



# A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning

Arshia Rehman<sup>1</sup> · Saeeda Naz<sup>1</sup> · Muhammad Imran Razzak<sup>2</sup>  · Faiza Akram<sup>3</sup> · Muhammad Imran<sup>4</sup>

Received: 3 February 2019 / Revised: 12 August 2019 / Accepted: 17 August 2019 /  
Published online: 10 September 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Brain tumors are the most destructive disease, leading to a very short life expectancy in their highest grade. The misdiagnosis of brain tumors will result in wrong medical intercession and reduce chance of survival of patients. The accurate diagnosis of brain tumor is a key point to make a proper treatment planning to cure and improve the existence of patients with brain tumors disease. The computer-aided tumor detection systems and convolutional neural networks provided success stories and have made important strides in the field of machine learning. The deep convolutional layers extract important and robust features automatically from the input space as compared to traditional predecessor neural network layers. In the proposed framework, we conduct three studies using three architectures of convolutional neural networks (AlexNet, GoogLeNet, and VGGNet) to classify brain tumors such as meningioma, glioma, and pituitary. Each study then explores the transfer learning techniques, i.e., fine-tune and freeze using MRI slices of brain tumor dataset—Figshare. The data augmentation techniques are applied to the MRI slices for generalization of results, increasing the dataset samples and reducing the chance of over-fitting. In the proposed studies, the fine-tune VGG16 architecture attained highest accuracy up to 98.69 in terms of classification and detection.

**Keywords** Brain tumor · Deep learning · Transfer learning · AlexNet · GoogLeNet · VGG · Figshare dataset

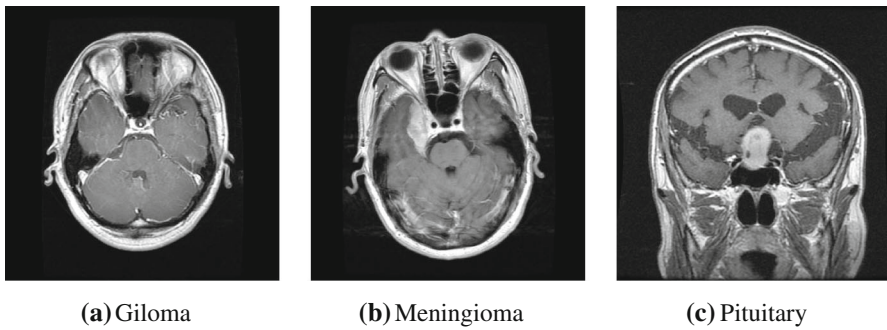
## 1 Introduction

Over the past decades, diseases have stumbled that are overcome with the human intelligence and biomedical advance, but still cancer, by virtue of its unstable nature,

---

Imran's work is supported by the Deanship of Scientific Research, King Saud University through research group Project Number RG-1435-051.

Extended author information available on the last page of the article



**Fig. 1** Types of brain tumor

remains a curse to the mankind. One of the fatal and most growing diseases is brain tumor cancer. Brain is the core and most complex organ of human body that comprises nerve cells and tissues to control the foremost activities of the entire body like breathing, movement of muscles and our senses. Every cells have their own capabilities; some cells grow with their own functionality, and some lose their capacity, resist, and grow aberrant. These mass collections of abnormal cells form the tissue are called as tumor. Cancerous brain tumors are uncontrolled and unnatural growth of brain cells [12]. It is one of the most life-threatening and lethal cancers. In 2015 [23], approximately 23,000 patient were diagnosed brain tumor in the USA. According to 2017 cancer statistics [22], brain tumor is measured as one of the foremost causes of cancer-related indisposition, morbidity, and mortality around the world both in children and in adults.

Generally, brain tumor can be classified into two types, i.e., benign and malignant tumors. Benign tumor is a non-cancerous type (non-progressive), and it is originated in the brain and is growing slowly. This type of tumor cannot spread anywhere else in the body thus assumed to be less aggressive. The abnormal growth of cell can press tissue or part of brain which can be removed on time. On contrary, malignant tumor type is a cancerous, produce quickly with undefined boundaries, invade other healthy cells, and spread other parts of the body. When this type of tumor is originated in the brain, then it is known as primary malignant tumor. When it is emanated elsewhere in the body and spread to the brain, then it is known as secondary malignant tumor [2]. However, meningioma, glioma, and pituitary tumors are the other common types of brain tumors. Meningiomas are the most common benign tumors that instigate in the thin membranes that surround the brain and spinal cord. The gliomas are assortment of tumors that grow within the substance of the brain [1]. High-grade gliomas are one of the aggressive brain tumors with a minimum survival of almost two years. Pituitary tumors are irregular growth of the brain cells. Pituitary tumors develop in the pituitary gland of the brain. These tumors have uniform shape and intrinsic nature that can produce anywhere in the brain. These types of brain tumors are depicted from Fig. 1.

The progression of brain tumor classification is one of the foremost challenging tasks due to the heterogeneity, isointense and hypo-intense properties, and related perilesional edema creates obscurity in tumor classification. Normally, T1-weighted

contrast-enhanced images are used for classification of primary tumors such as meningioma (MEN), glioblastoma multiforme (GBM), astrocytoma (AS), medulloblastoma (MED), and secondary tumors like metastases (MET). These tumors are significantly better visualized on T1-weighted contrast-enhanced images due to the stimulation of 0.150.20 mMol/kg of contrast material (gadolinium) in the patients. Classification of brain tumors is accomplished with the aid of features. Expedient features are extracted using the intricate structures of diverse tumors on brain magnetic resonance images (MR). The main methodologies of brain tumors classification conventionally rely on region-based tumor segmentation rather than feature extraction and classification. Thus, the paradigm shifted toward the classification tasks with the aid of deep learning approaches.

Deep learning is the subfield of machine learning that provides the capability to the computer to make predictions and take conclusions on data with its ability of learning data representations. Specifically, these approaches are extensively used for medical imaging classification and act as one of major computational intelligence techniques. Although deep learning approaches have revealed incredible success in a diversity of applications in numerous domains in different fields [10,11,13,14,14,16,17,21], nevertheless, it is data starving approach and necessitates as a minimum ten times the degree of freedom data samples. To address the challenge of limited training samples, transfer learning could be employed to fine-tune the already gained storing information on similar problem. Transfer learning is the deep learning technique in which the network is trained on the large dataset (base dataset) and transfers the learned knowledge to the small dataset (target dataset) [18]. Two main scenarios of transfer learning are: fine-tune the ConvNet and freeze the layers of ConvNet. In former scenario, replace and retain the pretrained ConvNet on the target dataset for continuing backpropagation. Finally, the last fully connected layer classifies the target dataset. While in later case, ConvNet is pretrained on the base dataset and the last fully connected layers are removed. In this way, fully connected or desired convolutional layer acts as features were passed to the linear classifier (SVM, NN, LDA, Naive Bayes, etc.) for classification. The motivation of this paper is to perform an extensive experiments using deep convolutional neural network (CNN), transfer learning, and its scenarios.

