Received: 14 August 2020 | Revised: 21 November 2020 | Accepted: 4 January 2021

DOI: 10.1002/ima.22546

RESEARCH ARTICLE

WILEY

Transfer learning networks with skip connections for classification of brain tumors

Saleh Alaraimi[1] | **Kenneth E. Okedu**[1] | **Hugo Tianfield**[2] | **Richard Holden**[2] | **Omair Uthmani**[2]

[1]Department of Electrical and Computer Engineering, College of Engineering, National University of Science and Technology, Muscat, Oman

[2]School of Computing, Engineering and Built Environment, Glasgow Caledonian University, Glasgow, UK

Correspondence
Saleh Alaraimi and Kenneth E. Okedu, Department of Electrical and Computer Engineering, College of Engineering, National University of Science and Technology, Muscat, Oman.
Email: salehalaraimi@nu.edu.om (S. A.) and okedukenneth@nu.edu.om (K. E. O.)

Abstract

This article presents a transfer learning model via convolutional neural networks (CNNs) with skip connection topology, to avoid the vanishing gradient and time complexity, which are usually common in transfer learning networks. Three pretrained CNN architectures, namely AlexNet, VGG16 and GoogLeNet are employed to equip with skip connections. The transfer learning is implemented through fine-tuning and freezing the CNN architectures with skip connections based on magnetic resonance imaging (MRI) slices of brain tumor dataset. Furthermore, in the preprocessing, a frequency-domain information enhancement technique is employed for better image clarity. Performance evaluation is conducted on the transfer learning networks with skip connections to obtain improved accuracy in brain MRI classifications.

KEYWORDS

AlexNet, convolutional neural network (CNN), deep learning, GoogLeNet, transfer learning, VGG

1 | INTRODUCTION

Brain tumors are the pathological complications caused by uncontrolled cell division, and abnormal growth of tissues. The common types of brain tumor are glioma, meningioma and pituitary. The size of the tumor is quite large in case of glioma, and can lead to short life expectancy.[1] Abnormal growth of brain cell in pituitary gland results in development of pituitary gland tumor that causes abnormal production of pituitary hormones. Due to their intrinsic nature, meningioma, glioma and pituitary gland tumors can occur anywhere in brain tissues. However, these tumors have different shapes, sizes and contrasts.[2]

Early detection of brain tumor is necessary as delay could lead to death. The detection of location and size of complex cases of brain tumor is a complex task, as the whole treatment plan is based on diagnosis made in the early stages. Thus, locating the exact region of tumor tissues, the extent to which they are spread in nodules, and the grading and type of cancer is necessary. This would help in assessment for the development of treatment plan as well as for the tracking of progression of certain diseases.[3-5]

Deep learning (DL) is considered as one of the most widely used techniques for diagnosis of Alzheimer's, normal and abnormal brain tumor classifications, and stroke lesion segmentation. Thus, nowadays, brain image analysis by radiologists and oncologists often consider DL. In the literature, DL models mostly employ convolutional neural network (CNN) for classification and segmentation of medical images, as it tends to learn the spatial relationship in a hierarchical manner between the pixels.[6] CNN learns how to segment images into many layers; it builds a hierarchy of feature maps by convolving the images through learned filters. CNN conducts minimal processing and extracts the medical features directly from the pixel images.[7-9]

The effectiveness of DL models in brain tumor imaging and processing is due to the potential to grade and categorize the images.[10] The high performance of CNN enables the radiologist to precisely identify the location, shape, and grade of tumors and assists in the development of effective personalized treatment plan.[11]

DL models for neuroimaging involve certain architectures of CNN which are highly efficient by segmenting the medical image into 1000 images for higher resolution and contrast.[12-18] The CNN models mostly utilized for neuroimaging as well as imaging of biopsy samples for other cancer types include AlexNet, VGG16 and VGG19.[19-23] DL models are efficient and robust as they precisely address the cancer type, the tissues and neural networks to which tumor has spread, and supplement the data for development of effective treatment plan.[24-29]

In Reference 30, a DL approach used transfer learning (TL) network to classify brain tumors, but skip connection topology was not considered. Reference 31 classified brain tumor images using CNN pretrained networks. Pretrained CNN architectures were used as feature extractor for image splicing detection in Reference 32. Although TL networks were able to achieve high accuracy in the classifications of brain tumor, the data goes through many layers, which causes vanishing gradient and time complexity.

In this article, a TL model is presented via CNN with skip connection topology, to avoid the vanishing gradient and time complexity, which are common in TL networks.

In particular, three pretrained CNN architectures, namely GoogLeNet, AlexNet and VGG16 are employed to equip with skip connections.

