proximal SVM (GEPSVM)	16	99.85	99.62	8.4430# 0.1059*	Stratified CV: 5 × 6-fold Stratified CV: 5 × 5-fold
[55]	DWPT + Tsallis entropy (TE) + GEPSVM	16	100.00	100		
[61]	Bi-dimensional EMD (BEMD), Kruskal–Wallis, and LS-SVM	Multiple AR coefficients	100.00	–	–	Tenfold CV
[61]	BEMD, Kruskal–Wallis, and LS-SVM	Multiple AR coefficients	100.00	–	–	LOOM
[44]	DWT + PCA + AdaBoost with random forests (ADBRF)	13	100.00	99.30	0.0243*	Stratified CV: 5 × 6-fold Stratified CV: 5 × 5-fold
[44]	DWT + probabilistic principal component analysis (PPCA) + ADBRF	13	100.00	100.00		
[62]	Ripplet II Features + PCA + LDA + PSO-based extreme learning machine classifier preceded by contrast enhancement procedure	3	100	100	–	Sixfold Fivefold
[63]	Pseudo-Zernike moment + kernel SVM	400	100	99.75	9.12*	Stratified CV: 10 × 6-fold Stratified CV: 10 × 5-fold
[64]	Density measure coupled with quantitative metric-based classification	1	100.00	100.00	5.0000# 0.23985*	Fivefold CV
[31]	ResNet34	00	–	100	310#	Fivefold CV
[32, 33]	Alexnet	00	100.00	100.00	811.0056# 1838.4124#	Fixed partition

–Performance metrics and other statistics not reported

*Refer to the time for feature extraction, reduction and categorization excluding the time required for classifier training

#Signify the total computational time

Table 5 Fivefold classification recall, specificity, precision, and F1-score by proposed method on the test datasets

Tumor types	Test set	Recall [30]	Recall proposed	Specificity [30]	Specificity proposed	Precision [30]	Precision proposed	F1-score [30]	F1-score proposed
Glioma	Test set 1	96.59	98.601	96.41	95.122	95.86	94.631	96.23	96.575
	Test set 2	98.63	96.491	88.12	97.866	87.05	97.518	92.48	97.002
	Test set 3	96.71	96.842	95.30	98.476	94.00	98.221	95.33	97.527
	Test set 4	91.40	98.246	97.31	97.554	96.59	97.222	93.92	97.731
	Test set 5	96.54	94.737	91.80	97.248	92.80	96.774	94.63	95.745
	Average	95.97	**96.9834**	93.79	**97.2532**	93.26	**96.8732**	94.52	**96.916**
Meningioma	Test set 1	93.01	87.324	97.66	98.941	93.01	96.124	93.01	91.513
	Test set 2	82.14	92.958	96.91	97.452	89.03	91.667	85.45	92.308
	Test set 3	92.98	94.366	96.65	98.301	87.60	94.366	90.21	94.366
	Test set 4	93.48	90.007	94.95	98.089	84.31	93.382	88.66	91.697
	Test set 5	88.28	92.199	95.95	95.754	85.91	86.667	87.07	89.347
	Average	89.98	**91.3708**	96.42	**97.7074**	87.97	**92.4412**	88.88	**91.8462**
Pituitary	Test set 1	99.39	98.387	95.33	99.065	89.50	97.861	94.19	98.123
	Test set 2	93.30	97.849	92.62	98.829	84.19	97.326	88.51	97.587
	Test set 3	96.59	99.462	95.52	98.829	92.52	97.368	94.51	98.404
	Test set 4	100.00	97.849	92.09	98.592	85.59	96.809	92.24	97.326
	Test set 5	94.77	97.312	94.11	99.531	84.90	98.907	89.56	98.103
	Average	96.81	**98.1718**	93.93	**98.9692**	87.34	**97.6542**	91.8	**97.9086**