The main contributions of our proposed studies are:

- A novel and robust techniques of transfer and deep learning are presented for classification and automated detection of brain tumor that is effective in extraction of important and rich features on a benchmark Figshare dataset
- To explore the three different architectures of deep convolutional networks like AlexNet, GoogLeNet, and VGGNet using MRI images of brain tumor and deploy transfer learning techniques on the target dataset.
- To deliver an in-depth performance assessment of the critical factors affecting the fine-tuning approach of pretrained models.
- To employ the different freeze layers of pretrained model and then pass to the support vector machine (SVM) for classification.
- To conduct the comparative analysis in terms of accuracy of each architecture of CNN for brain tumor classification and detection.

The rest of paper is structured as: Sect. 2 deliberates the closely related work. Section 3 explains the brain tumor classification and the CNN architectures in detail. Experimental design of the proposed brain tumor classification and detection system thoroughly discusses in Sect. 4. Section 5 details the experimental results along with the comparison of existing systems. Lastly, the conclusion with future direction has been drawn in Sect. 6.

## 2 Related Work

Numerous methods with solutions for identification of the brain tumor using MRI images had been proposed by number of researchers in the past years. These methods vary from conventional machine learning algorithms to the deep learning models. We here present the related work of brain tumor detection rely on brain tumor dataset—Figshare [4]. In this regard, Cheng et al. [6] conducted experiment on brain tumor dataset—Figshare. They used augmented tumor region as region of interest and split these regions into subregions by employing adaptive spatial division method. They extracted intensity histogram, gray-level co-occurrence matrix (GLCM), and bag-of-words (BoW) model-based features. They reported highest accuracy of 87.54%, 89.72%, and 91.28% on extracted features using ring-form partition method. Another contribution of same authors was presented in [5]. They deployed Fisher Vector for the aggregation of local features from each subregion. Mean average precision (map) of 94.68% was retrieved. Similarly, Ismael and Abdel-Qader [8] extracted statistical features from MRI slices with the aid of 2D discrete wavelet transform (DWT) and Gabor filter techniques. They classified the brain tumors using backpropagation multilayer perceptron neural network and retrieved highest accuracy of 91.9%. Abir et al. [1] deployed probabilistic neural network (PNN) for classification of brain tumors. They performed image filtering, sharpening, resize, and contrast enhancement in pre-processing and extracted GLCM features. They attained highest accuracy of 83.33%.

Still the available automated tumor detection systems are not providing satisfactory output, and there is a big demand to get a robust automated computer-aided diagnosis systems for brain tumor detection. The conventional machine learning-based algorithms and models require domain specific expertise and experience. These methods need efforts for segmentation and manual extraction of structural or statistical features which may result in degradation of accuracy and efficiency of the system performance [15]. The deep transfer learning-based techniques overcome these issues due extraction of visual and discriminative features using different convolutional layers automatically. These extracted features are supposed to be rich and robust for classification purpose. Widhiarso et al. [26] computed GLCM and fed to the convolutional neural network. They claimed that GLCM combined with contrast feature gave 20% improved accuracy. They achieved highest accuracy of 82% using this scenario. Table 1 elaborates the related work of Figshare dataset. Abiwinanda et al. [2] employed five different architectures of CNN and reported highest accuracy on architecture 2. The architecture 2 contains two convolutional layers, ReLU layer, and max-pool followed by 64 hidden neurons. They achieved 98.51% and 84.19% on training and validation sets, respectively. Afshar et al. [3] proposed a novel model Capsule networks

**Table 1** Related work of brain tumor detection using Figshare dataset

References	Features	Model	Accuracy (%)
Cheng et al. [6]	Intensity histogram	SVM	87.54
	GLCM		89.72
	Bag of words		91.28
Cheng et al. [5]	Local features using Fisher Vector	SVM	94.68
Ismael and Qader [8]	Fabor features	MPNN	98.9
Abir et al. [1]	GLCM	PNN	83.33
Widhiarso et al. [26]	GLCM	CNN	82
Abiwinanda et al. [2]	Model based	CNN	84.19
Afshar et al. [3]	Model based	CapsNet	86.56

*MPNN* Multilayer perceptron neural network, *SVM* support vector machine, *CNN* convolutional neural network, *PNN* probabilistic neural network, *CapsNet* capsule network, *GLCM* gray-level co-occurrence matrix

(CapsNets) for the classification of brain tumor. They varied the feature maps in the convolutional layer of CapsNet in order to increase accuracy. They achieved highest accuracy of 86.56% using 64 feature maps with one convolutional layer of CapsNet. Table 1 summarizes the state-of-the-art techniques presented by numerous researchers for brain tumor detection and classification using manual features with conventional networks and model-based features with deep neural networks.

### 3 Materials, Methods, and Measurement's Metrics

We use the Caffee library 7 for our implementation. The Caffee library 7 is used with the single GPU (NVIDIA CUDA) having the multiple processors of 2.80 GHz, 16GB DDR4-SDRAM, 1TB HDD, 128 GB SSD. In this section, we present different architectures of deep convolution neural network for proposed framework using brain tumor dataset—Figshare [4], for brain tumor classification and detection. We explore and evaluate the well-known CNN architectures (AlexNet, GoogLeNet, and VGG16) using augmented MRI slices of brain tumor dataset. These pretrained CNN architectures are used to deploy the transfer learning techniques to extract the visual discriminative and rich features. Finally, the visual patterns are classified using log-based softmax layer or support vector machine (SVM). The key elements of proposed framework are discussed in the subsections. Then, the measurement matrices are discussed for evaluation for performance of our proposed systems.

#### 3.1 Dataset

We used publicly available brain tumor dataset—Figshare [4], to analyze and evaluate our proposed framework using different architectures of CNN. It is developed by Cheng in 2017. The dataset comprises 3064 brain MRI slices collected from 233 patients. It contains three kinds of brain tumors: meningioma, pituitary, and glioma.