Also, the CNN architectures with skip connections have slight changes in the input datasets and in the number of kernels for the output networks. In addition, a frequency-domain information enhancement technique is employed for better image clarity in the preprocessing.

Furthermore, TL is implemented through fine-tuning and freezing the CNN architectures with skip connections based on magnetic resonance imaging (MRI) slices of brain tumor dataset Figshare. Performance evaluation is conducted on the fine-tuned CNN architectures with skip connections in brain MRI classifications for common tumors such as meningioma, glioma and pituitary.

2 | TL NETWORKS WITH SKIP CONNECTIONS

2.1 | TL model via CNN with skip connection topology

The CNN architecture, as shown in Figure 1, includes a number of layers having its own functions. In order to train parameters in CNN, the data has to undergo each layer serially. The convolutional layer is known as the first layer in CNN. It uses simple filtering to extract the features of images. The pixels are analyzed based on similarities and relationships that exist in a smaller number of samples.

The main function of pooling layer in CNN is to decrease the samples required, by keeping the amount of important information, in order to increase the speed of its operations. There are three different types of pooling layers, namely Max, Average and Sum pooling. They play an important role in controlling the overfitting of datasets.

The fully connected layer in CNN is linking all the activation functions that are present in the previous layers, by bringing appropriate classification. The main task is to retrieve the required information present in the data, that is, image in this article. The disadvantage of this layer is the massive requirement of parameters than the previous layers to produce a single output. A fixed number of inputs are sufficient to give a similar output. The aspect ratio input data can be maintained by converting the fully connected layer to convolutional layers of 1×1 kernel and $1 \times 1 \times 1$ in case of 3D images. For better accuracy, the layer uses nonlinear transformation in order to extract features.

The training of deep neural networks is difficult to proceed due to vanishing or exploding gradient issues. Reaching control of deep network by using normalization only may expose degradation problem, while adding a number of layers may increase training errors rate. In 2015,

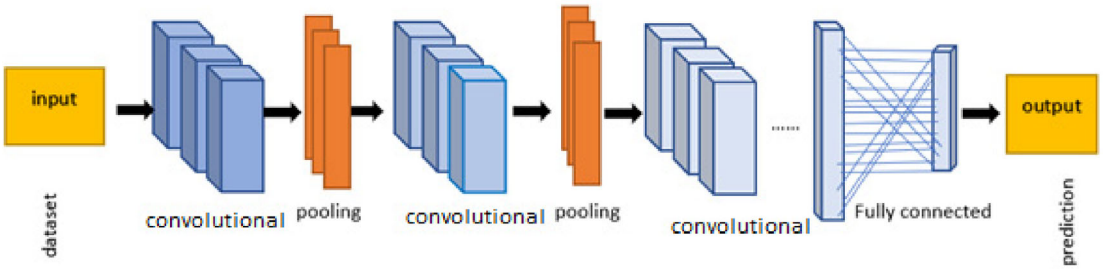


FIGURE 1 General architecture of convolutional neural networks (CNNs)

Reference 33 introduced a model using Resnet blocks that connects the output of one layer with the input of an earlier layer. This solution is referred to as a skip connection.

In this article, we employ three pretrained CNN architectures, namely AlexNet, VGG16 and GoogLeNet based on the Resnet blocks in Reference 33. In our TL model, the output of each group of layers is connected to the input data with a skip connection, which is a parameter that is also to be learned during the training process. Therefore, in our TL model, the skip connection topology connects input image to groups of layers through a single skip connection in AlexNet and VGG16 networks, and through multiple skip connections in GoogLeNet, at frequent periods. The TL model is implemented with brain tumor dataset for three different classes of brain tumors, that is, meningioma, glioma and pituitary in this article.

Skip connection topology is a nonlinear process that connects neural network layers and skips one or more layers, as shown in Figure 2. The mechanism of skip connections is to skip some layers in neural network and feed the output of one layer, as the input of the next layers. The construction of $F(x) + X$ can be achieved by feedforward neural networks, with "skip connections." The skip connection should learn at least identity mappings from one point in the network, in order to forward them to another point, thus, making the network learn extra F(x). The process will continue until $F(x)$ becomes 0. Consequently, this will help the network to easily learn a mapping closer to 0 than identity mapping.