Bold specifies the best value of the performance measures

Table 6 Comparison with related works using Figshare dataset

Work	Method	Training data	Accuracy (%)
Jun Cheng [65]	BoW-SVM	80%	91.28
Ismael [67]	DWT–Gabor–NN	70%	91.90
Pashaei [66]	CNN–ELM	70%	93.68
Nyoman [24]	CNN		84.19
Afshar [25]	CapsNet	–	90.89
Deepak and Ameer [29]	Pre-trained GoogLeNet (stand-alone)	Fivefold cross-validation	92.3 ± 0.7
Swati [30]	Pre-trained VGG-16 with block-wise fine-tuning strategy	Fivefold cross-validation	94.82
Proposed method	Pre-trained Alexnet with transfer learning	70%	96.95
		Fivefold cross-validation	96.1 ± 0.68

value of 96.95% and $96.1\% \pm 0.68$ for fixed partition and cross-validation partitioning scheme better than the 91.28%, 93.68%, and 94.82% obtained by [30, 65, 66].

5.7 Advantages of the proposed method

The advantages of using pre-trained DCNN models with transfer learning for classification are manifold: firstly, the classification mechanism is fully automated, secondly it eliminates the conventional steps of noise filtering, ROI delineation, feature extraction, and selection, thirdly no inter- and intra-observer biases are there and the predictions made by the pre-trained DCNN models are reproducible, and finally, ceiling level of accuracy is achieved in contrast to similar works reported using DCNN [24, 25, 29–31, 65–67].

5.8 The pitfalls of the using the pre-trained CNN models with transfer learning for classification are as follows

The models are computationally intensive to run on a single CPU system. The computation time for transfer learning is nearly 10 min, 13.5 min, 31 min, 44 min 14 s, and 72 min 08 s for the best performing pre-trained Alexnet model with transfer learning.

The reason behind higher computation time this that all the models are implemented using the MATLAB 2017b platform

and executed them on a system with configuration as Intel Xeon CPU E3-1225, 3.3 GHz, 8 GB RAM. Since we have executed the code on a MATLAB platform with a single CPU as the hardware resource, the computation time is large. Just for indication, experimenting on the challenging dataset from Figshare took 66 min and 42 s using transfer learning which is comparable to a value of 55 min reported in the work by [29] using MATLAB 2018b platform with single CPU, 32 GB RAM (Intel E3-1245v6, 3.70 GHz) and 25% training data. In another work by [30], the time for transfer learning was approximately 20-30 min when implemented on a MATLAB 2017b platform with GPU capability (NIVIDA TITAN X) with 12 GB onboard memory.

Future works will be focused on running the models on the systems with GPU-enabled capability that is expected to reduce the computational overhead.

6 Conclusion

In the present paper, we have compared different pre-trained DCNN models with transfer learning for MR brain image classification. The approach of using pre-trained DCNN models with transfer learning was successful in providing a ceiling level of recognition rate. Out of all the evaluated models, the Alexnet model proved to be the finest by rendering a 100%, 94%, and 95.92% classification results for all the three datasets. The results were superior to the existing classical and deep learning methods reported on brain image classification. It is superior in the context of the elimination of the preprocessing stages of ROI delineation, feature extraction, and selection in contrast to the classical methods. Moreover, in contrast to the recent works using deep learning the transferred Alexnet model achieve the higher value of performance measures. Future works will be focused on running the models on the systems with GPU-enabled capability that is expected to reduce the computational overhead and exploring different fine-tuning strategies.

Acknowledgements The authors would like to thank Head, Neurocomputing Lab (Dr. Tapan Kumar Gandhi), Department of Electrical Engineering, IIT Delhi, for providing the resources for carrying out the research work.