The total number of images(slices) for meningioma is 708, 930 for pituitary, and 1426 for glioma tumor. The dataset is publicly available on Figshare Web site in MATLAB “.mat” format. Each MAT-file contains a structure comprising a patient ID, unique label that demonstrates the type of brain tumor,  $512 \times 512$  image data in uint16 format, vector containing tumor border with the coordinates of discrete points, and ground truth in binary mask image. In our experiments, each CNN model takes image as input unit; thus, we only use the image data from the .mat files as depicted in Fig. 1.

### 3.2 CNN-Based Architectures

A CNN or ConvNet is the acronym of convolutional neural network. It has designed for recognizing important visual patterns from raw pixels of images automatically with less preprocessing efforts. The competition of ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) has brought a breakthrough in 2012 [9] and new architectures of deep CNNs presented by increasing the complexity and number of convolutional layers for fantastic and better accuracy on ImageNet dataset [7]. The different CNN architectures coupled with transfer learning techniques, i.e., fine-tune and freeze layers, have made great successes for their improved performance on image classification and beat the traditional machine learning models in the past few years.

We are employing and exploring the three popular and powerful state-of-the-art architectures of CNN for classification and identification task of brain tumor using MRI images of brain tumor dataset—Figshare. A brief and comprehensive framework of architectures of AlexNet, Google, and VGG16 is illustrated in Fig. 2 for our proposed automatic brain tumor classification and identification system.

#### 3.2.1 AlexNet

AlexNet [9] was proposed by Alex Krizhevsky in 2012. It has been successfully trained on ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) dataset [20] that contains 1.2 million natural images of 1,000 different categories. It was the winner of ILSVRC 2012. Its architecture encompasses 60 million parameters, 650,000 neurons, and 630 million connections, with five convolutional layers, max-pooling layer at each three convolution layers, and three fully connected layers. The input layer takes the image size of  $227 \times 227$ .

The first convolutional layer applies 96 filter of  $11 \times 11$  on the input images at stride 4 in Conv1, whereas  $3 \times 3$  filters are applied at stride 2 in pool1. Likewise, the second convolutional layer applies 256 filters of  $5 \times 5$ . Similarly,  $3 \times 3$  filter size is used in the third, fourth, and fifth layer with the filters of 384, 384, and 256. ReLU activation function is applied in each convolution layer. Fully connected layers are fc6 and fc7, and each has 4096 neurons. Furthermore, output layer fc8 uses softmax classifier that initiates 1000 neurons according to the classes of ImageNet.

#### 3.2.2 GoogLeNet

GoogLeNet [25] was developed by Szegedy et al. in 2014. It is the first winner of ILSVRC 2014 trained on ILSVRC dataset. The architecture contains approximately

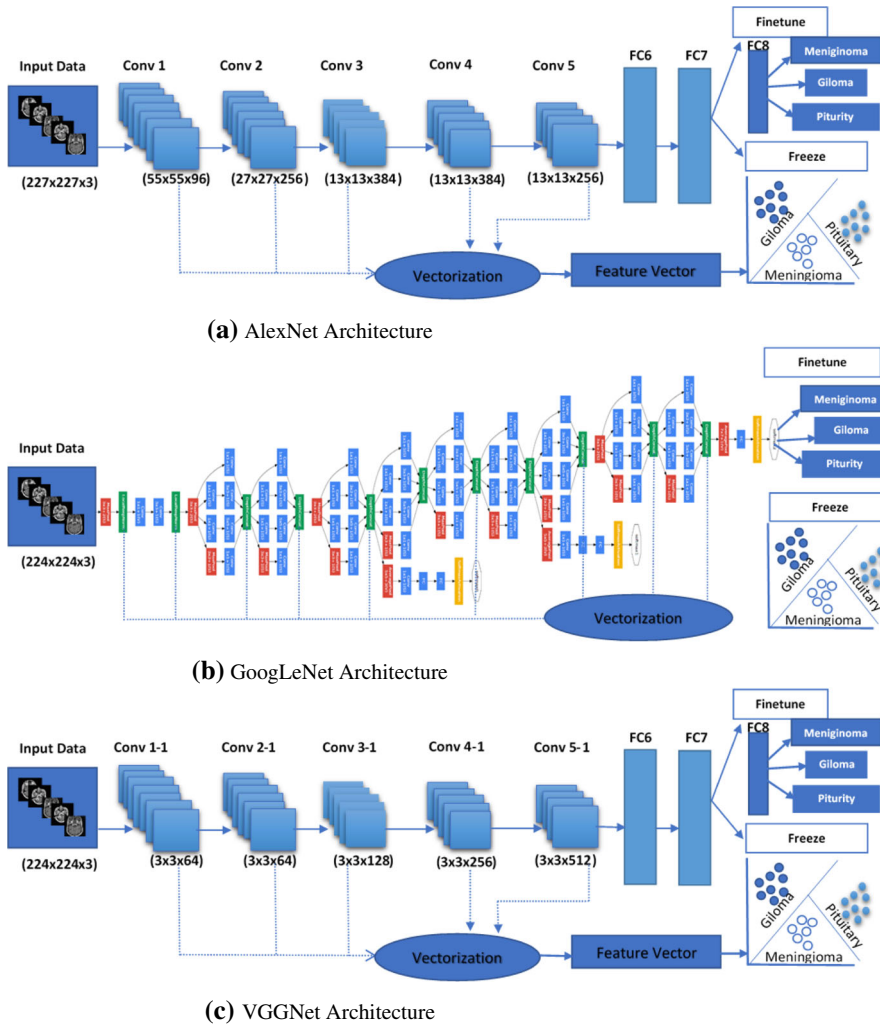


Fig. 2 The overall framework for proposed automatic brain tumor classification system

6.8 million parameters comprising nine inception modules, two convolutional layers, four max-pooling layers, one convolutional layer for dimension reduction, one average pooling, two normalization layers, one fully connected layer, and finally a linear layer with softmax activation in the output. Furthermore, each inception module contains one max-pooling layer and six convolutional layers, from that four convolutional layers are used for dimension reduction. ReLU activation function is applied in all the fully connected layers, and dropout regularization is used in the fully connected layers. Moreover, it is more precise than AlexNet on original ILSVRC dataset.

### 3.2.3 VGGNet

Another CNN architecture proposed in 2014 was VGGNet16 [24], which secured the second position in challenge with respect to accuracy but the first runner-up in ILSVRC 2014. The largest VGGNet16 architecture contains roughly 144 million parameters from 16 convolutional layers, three fully connected layers, five max-pooling layers in each convolutional layer with size of  $2 \times 2$ , and a softmax linear layer in the output. ReLU activation function is applied in all the fully connected layers, and dropout regularization is used in the fully connected layers. It is computationally more expensive CNN model due to the large parameters as compared to AlexNet and GoogLeNet.