In this article, three pretrained CNN architectures, that is, AlexNet, VGG, GoogLeNet, are equipped with skip connections, in order to improve their performances. The CNN with skip or shot connection employed for implementing the TL model is shown in Figure 3. The skip or shot connection works by jumping or transferring data from one convolution layer to another. It should be noted that, when implementing TL, a frequency-domain information enhancement technique is employed for better image clarity in the preprocessing for all the CNN architectures with skip connections. In this work, the open source available brain tumor dataset Figshare was used to evaluate the performance of the proposed CNN. The dataset was developed by Cheng in 2017.[34]

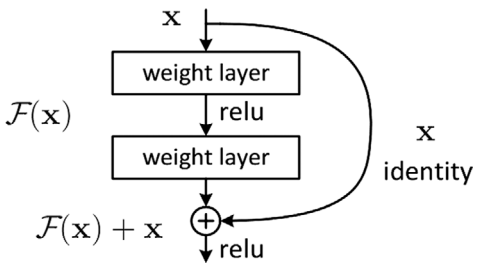


FIGURE 2 Skip connection topology

2.2 | CNN architectures with skip connections

2.2.1 | AlexNet with skip connection

AlexNet is the winner of ISLVRC 2012. It is the first deep neural network that classifies thousands of objects in different classes. The network comprises of convolutional, rectified linear unit, and max-pooling layers. The network consists of five convolutional layers and three fully connected layers. The input size of the image to be fed to this network is $227 \times 227 \times 3$.

Before feeding the input image to AlexNet, images are preprocessed using Z-score image normalization that depends on the mean and SD values of the image pixels. We use the contrast enhancement in addition to the normalization technique to train the images effectively. Contrast enhancement in images refers to the enhancement of the intensity of image pixels, in order to make them more informative.

In this article, the contrast enhancement is performed by contrast limited adaptive histogram equalization. Figure 4 shows the architectural model of AlexNet with skip connection, employed in this article. Generally, the normal AlexNet architecture is pretrained for 1000 classes. However, the AlexNet with skip connection employed in this article is changed to three classes for cancer detection, based on the dataset classes or requirements for brain tumor.

From Figure 4, the AlexNet with skip connection is different from the normal architecture because of the skip connection, dataset classes, changes in the frequency domain, and the number of kernels used in the network.

Table 1 lists the summary of the convolutional layers in the AlexNet architecture, and Table 2 shows the number of layers in AlexNet network. The first convolutional layer in the AlexNet architecture has 96 filters with a window size of 11x11. The second convolutional layer consists of 256 filters with a window size of 5×5. Likewise, the third, fourth, and fifth layers consist of 384, 384, and 256 filters, respectively. The rectified linear unit (relu) activation function is used between the layers as an activation function that takes the negative values as zero. Therefore, the nonactivated neurons become zero which helps to improve the feature extraction for better classification.

For TL, we use the AlexNet which is a pretrained network, and freeze the layers for fine turning. Thus, the final fully connected layer is set as three as there are three different classes of brain tumors, that is, glioma, meningioma, and pituitary.

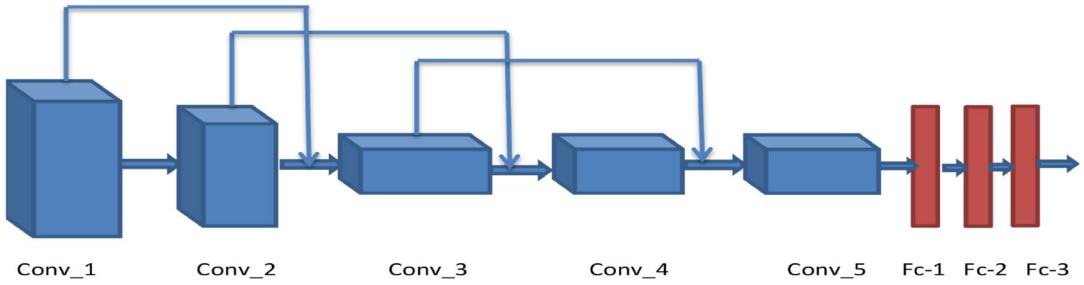
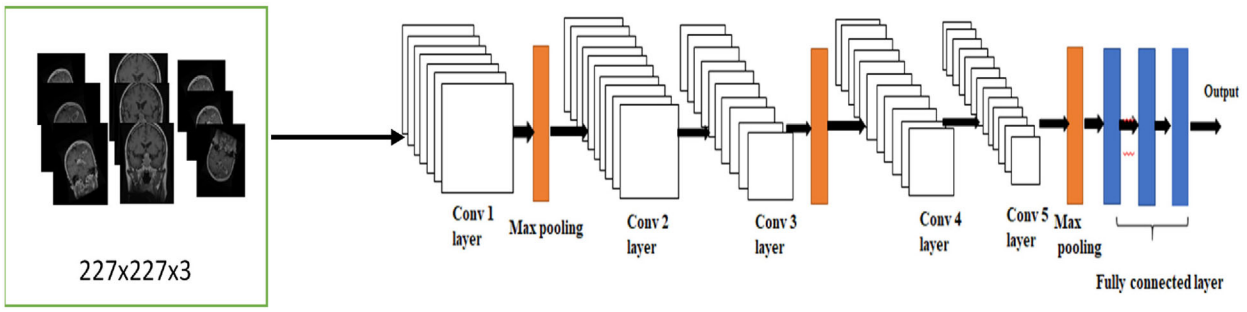
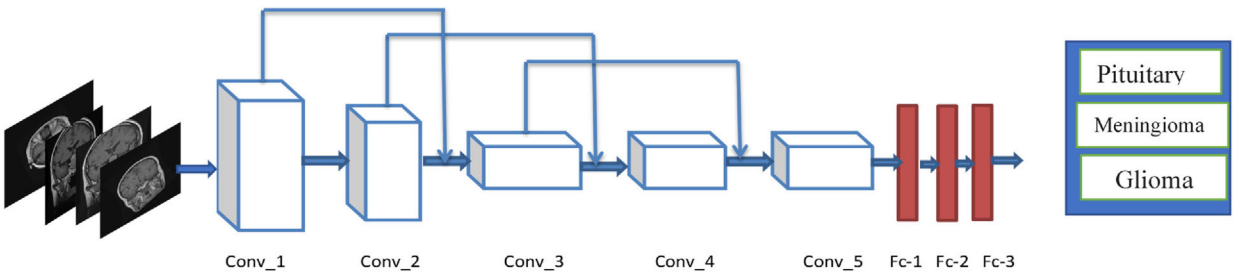