References

1. Lee, M.C., Nelson, S.J.: Supervised pattern recognition for the prediction of contrast-enhancement appearance in brain tumors from multivariate magnetic resonance imaging and spectroscopy. Artif. Intell. Med. **43**, 61–74 (2008)
2. Zacharaki, E.I., Wang, S., Chawla, S., Soo, D.: Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn. Reson. Med. **62**, 1609–1618 (2009)
3. Wang, S., Kim, S., Chawla, S., Wolf, R.L., Zhang, W., Rourke, D.M.O., Judy, K.D., Melhem, E.R., Poptani, H.: Differentiation between glioblastomas and solitary brain metastases using diffusion tensor imaging. Neuroimage **44**, 653–660 (2010)
4. Hemanth, D.J., Vijila, C.K.S., Selvakumar, A.I., Anitha, J.: Performance enhanced hybrid kohonen-hopfield neural network for abnormal brain image classification. In: Kim, T., Adeli, H., Ramos, C., Kang, B.-H. (eds.) Signal processing, image processing and pattern recognition, pp. 356–365. Springer (2011)
5. Zollner, F.G., Emblem, K.E., Schad, L.R., Zöllner, F.G., Emblem, K.E., Schad, L.R.: SVM-based glioma grading: optimization by feature reduction analysis. J. Med. Phys. **22**, 205–214 (2012)
6. Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N., Ahuja, C.K.: Segmentation, feature extraction, and multiclass brain tumor classification. J. Digit. Imaging **26**, 1141–1150 (2013)
7. Skogen, K., Schulz, A., Dormagen, J.B., Ganeshan, B., Helseth, E., Server, A.: Diagnostic performance of texture analysis on MRI in grading cerebral gliomas. Eur. J. Radiol. **85**, 824–829 (2016)
8. Bahadure, N.B., Ray, A.K., Thethi, H.P.: Comparative approach of MRI-based brain tumor segmentation and classification using genetic algorithm. J. Digit. Imaging **31**, 477–489 (2018)
9. Lahmiri, S.: Glioma detection based on multi-fractal features of segmented brain MRI by particle swarm optimization techniques. Biomed. Signal Process. Control **31**, 148–155 (2017)
10. Gupta, N., Bhatele, P., Khanna, P.: Glioma detection on brain MRIs using texture and morphological features with ensemble learning. Biomed. Signal Process. Control **47**, 115–125 (2019)
11. Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N., Ahuja, C.K.: A package-SFERCB-"Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors". Appl. Soft Comput. **47**, 151–167 (2016)
12. Zhang, Y., Wang, S., Ji, G., Dong, Z.: An MR brain images classifier system via particle swarm optimization and kernel support vector machine. Sci. World J. **2013**, 1–9 (2013)
13. Yang, G., Zhang, Y., Yang, J., Ji, G., Dong, Z., Wang, S., Feng, C., Wang, Q.: Automated classification of brain images using wavelet-energy and biogeography-based optimization. Multimed. Tools Appl. **75**, 15601–15617 (2016)
14. Lu, S., Qiu, X., Shi, J., Li, N., Lu, Z.-H., Chen, P., Yang, M.-M., Liu, F.-Y., Jia, W.-J., Zhang, Y.: A pathological brain detection system based on extreme learning machine optimized by bat algorithm. CNS Neurol. Disord. Targets (Formerly Curr. Drug Targets-CNS Neurol. Disord) **16**, 23–29 (2017)
15. Wang, S., Zhang, Y., Dong, Z., Du, S., Ji, G., Yan, J., Yang, J., Wang, Q., Feng, C., Phillips, P.: Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. Int. J. Imaging Syst. Technol. **25**, 153–164 (2015)
16. Zhang, Y., Wang, S., Dong, Z., Phillip, P., Ji, G., Yang, J.: Pathological brain detection in magnetic resonance imaging scanning by wavelet entropy and hybridization of biogeography-based optimization and particle swarm optimization. Prog. Electromagn. Res. **152**, 41–58 (2015)
17. Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adeli, H.: Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. **100**, 270–278 (2017)
18. Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., Greenspan, H.: Chest pathology detection using deep learning with non-medical training. In: ISBI. pp. 294–297 (2015)
19. Zhou, M., Tian, C., Cao, R., Wang, B., Niu, Y., Hu, T., Guo, H., Xiang, J.: Epileptic seizure detection based on EEG signals and CNN. Front. Neuroinform. **12**, 95 (2018)
20. Hoo-Chang, S., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset

characteristics and transfer learning. IEEE Trans. Med. Imaging **35**, 1285–1298 (2016)
21. Yousefi, M., Krzyżak, A., Suen, C.Y.: Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning. Comput. Biol. Med. **96**, 283–293 (2018)
22. Zuo, H., Fan, H., Blasch, E., Ling, H.: Combining convolutional and recurrent neural networks for human skin detection. IEEE Signal Process. Lett. **24**, 289–293 (2017)
23. Charron, O., Lallement, A., Jarnet, D., Noblet, V., Clavier, J.-B., Meyer, P.: Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. Comput. Biol. Med. **95**, 43–54 (2018)
24. Abiwinanda, N., Hanif, M., Hesaputra, S.T., Handayani, A., Mengko, T.R.: Brain tumor classification using convolutional neural network. World Congr. Med. Phys. Biomed. Eng. **2018**, 183–189 (2019)
25. Afshar, P., Plataniotis, K.N., Mohammadi, A.: Capsule networks for brain tumor classification based on mri images and coarse tumor boundaries. In: ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP). pp. 1368–1372 (2019)
26. Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., Budak, Ü.: Transfer learning based histopathologic image classification for breast cancer detection. Heal. Inf. Sci. Syst. **6**, 18 (2018)
27. Hussein, S., Kandel, P., Bolan, C.W., Wallace, M.B., Bagci, U.: Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches. IEEE Trans. Med. Imaging **38**, 1777–1787 (2019)
28. Ahmed, K.B., Hall, L.O., Goldgof, D.B., Liu, R., Gatenby, R.A.: Fine-tuning convolutional deep features for MRI based brain tumor classification. In: Medical imaging 2017: computer-aided diagnosis, p. 101342E (2017)
29. Deepak, S., Ameer, P.M.: Brain tumor classification using deep CNN features via transfer learning. Comput. Biol. Med. **111**, 103345 (2019)
30. Swati, Z.N.K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., Lu, J.: Brain tumor classification for MR images using transfer learning and fine-tuning. Comput. Med. Imaging Graph. **75**, 34–46 (2019)
31. Talo, M., Baloglu, U.B., Yıldırım, Ö., Acharya, U.R.: Application of deep transfer learning for automated brain abnormality classification using MR images. Cogn. Syst. Res. **54**, 176–188 (2019)
32. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M.: others: imagenet large scale visual recognition challenge. Int. J. Comput. Vis. **115**, 211–252 (2015)
33. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105. Lake Tahoe, NV (2012)
34. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9 (2015)
35. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)
36. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv Prepr. arXiv:1409.1556 (2014)
37. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826 (2016)
38. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: AAAI, p. 12 (2017)
39. Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J.: Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans. Med. Imaging **35**, 1299–1312 (2016)
40. Harvard Medical School, http://med.harvard.edu/AANLIB/
41. Gupta, T., Gandhi, T.K., Gupta, R.K., Panigrahi, B.K.: Classification of patients with tumor using MR FLAIR images. Pattern Recognit. Lett. (2017). https://doi.org/10.1016/j.patrec.2017.10.037
42. Gupta, T., Gandhi, T.K., Panigrahi, B.K.: Multi-sequential MR brain image classification for tumor detection. J. Intell. Fuzzy Syst. **32**, 3575–3583 (2017)
43. Cheng, J.: brain tumor dataset, https://figshare.com/articles/brain_tumor_dataset/1512427
44. Nayak, D.R., Dash, R., Majhi, B., Ranjan, D., Dash, R., Majhi, B.: Brain MR image classification using two-dimensional discrete wavelet transform and AdaBoost with random forests. Neurocomputing **177**, 188–197 (2016)
45. Fawcett, T.: ROC graphs: notes and practical considerations for researchers. Mach. Learn. **31**, 1–38 (2004)
46. Zollner, F.G., Emblem, K.E., Schad, L.R.: SVM-based glioma grading: Optimization by feature reduction analysis. J. Med. Phys. **22**, 205–214 (2012)
47. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. Mach. Learn. **46**, 389–422 (2002)
48. Rakotomamonjy, A.: Variable selection using SVM-based criteria. J. Mach. Learn. Res. **3**, 1357–1370 (2003)
49. Emblem, K.E., Nedregaard, B., Hald, J.K., Nome, T., Due-Tonnessen, P., Bjornerud, A.: Automatic glioma characterization from dynamic susceptibility contrast imaging: brain tumor segmentation using knowledge-based fuzzy clustering. J. Magn. Reson. Imaging **30**, 1–10 (2009)
50. Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N., Ahuja, C.K.: A novel content-based active contour model for brain tumor segmentation. Magn. Reson. Imaging **30**, 694–715 (2012)
51. Hemanth, J.D., Anitha, J.: Modified Genetic Algorithm approaches for classification. Appl. Soft Comput. J. **75**, 21–28 (2019)
52. Chaplot, S., Patnaik, L.M., Jagannathan, N.R.: Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomed. Eng. Online **1**, 86–92 (2006)
53. El-Dahshan, E.-S.A., Hosny, T., Salem, A.-B.M.: Hybrid intelligent techniques for MRI brain images classification. Digit. Signal Process. **20**, 433–441 (2010)
54. Lahmiri, S.: An Application of the empirical mode decomposition to brain magnetic resonance images classification. In: Fourth latin American symposium on circuits and systems (LASCAS), pp. 1–4 (2013)
55. Zhang, Y., Dong, Z., Wang, S., Ji, G., Yang, J.: Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine (GEPSVM). Entropy **17**, 1795–1813 (2015)
56. Zhang, Y., Dong, Z., Wu, L., Wang, S.: A hybrid method for MRI brain image classification. Expert Syst. Appl. **38**, 10049–10053 (2011)
57. Zhang, Y., Wu, L.: An MR brain images classifier via principal component analysis and kernel support vector machine. Prog. Electromagn. Res. **130**, 369–388 (2012)
58. Saritha, M., Joseph, K.P., Mathew, A.T.: Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. Pattern Recognit. Lett. **34**, 2151–2156 (2013)
59. El-Dahshan, E.-S.A., Mohsen, H.M., Revett, K., Salem, A.-B.M.: Computer-aided diagnosis of human brain tumor through MRI: a