### 3.3 Evaluation Metrics

The effectiveness of proposed brain tumor classification and detection system is evaluated by computing evaluation measures based on four major outcomes that are used to test the classifier: true positives (tp), false positives (fp), true negatives (tn), and false negatives (fn). The performance of the proposed system is computed using the following measures:

*Accuracy* determines the ability to differentiate the brain tumor types correctly. To estimate the accuracy of a test, we calculate the proportion of true positive and true negative in all evaluated cases computed by the following relations:

$$\text{Accuracy} = \frac{\text{tp} + \text{fn}}{\text{tp} + \text{tn} + \text{fp} + \text{fn}} \quad (1)$$

*Sensitivity* measures the capability of system to accurately classify the brain tumors and is calculated from the proportion of true positives using relation:

$$\text{Sensitivity} = \frac{\text{tp}}{\text{tp} + \text{fn}} \quad (2)$$

*Specificity* is the capability of the model to accurately classify the actual brain tumor type and is computed as:

$$\text{Specificity} = \frac{\text{tn}}{\text{tn} + \text{fp}} \quad (3)$$

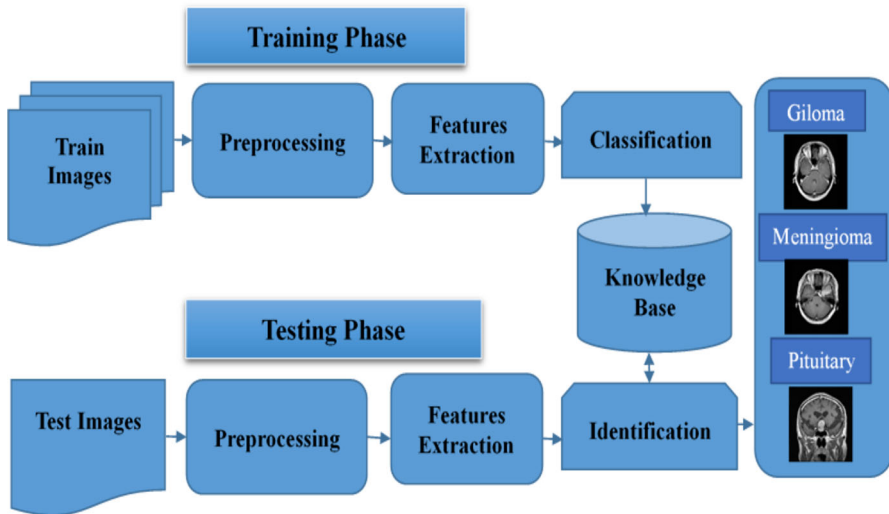
*Precision* is the true positive measure and is computed using relation:

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}} \quad (4)$$

## 4 Experimental Setting

In this section, we provide the comprehensive methods of evaluation for proposed automated brain tumor classification and detection system. We perform the brain





**Fig. 3** Proposed block diagram of brain tumor classification system

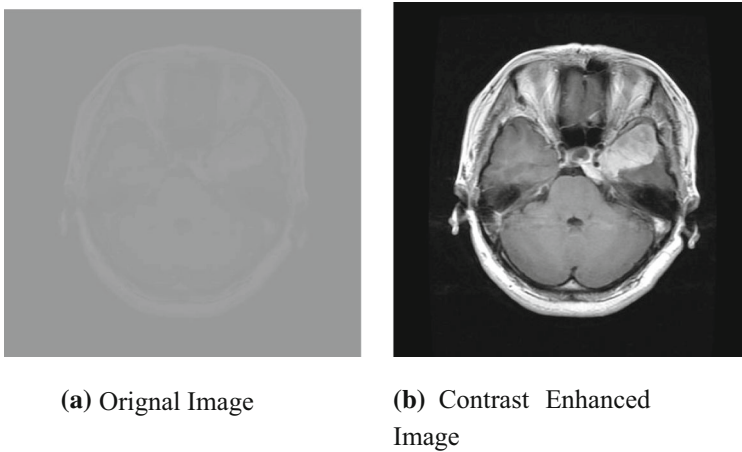
tumor classification using T1-weighted contrast-enhanced images from 233 patients with three kinds of brain tumors. The proposed work employs three pretrained architectures of CNN, i.e., AlexNet, GoogLeNet, and VGGNet. The framework of proposed system contains three main phases: preprocessing, feature extraction, and classification/detection as illustrated in Fig. 3.

In the first phase, MRI images are enhanced using contrast stretching technique. Data augmentation techniques like rotation and flipping are applied to generate the large amount of data for CNN architectures and to reduce over-fitting. In the next step, three pretrained architectures of CNN on ImageNet dataset [7] are employed on a target brain tumor dataset—Figshare, to extract the discriminating visual and distinct features from MRI images. In this phase, the fine-tune and freeze strategies of transfer learning are deploying using each CNN architecture (AlexNet, GoogLeNet, or VGGNet). In final phase, automated features are classified in the last step using linear classifier in case of freeze layers strategy and using log-softmax layer in case of fine-tune layers strategy.

#### 4.1 Preprocessing and Data Augmentation

Preprocessing is the data cleaning phase that enhances and improves the input data for further tasks.

The primary task of medical imaging analysis is to clean the MRI images and to enhance the contrast. The MRI images were obtained from different modalities that cause artifacts and false intensity levels. Thus, different machine learning and image processing algorithms were deployed to enhance the contrast of MRI images. We used the contrast stretching algorithm of preprocessing to generate the high-resolution contrast images. The objective of contrast stretching is to increase the dynamic range



**Fig. 4** Result of contrast enhancement

of gray levels for low-contrast MRI images. The contrast enhancement of MRI images is obtained using the relation:

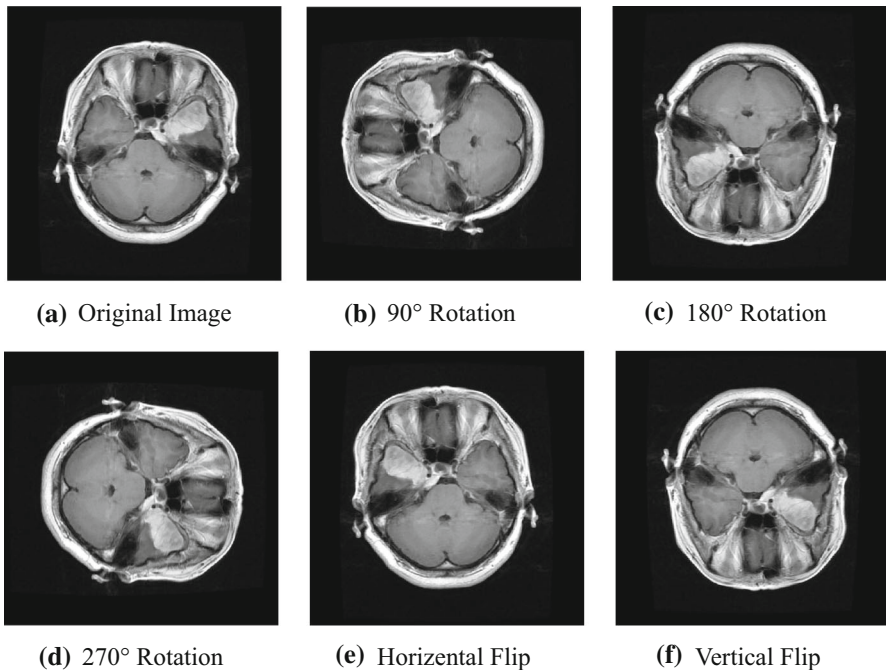
$$g(x, y) = \frac{f(x, y) - f_{\min}}{f_{\max} - f_{\min}} * 2^{\text{bpp}} \quad (5)$$

In our case, MRI images of brain tumor dataset—Figshare, are 8 bpp, so levels of grayscale are 256. The minimum value is 0, and the maximum value is 255. Figure 4 depicts the original and the resultant enhance image.

In order to increase the dataset samples, different variations of images created using traditional data augmentation techniques which help in reducing the over-fitting during the CNNs training. We have applied several data augmentation techniques (rotations and flipping) to increase the training dataset for providing large input space to CNNs. One of the basic data augmentation techniques is rotation in which the input images are rotated on various angles like angle of 90, 180, and 270. Another employed technique is flipping in which image is mirrored from vertical and horizontal direction. Figure 5 shows the resultant images of data augmentation.

## 4.2 CNN-Based Feature Extraction

After data augmentation, an enormous image samples are generated for the training set, based on which the next step is to extract the discriminative and visual features to represent their properties. Regarding different feature extraction strategies [19], the success of deep neural networks is revealed tremendously due to the large spread of ConvNets. One of the significant achievements of CNNs is the transfer learning that is deployed where less instances of dataset samples are available like the case under consideration. In this study, we have deployed three pretrained architectures such as AlexNet, GoogLeNet, and VGGNet of CNN for features extraction. The discriminative



**Fig. 5** Data augmentation applied on a meningioma image

visual features are extracted using two scenarios of transfer learning, i.e., fine-tune and freeze using each architecture, independently. Fine-tuning of transfer learning is used to increase the efficiency and productivity of a CNN network by making replacing the last layers of the pretrained network. In this scenario, ConvNet weights are initialized from the top of the pretrained network rather than replacing and retraining the whole architecture of CNN classifier. This scenario works by transferring the weights of the pretrained network from source dataset (ImageNet) to our target dataset (Figshare). The common practice is to truncate the softmax layer of the pretrained network and replace it with our new softmax layer that is relevant to our own problem. In this paper, using each architecture of CNN, last fully connected layer is replaced with the neurons of target dataset. In other words, 1000 classes of ImageNet are replaced with the three classes of brain tumor dataset—Figshare. In the second scenario, pretrained network layers are frozen and worked as fixed features. This scenario works by deriving the weights of the pretrained model from source dataset (ImageNet), and the desired features vector can be used from fully connected layers or from convolutional layers for training a linear classifier (SVM) on the data of target task (Figshare dataset). More detail is provided in Sect. 5 using each architecture with different experiments.

### 4.3 Classification and Detection of Brain Tumors

As we extract the useful and important visual features and patterns successfully using transfer learning techniques, then we perform classification and detection on target

**Table 2** Experimental results of fine-tune AlexNet on test of brain tumor dataset—Figshare

Parameters	Original dataset				Augmented dataset				
	Solver name	Adam	RMS prop	SGDM	SGDM	Adam	RMS prop	SGDM	SGDM
Batch size	10	10	7	10	10	10	7	<b>10</b>	
Initial learn rate	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$
Time (min:s)	49:46	9:19	38:51	13:46	40:20	39:10	50:40	<b>42:36</b>	
Validation frequency	300	300	500	300	300	300	500	<b>300</b>	
Max epochs	17	8	17	15	14	13	15	<b>7</b>	
Test accuracy	95.86	90.41	95.42	97.17	92.93	93.13	94.42	<b>97.39</b>	

The numbers in bold show the best results/performing-parameters

dataset. The classification is performed in two ways, i.e., log-softmax layers of pre-trained network using fine-tune features and linear classifier (SVM) using freeze layers. In first scenario, the learned visual features from each CNN architecture are fine-tuned to target dataset and brain tumor is classified using softmax layer by initializing the number of neurons to three classes. The parameters of fine-tuning method are not set by the network itself, and it is essential to set and optimize these parameters according to the results of training the MRI images in improving the performance. In our case, stochastic gradient descent momentum (SGDM) is trained with 0.9 momentum in each architecture. The value of batch size is set to 10 with the initial learn rate of  $1e^{-4}$  and the maximum epochs of 30. The number of best epochs is varied according to validation criteria with the validation frequency of 300 iterations. The highest accuracy of best network is achieved up to 98.69% using VGG16 on epoch 7. In the second scenario of classification, the freeze layers (different ConvNets or fully connected layers) of pretrained networks are passed to the linear classifier, i.e., support vector machine (SVM). Different feature vectors from each architecture are passed to the SVM for the brain tumor classification, independently. In case of AlexNet, Conv5 features are more discriminative, abstract, and more expressive; thus, the dimensions are unified through vectorization and fed to the SVM. The best network is attained with highest of 96.73% accuracy in this case. The GoogLeNet inception layers are explored and vectorized through activation process for unifying dimension. In this case, highest accuracy of 97% is achieved using inception-4e-output features. Similarly in case of VGG16, Fc7 features are more expressive thus passed to the SVM for classification. The highest accuracy in case is 89.76%. The detailed results' analysis is given in next section.