FIGURE 3 Convolutional neural networks (CNNs) with skip or shot connection for transfer learning



(A) AlexNet architecture



(B) AlexNet with skip connection

FIGURE 4 Architectural model of AlexNet with skip connection

TABLE 1 Summary of convolutional layers in AlexNet

Convolutional layer number	Number of filters	Filter size	Stride	Padding
Convolutional layer-1	96	3 × 3	4	0
Convolutional layer-2	256	3 × 3	1	2
Convolutional layer-3	384	3 × 3	1	1
Convolutional layer-4	384	3 × 3	1	1
Convolutional layer-5	256	3 × 3	1	1

2.2.2 | VGG16 with skip connection

VGG16 is a 16-layer CNN architecture. Similar to AlexNet, it is made up of convolutional, max-pooling Relu layers.[35] The number of filters changes from 64 to 512, and there are three fully connected layers. All the convolutional layers have a filter window size of 3 × 3. VGG16 has fully connected layers, the layer parameters and the number

TABLE 2 Summary of layers in AlexNet

Layer type	Feature map	Kernel size + stride	Activation
Input image	1	—	—
Convolution	96	3 × 3 _ stride 4	relu
Max pooling	60	3 × 3 _ stride 2	relu
Convolution	256	3 × 3 _ stride 1	relu
Max pooling	256	3 × 3 _ stride 2	relu
Convolution	384	3 × 3_ stride 1	relu
Convolution	256	3 × 3 _ stride 1	relu
Max pooling	256	3 × 3 _ stride 2	relu
FC	—	—	relu
Softmax	—	—	Softmax