survey and a new algorithm. Expert Syst. Appl. **41**, 5526–5545 (2014)
60. Das, S., Chowdhury, M., Kundu, M.K.: Brain MR image classification using multiscale geometric analysis of ripplet. Prog. Electromagn. Res. **137**, 1–17 (2013)
61. Sahu, O., Anand, V., Kanhangad, V., Pachori, R.B.: Classification of magnetic resonance brain images using bi-dimensional empirical mode decomposition and autoregressive model. Biomed. Eng. Lett. **5**, 311–320 (2015)
62. Nayak, D.R., Dash, R., Majhi, B.: Discrete ripplet-II transform and modified PSO based improved evolutionary extreme learning machine for pathological brain detection. Neurocomputing **282**, 232–247 (2018)
63. Zhang, Y.-D., Jiang, Y., Zhu, W., Lu, S., Zhao, G.: Exploring a smart pathological brain detection method on pseudo Zernike moment. Multimed. Tools Appl. **77**, 22589–22604 (2018)
64. Kaur, T., Saini, B., Gupta, S.: Quantitative metric for MR brain tumor grade classification using sample space density measure of analytic intrinsic mode function representation. IET Image Process. **11**, 620–632 (2017)
65. Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z., Feng, Q.: Enhanced performance of brain tumor classification via tumor region augmentation and partition. PLoS ONE **10**, e0140381 (2015)
66. Pashaei, A., Sajedi, H., Jazayeri, N.: Brain tumor classification via convolutional neural network and extreme learning machines. In: 2018 8th International conference on computer and knowledge engineering (ICCKE), pp. 314–319 (2018)
67. Ismael, M.R., Abdel-Qader, I.: Brain tumor classification via statistical features and back-propagation neural network. In: 2018 IEEE international conference on electro/information technology (EIT), pp. 252–257 (2018)