## 5 Comparative Analysis, Results, and Discussion

As discussed earlier, the experimental study of the system is carried out using the publicly free available brain tumor dataset— Figshare [4], comprising 3064 brain MRI slices with 708 slices for meningioma tumor, 1426 for glioma tumor, and 930 for pituitary tumor. We used 70% of dataset in training, 15% for validation, and 15%

**Table 3** Experimental results of freeze AlexNet on test set of brain tumor dataset—Figshare

Freeze layer	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
<i>Comparative analysis of accuracy using different features vector</i>				
Conv1	49.67	54.32	76.24	55.67
Conv2	56.21	49.19	75.27	56.97
Conv3	62.54	68.13	82.96	70.55
Conv4	81.92	79.33	90.16	80.84
Conv5	94.79	93.78	97.30	94.37
FC6	95.44	94.96	97.53	95.30
FC7	91.21	90.48	95.48	90.36

**Table 4** Experimental results of fine-tune GoogLeNet on brain tumor dataset— Figshare

Parameters	Original dataset				Augmented dataset			
	Solver name	Adam	RMS prop	SGDM	SGDM	Adam	RMS prop	SGDM
Batch size	10	10	7	10	10	10	7	<b>10</b>
Initial learn rate	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$
Time (min:s)	14:15	13:59	27:51	21:40	40:20	46:47	60:45	79:25
Validation frequency	300	300	500	300	300	300	500	500
Max epochs	5	6	10	19	6	9	7	7
Test accuracy	95.61	91.07	92.59	94.55	96.60	92.89	94.39	98.04

The number in bold shows the best results/performing-parameters

**Table 5** Experimental results of freeze GoogLeNet on test set of brain tumor dataset—Figshare

Freeze layer	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
<i>Comparative analysis of accuracy using different features vector</i>				
Inception 3a-Output	46.62	51.23	55.65	45.65
Inception 3b-Output	55.99	63.26	80.58	71.64
Inception 4a-Output	81.92	75.24	89.40	83.86
Inception 4b-Output	94.55	93.40	97.16	94.38
Inception 4c-Output	91.94	90.20	95.92	90.85
Inception 4d-Output	95.77	96.38	95.74	94.70
Inception 4e-Output	95.42	95.38	97.74	94.69
Inception 5a-Output	91.94	90.09	95.69	91.41
Inception 5b-Output	93.46	95.38	97.74	94.30

for testing. Three studies are carried out to evaluate the efficiency and performance of three CNN architectures named as AlexNet, GoogLeNet, and VGG16. In each study, two cases of transfer learning techniques such as fine-tune and freeze are explored in depth. Series of experiments conducted in both cases of each study independently and compared their results in Tables 2, 3, 4, 5, 6, 7. We explored different parameters

**Table 6** Experimental results of fine-tune VGG16 on test set of brain tumor dataset—Figshare

Parameters	Original dataset				Augmented dataset			
	Solver name	Adam	RMS prop	SGDM	SGDM	Adam	RMS prop	SGDM
Batch size	10	10	7	10	10	10	7	<b>10</b>
Initial learn rate	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$	$1e^{-4}$
Time (min:s)	19:4	15:1	10:45	17:52	40:20	39:10	50:40	89:30
Validation frequency	300	300	500	300	300	300	500	<b>300</b>
Max training epochs	7	7	4	8	14	13	15	<b>7</b>
Test accuracy	88.24	91.07	91.29	95.42	93.93	95.13	95.42	<b>98.69</b>

The numbers in bold show the best results/performing-parameters

**Table 7** Experimental results of freeze VGG16 on test set of brain tumor dataset—Figshare

Freeze layer	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
<i>Comparative analysis of accuracy using different features vector</i>				
Conv1-1	43.62	51.23	55.65	43.65
Conv2-1	50.90	67.70	66.99	50.00
Conv3-1	65.99	63.26	80.58	71.64
Conv4-1	75.90	73.31	86.95	75.36
Conv5-1	79.69	80.78	88.98	79.01
FC6	88.89	87.87	94.49	87.47
FC7	<b>89.76</b>	87.81	94.64	88.67

The number in bold shows the best results/performing-parameters

and reported the best network with the optimal parameters as shown in Tables 2, 4, 6. These tables show trained work of each model with different solvers and keep the default parameters, i.e., 30 epochs, 0.9 momentum, positive scalar for initial learning rate, 0.0001 L2 regularization, 0.1 learn rate drop factor, and 10 learn rate drop period. Each architecture of CNN in each study uses different numbers of parameters and features depending on the depth of the convolution layer and the fully connected network.

In first study, AlexNet is deployed to explore the transfer learning techniques as depicted in Fig. 2. In case of fine-tune approach of the pretrained AlexNet, we have evaluated number of parameters to get best performance of network and get highest accuracy. We have trained the network with three basic solvers such as sgdm, adam, and rms prop with different batch sizes and validation frequencies as reported in Table 2. The best network produces highest accuracy of 97.39% when trained with SGDM with the batch size of 10, validation frequency of 300. In case of freeze approach of the pretrained AlexNet, convolutional and fully connected layers are investigated for feature extraction as illustrated in Table 3. We have extracted the features from convolutional layers (Conv1, Conv2, Conv3, Conv4, and Conv5) that contain generic representation of source dataset's images (like edge, blob, etc.). We also extracted features from fully connected layers (FC6 and FC7) that represent a more detailed

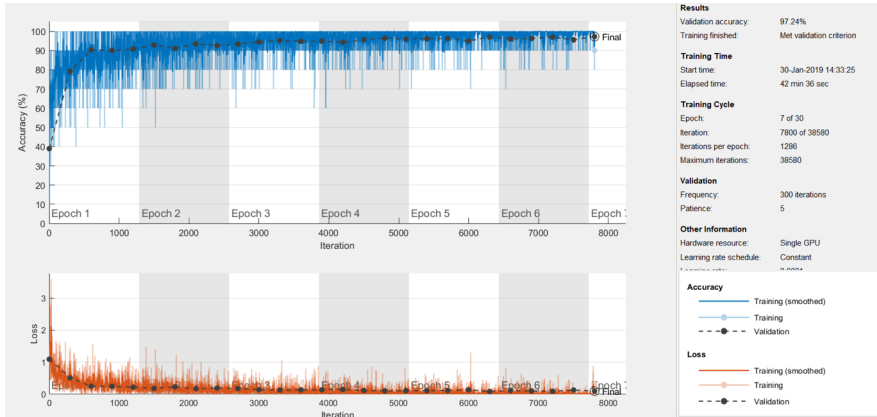
and specific features of the images of source dataset (ImageNet). The best results are reported using Conv5 of AlexNet having 95.46% accuracy.

In second study, GoogLeNet model is employed to investigate the transfer learning techniques as illustrated from Fig. 2. Eight series of experiments are conducted using fine-tuning technique of the pretrained GoogLeNet by varying the parameters as shown in Table 4. The highest identification rate of 98.04% is achieved when the pretrained GoogLeNet is trained with SGDM solver with the batch size of 10. Furthermore in case of freeze technique of the pretrained GoogLeNet, inception layers are investigated to extract the features and to explore which layers outperform the network as reported in Table 5. Nine inception layers are vectorized, and the extracted features from vectorization are passed to the SVM for classification. We have realized 95.77% accuracy using inception-4d-output layer.