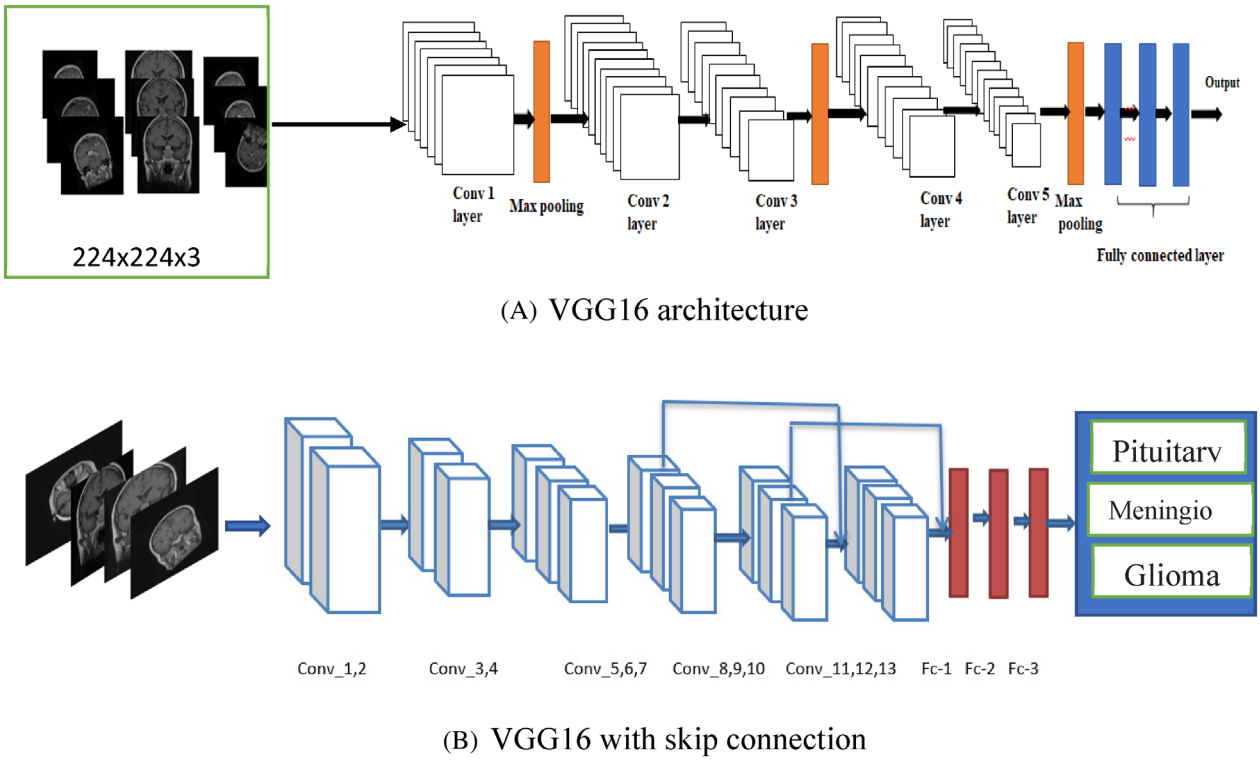


FIGURE 5 Architectural model of VGG16 with skip connection

of layers is higher compared to AlexNet. In the VGG16 architecture, the first, second convolutional layers have the same filter size of 64; the third and fourth convolutional layers have 128; the fifth, sixth and seventh convolutional layers have 256 filters. Likewise, the 8th, 9th, 10th, 11th, 12th, and 13th layers have 512 filters. At the end, it consists of three fully connected layers.

The computational complexity of this VGG16 architecture is higher compared to other networks, since it consists of two stages 512 filters each. To perform fine-tuning for TL, the final fully connected layer is tuned to produce three output classes instead of 1000 classes. The activation function used here is the rectified linear unit. Each convolutional layer extracts suitable features and it is a pyramid architecture that helps to improve the classification accuracy.

Figure 5 shows the VGG16 with skip connection topology employed in this article, with an input dataset of 224 × 224, and kernel output for the datasets under consideration. Table 3 shows the number of layers in VGG16 network.

TABLE 3 Summary of layers in VGG16

Layer type	Feature map	Kernel size + stride	Activation
Input image	1	—	—
Convolution	64	3 × 3 _ stride 4	relu
Convolution	64	3 × 3 _ stride 4	relu
Max pooling	64	3 × 3 _ stride 2	relu
Convolution	128	3 × 3 _ stride 4	relu
Convolution	128	3 × 3 _ stride 4	relu
Max pooling	256	3 × 3 _ stride 2	relu
Convolution	256	3 × 3_ stride 1	relu
Convolution	256	3 × 3 _ stride 1	relu
Convolution	256	3 × 3 _ stride 1	relu
Max pooling	256	3 × 3 _ stride 2	relu
Convolution	512	3 × 3 _ stride 1	relu
Convolution	512	3 × 3 _ stride 1	relu
Convolution	512	3 × 3 _ stride 2	relu
Max pooling	512	3 × 3 _ stride 1	relu
Convolution	512	3 × 3 _ stride 1	relu
Convolution	512	3 × 3 _ stride 2	relu
Convolution	512		
Max pooling	256	3 × 3 _ stride 2	relu
FC	—	—	relu
Softmax	—	—	Softmax

2.2.3 | GoogLeNet with skip connection

GoogLeNet with a shot connection is a network architecture in which the previous layer connections will be added with the current layers to avoid the vanishing gradients. Here, the previous layer features are added with the present layer features such that there is no feature loss. Also, the present layer feature values will improve the learning of effective features.

GoogLeNet[2,36] has nine inception modules. Furthermore, each inception module contains one max-pooling layer and six convolutional layers. From this, 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.

GoogLeNet has 6.8 million parameters, because of additional feature learning. It comprises of 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.