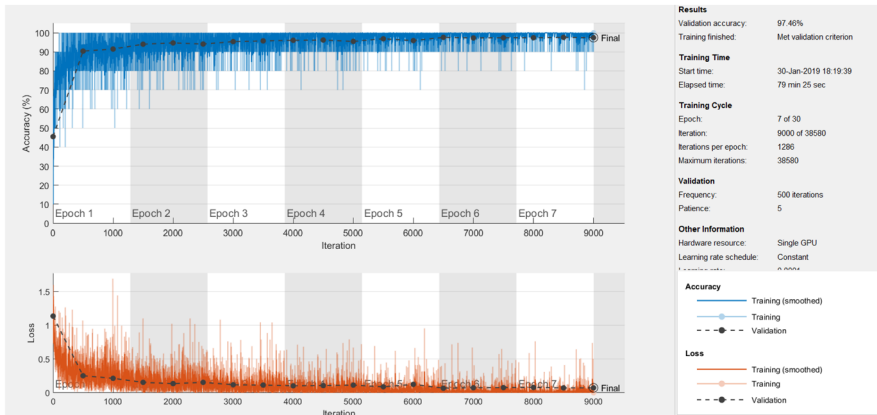
In third study, VGG16 architecture of CNN is implemented to explore the effectiveness of the transfer learning techniques as shown in Fig. 2 on the given data. Table 6 shows accuracy of different networks using different solvers and other parameters using fine-tuned of VGG15 classifier. The VGG16 fine-tuned-based best network has got on epoch 7. In this case, highest identification rate of 98.69% is achieved when the network is trained with SGDM solver, batch size of 10, and maximum epoch of 30. In case of freeze of the pretrained VGG16, convolutional and fully connected layers are investigated for feature extraction as illustrated in Table 7. The best with network with accuracy of 89.77% is attained using FC7 layer.

By analyzing and comparing the results of each architecture using fine-tuned technique as depicted in Fig. 6, we could observe that all architectures of convolutional neural networks outperform with the marginal difference. However, the highest accuracy is attained using VGG16 out of the three CNN architectures by generalizing the brain tumor images. The VGG16 network with fine-tuned approach achieved 98.69% accuracy on test set, 98.79% on validation set, and 99.02% on train set (Fig. 6c). The best network achieved when the validation criteria meet at epoch 7 at 8500 iterations with SGDM with batch size of 10 and training stopped automatically.

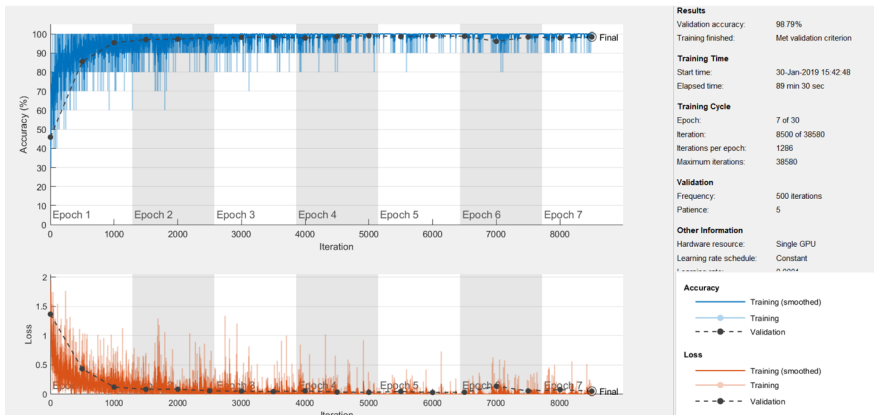
We have investigated and compared the evaluation of proposed system with existing brain tumor classification and detection-based systems on brain tumor dataset—Figshare, as given in Table 1. A significant comparison of our proposed system is possible with researches of Afshar et al. [3] and Abiwinanda et al. [2] as deliberated in Table 8. As they provided the insight work of brain tumors classification using deep learning model-based automated features, Afshar et al. [3] deployed model Capsule networks (CapsNets) for the classification of brain tumor. They vary the feature maps in the convolutional layer of CapsNet in order to increase accuracy. They achieved highest accuracy of 86.56% using 64 feature maps with one convolutional layer of CapsNet. However, Abiwinanda et al. [2] employed five different architectures of CNN for the classification of brain tumors and reported highest accuracy on architecture 2. The architecture 2 comprises of two convolutional layers, ReLU layer, and max-pool followed by 64 hidden neurons. They achieved 84.19% on validation set and 98.51% on training set, respectively. The literature divulges that transfer learning techniques using pretrained deep learning models had not been explored yet. Our work is the pioneer study to explore the transfer learning techniques using three renowned and pioneer pretrained CNN architectures. To increase the size of training dataset, we have



(a) AlexNet-Finetune



(b) GoogLeNet-Finetune



(c) VGG16-Finetune

Fig. 6 Progress graph of fine-tune approach of transfer learning using train set and validation set of brain tumor dataset—Figshare



**Table 8** Comparison with the closely related existing systems

References	Features	Model	Accuracy (%)
Afshar et al. [3]	Model based	CapsNet	86.56
Abiwinanda et al. [2]	Model based	CNN	84.19
Proposed systems	Fine-tune AlexNet	Log-based softmax	97.39
	Fine-tune GoogLeNet		98.04
	Fine-tune-VGG16		<b>98.69</b>
	Freeze AlexNet-Conv5		95.77
	Freeze GoogLeNet-inception-4e	SVM	95.44
	Freeze-VGG16-Conv5-1		89.79

The numbers in bold show the best results/performing-parameters

employed several techniques of data augmentation (rotation and flipping) with raw images. We have investigated the number of parameters using fine-tune of pretrained CNN networks. In reusing of freeze features of transfer learning, we have explored the features at each convolutional layer, inception layer and fully connected layer in three studies composed of architectures of AlexNet, GoogLeNet, and VVG16. Our proposed framework reveals the promising results on MRI images of Figshare using different techniques of data augmentation and as compared to the existing systems in the literature as shown in Table 8. We have attained the highest accuracy of 98.69% using fine-tuned VGG16 among fine-tune networks of AlexNet and GoogLeNet. In case of freeze technique of transfer learning, the highest accuracy 95.77% uses freeze Conv5–AlexNet layer as compared to other layers of AlexNet and also to all layers of architectures of GoogLeNet and VGG16. We have presented a pioneer study for brain tumors classification using brain tumor dataset—Figshare, and achieved the highest accuracy up to 98.69% using fine-tuned VGG16 approach.

## 6 Conclusion

In summary, the presented work is a pioneer study in the domain of brain tumor classification using transfer learning and deep CNN architectures. We applied transfer learning techniques using natural images of ImageNet dataset (source task) and classified the brain tumor type from glioma, meningioma, and pituitary using Figshare dataset (target task). We deployed three powerful deep CNN architectures (AlexNet, GoogLeNet, and VGGNet) on MRI slices of Figshare to identify the tumor type. To evaluate and explore the performance of deep networks, two studies of transfer learning (fine-tune and freeze) are conducted to extract the discriminative visual features and patterns from MRI slices. We have attained the highest accuracy of 98.69% using fine-tune VGG16 network among all experiments.

Although this paper explored three architectures of deep CNN and transfer learning approaches for brain tumor in the medical imaging domain, still much remains to be investigated. In future, we explore other essential powerful deep neural network's architectures for brain tumor classification with less time complexity.


## References

1. T.A. Abir, J.A. Siraji, E. Ahmed, B. Khulna, Analysis of a novel MRI based brain tumour classification using probabilistic neural network (PNN). *Int. J. Sci. Res. Sci. Eng. Technol.* **4**(8), 65–79 (2018)
2. N. Abiwinanda, M. Hanif, S.T. Hesaputra, A. Handayani, T.R. Mengko, Brain tumor classification using convolutional neural network. In *World Congress on Medical Physics and Biomedical Engineering*, pp. 183–189. Springer (2019)
3. P. Afshar, A. Mohammadi, K.N. Plataniotis, Brain tumor type classification via capsule networks. arXiv preprint: [arXiv:1802.10200](https://arxiv.org/abs/1802.10200) (2018)
4. J. Cheng, Brain tumor dataset. figshare. dataset. <https://doi.org/10.6084/m9.figshare.1512427.v5>. Accessed 30 May 2018
5. J. Cheng, W. Yang, M. Huang, W. Huang, J. Jiang, Y. Zhou, R. Yang, J. Zhao, Y. Feng, Q. Feng, Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation. *PLoS ONE* **11**(6), e0157112 (2016)
6. J. Cheng, W. Huang, R. Shuangliang Cao, W.Y. Yang, Z. Yun, Z. Wang, Q. Feng, Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLoS ONE* **10**(10), e0140381 (2015)
7. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, ImageNet: a large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009 CVPR 2009*, pp. 248–255. IEEE (2009)
8. M.R. Ismael, I. Abdel-Qader, Brain tumor classification via statistical features and back-propagation neural network. In *2018 IEEE International Conference on Electro/Information Technology (EIT)*, pp. 0252–0257. IEEE (2018)
9. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pp. 1097–1105 (2012)
10. A. Naseer, M. Rani, S. Naz, M.I. Razzak, M. Imran, G. Xu, Refining Parkinson’s neurological disorder identification through deep transfer learning. *Neural Comput. Appl.* (2019). <https://doi.org/10.1007/s00521-019-04069-0>
11. S. Naz, A.I. Umar, R. Ahmad, I. Siddiqi, S.B. Ahmed, M.I. Razzak, F. Shafait, Urdu Nastaliq recognition using convolutional–recursive deep learning. *Neurocomputing* **243**, 80–87 (2017)
12. I. Razzak, M. Imran, G. Xu, Efficient brain tumor segmentation with multiscalecancer statistics two-pathway-group conventional neural networks. *IEEE J. Biomed. Health Inf.* (2018). <https://doi.org/10.1109/JBHI.2018.2874033>
13. M.I. Razzak, Malarial parasite classification using recurrent neural network. *Int. J. Image Process.* **9**, 69 (2015)
14. M.I. Razzak, B. Alhaqbani, Automatic detection of malarial parasite using microscopic blood images. *J. Med. Imaging Health Inform.* **5**(3), 591–598 (2015)
15. M.I. Razzak, M. Imran, G. Xu, Big data analytics for preventive medicine. *Neural Comput. Appl.* (2019). <https://doi.org/10.1007/s00521-019-04095-y>
16. M.I. Razzak, S. Naz, Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 801–807. IEEE (2017)
17. M.I. Razzak, S. Naz, A. Zaib, Deep learning for medical image processing: overview, challenges and the future. In *Classification in BioApps*, pp. 323–350. Springer (2018)
18. A. Rehman, S. Naz, M.I. Razzak, H.A. Ibrahim, Automatic visual features for writer identification: a deep learning approach. *IEEE Access* **7**, 17149–17157 (2019)
19. A. Rehman, S. Naz, M.I. Razzak, Writer identification using machine learning approaches: a comprehensive review. *Multimed. Tools Appl.* **78**(8), 10889–10931 (2019)
20. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**(3), 211–252 (2015)
21. S.H. Shirazi, A.I. Umar, S. Naz, M.I. Razzak, Efficient leukocyte segmentation and recognition in peripheral blood image. *Technol. Health Care* **24**(3), 335–347 (2016)
22. R. Siegel, C.R. Miller, A. Jamal, Cancer statistics, 2017. *CA Cancer J. Clin.* **67**(1), 7–30 (2017)
23. R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2015. *CA Cancer J. Clin.* **65**(1), 5–29 (2015)
24. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint: [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)

25. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9 (2015)
26. W. Widhiarso, Y. Yohannes, C. Prakarsah, Brain tumor classification using gray level co-occurrence matrix and convolutional neural network. *IJEIS (Indones. J. Electron. Instrum. Syst.)* **8**(2), 179–190 (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Arshia Rehman<sup>1</sup> · Saeeda Naz<sup>1</sup> · Muhammad Imran Razzak<sup>2</sup>  ·  
Faiza Akram<sup>3</sup> · Muhammad Imran<sup>4</sup>

✉ Muhammad Imran Razzak  
imran.razzak@ieee.org  
Arshia Rehman  
arshiar29@gmail.com  
Saeeda Naz  
saeedanaz292@gmail.com

- <sup>1</sup> Computer Science Department, GGPGC No. 1, Abbottabad, Khyber Pakhtunkhwa, Pakistan
- <sup>2</sup> University of Technology Sydney, Ultimo, Australia
- <sup>3</sup> Radiology Department, Ayub Medical and Teaching Institute, Abbottabad, Pakistan
- <sup>4</sup> College of Applied Computer Science, King Saud University, Riyadh, Saudi Arabia