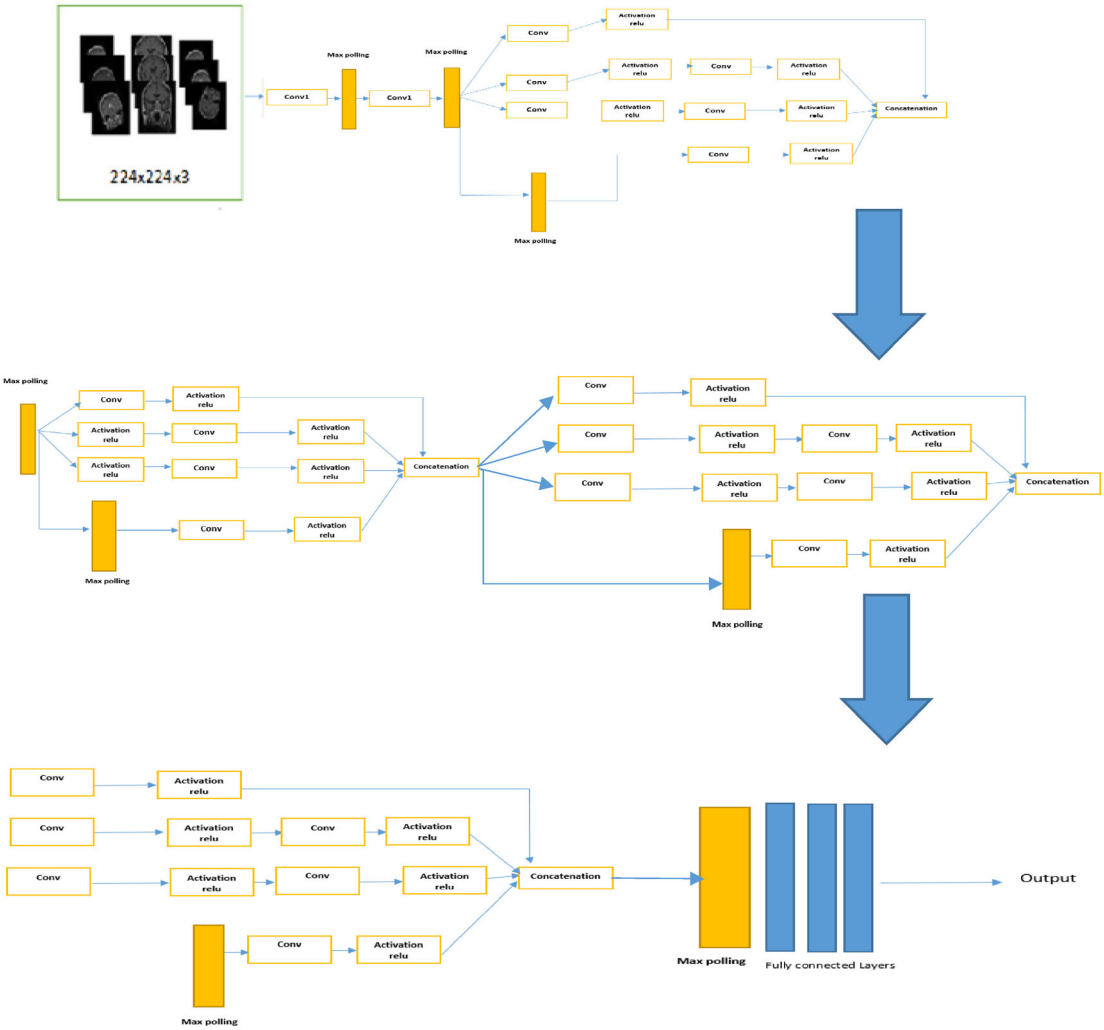
The performance of GoogLeNet is more accurate than other architectures, that is, VGG and AlexNet. Moreover, it is more precise than AlexNet on the original ILSVRC dataset.

The GoogLeNet with skip connection employed in this article is given in Figure 6. Table 4 lists the summary of layers in the GoogLeNet.

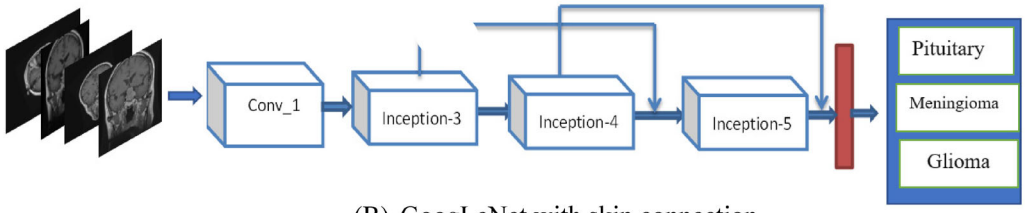
2.3 | Training strategies of CNNs with skip connections

In this section, the training methods used in the TL networks are discussed. To increase the classification performance, all the three pretrained networks are fine-tuned for TL. In all the pretrained networks, the fully connected layers are fixed as three. Learning rate and other learning parameters are also changed to achieve better results. In all three pretrained networks, activation has been analyzed, but the GoogLeNet activation function is more visible than the other two networks. Three different patch sizes, namely 16, 32, and 64 are analyzed, and it is finally fixed with 32 in this article.

The Adam optimization algorithm is used here for training. It helps to reduce the gradient value as quickly as possible. It is the combination of two optimization



(A) GoogLeNet architecture



(B) GoogLeNet with skip connection

FIGURE 6 Architectural model of GoogLeNet with skip connection

algorithms, such as gradient descent with momentum and RMSprop. This combined optimizer effectively works for the fine tuning of hyperparameters. Also, its low memory requirements help to improve the training accuracy with fewer epochs.

The loss function is an important criterion in training to recognize the class labels. This loss function needs to be minimized. The loss function used in this article is cross-entropy loss function,[37] which is defined as.

$$L(w) = \sum_{i=1}^{N}\sum_{c=1}^{4}\{-y_{ic}\log f_c(x_i)\} + \in\|w\|_2^2 \qquad (1)$$

$f_c(x_i)$ – predicted probability of class c for image x

The training algorithm is given as Algorithm 1.

The performance of the deep network can be improved by increasing the size of the training data. Image augmentation can be carried out to increase the

TABLE 4 Summary of layers in GoogLeNet

Layer type	Feature map	Kernel size + stride	Activation
Input image	1	—	—
Convolution	64	3 × 3 _ stride 4	relu
Max pooling	64	3 × 3 _ stride 4	relu
Convolution	60	3 × 3 _ stride 2	relu
Max pooling	128	3 × 3 _ stride 4	relu
Inception (3a)	128	3 × 3 _ stride 4	relu
Inception (3b)	256	3 × 3 _ stride 2	relu
Max pooling	256	3 × 3_ stride 1	relu
Inception (4a)	256	3 × 3 _ stride 1	relu
Inception (4b)	256	3 × 3 _ stride 1	relu
Inception (4c)	256	3 × 3 _ stride 2	relu
Inception (4d)	512	3 × 3 _ stride 1	relu
Inception (4e)	512	3 × 3 _ stride 1	relu
Max pooling	512	3 × 3 _ stride 1	relu
Inception (5a)	512	3 × 3 _ stride 1	relu
Inception (5b)	512	3 × 3 _ stride 2	relu
Avg pooling	256	3 × 3 _ stride 2	relu
FC	—	—	relu
Softmax	—	—	Softmax

Algorithm 1

Training algorithm of CNN with skip connection for TL

1. **Input**: Input data samples for Training Datasets (*TDs*) with the number of images *nIm*, class labels *Cl* (glioma, meningioma, and pituitary), validation datasets (*VDs*), learning rate *Ra*,
2. **Output**: Feature map, performance metrics such as accuracy, precision, recall, etc.
3. Reduce the cross entropy, train the network with Adam optimizer.
4. Loss is computed on the training datasets *TDs* and the validation datasets *VDs* by backpropagation algorithm based on the class labels.
5. Initialize the learning rate *Ra*
6. *For* $z = 1$ to the number of images *nIm*
   a. Take the feature map for the final convolutional layer
   b. Find softmax values of the output layer
7. Return Loss *Lr*, Accuracy *Ac*, feature map

size of the training data. If there is not much training data, image augmentation can be obtained by resizing, rotating, and flipping operations. However, while we increase the training data, the data memory requirement on the computer and the computation time will be high.

The selection of patch size is very important to get the optimal performance of the network. In this brain tumor detection problem, we set the minimum patch size as 32, which will help to improve the detection rate. In some scenarios, increasing the patch size leads to a high misclassification rate because most of the feature maps will be redundant, which decreases the optimal performance of the network.

Generally, it is not enough that the entire datasets are given to the network for only one time, because the network needs to calculate the loss for different training samples. Therefore, the maximum epoch must be higher than one. The number of epochs increases the weight value of the network. The epoch value should be increased until you get the training accuracy higher or 100. Here, we set the maximum epoch as 200.

For the preprocessing, two different methods are used here. Both are in the frequency domain. The two methods are histogram equalization and normalization. Stationary wavelet transformation is used for frequency-domain

conversion. These two methods are carried out in the frequency domain only. Four subbands are obtained in the frequency domain. Initially, the input image is converted from the spatial domain to the frequency domain by wavelet transformation, then the equalization is carried out on all the subbands and inverted to the original domain. After the equalization, the same frequency-domain process is carried out for Z score normalization in all the subbands.

Figure 7 shows frequency-domain information enhancement on the image improvement in the preprocessing, which gives higher resolution to fit into the training process, while Figure 8 is the frequency information enhancement process.

Equalization and normalization give contrast-enhanced images for better training. The Z score normalization is defined by.[38,39]

$$Zsc = \frac{X - \text{mean}}{std} \tag{2}$$

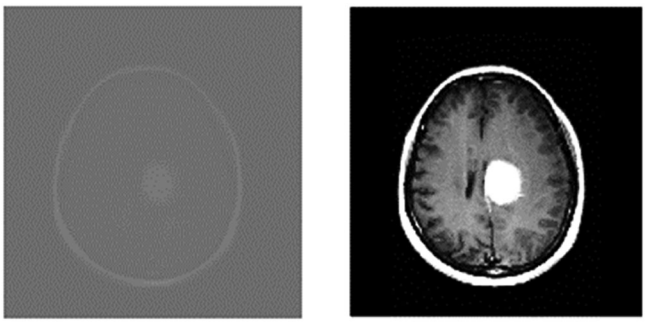
Zsc – Normalized outcome, X – input data

Once the image data augmentation is completed, the network is fed with this huge amount of training data. This increases the amount of training data and leads to improve classification performance.

The convolution layers are the second layer of the network after the input layer, which helps to extract special features such as low-level, medium, and high-level features. Each convolutional layer extracts different levels of features which helps to train the network in a better way. The three pretrained CNN architectures, AlexNet, VGG16, and GoogLeNet, use the pretrained weights that are transferred from one layer to another layer. For fine-tuning, the final fully connected layer of the three pretrained networks is replaced by three classes instead of 1000 classes.

The pretrained network weights are transferred to the target dataset and there is no need for retraining the network. The significance of TL is that a smaller number of datasets are needed under this circumstance because the weights are transferred. In this article, the softmax is used as the classification layer.

Image data augmentation is needed for improving the training accuracy, by increasing the number of training images. In this article, we have used five different types of image processing techniques which are image rotation, cropping, flipping, scaling, and translation. In total, there are 100 gliomas, 120 meningioma, and 150 pituitaries. Preprocessing is carried out on these three classes and augmented to get a total of 1500 images. Figure 9 shows the image data augmentation. In Figure 9A, the data are rotated 10°, while Figure 9B is normal orientation, Figure 9C, is flipped from upper orientation to laser and Figure 9D is rotated by 45°.



(A) Original image (B) Contrast enhanced image

FIGURE 7 Result of contrast enhancement

3 | PERFORMANCE EVALUATION

In order to implement our TL model, MATLAB 2020a is employed. The MATLAB tool is used with single graphical processing unit having multiple processers of 2.6 GHz and 16GB RAM, 1TB of HDD. This section presents a set of the experiments that are carried out using brain tumor dataset Figshare for brain tumor classification and detection. We evaluate the performance of three pretrained CNN architectures, AlexNet, VGG16 and GoogLeNet, using augmented image slices of brain tumor dataset. These pretrained CNN architectures are used to deploy the TL model, in order to extract the rich features. Then, the confusion

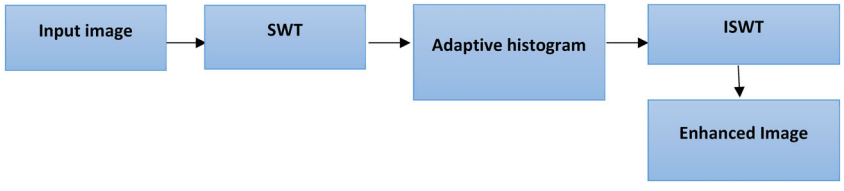


FIGURE 8 Frequency-domain information enhancement process in the preprocessing: ISWT, inverse SWT; SWT, stationary wavelet transformation

matrices are discussed for evaluation performance of the CNN architectures with skip connections.

In the experiments, we use the open-source brain tumor dataset Figshare to evaluate the performance of the CNN architectures with skip connections. The dataset was developed by Cheng in 2017. The number of brain MRI is 3064 collected from 233 patients, with three different types of brain tumors, that is, meningioma, glioma and pituitary. The images are distributed as follows: 930 images for pituitary, 708 images for meningioma and 1426 images for glioma. The main format is ".mat," each file including structure with a patient ID, unique label that denotes the type of brain tumor, 512×512 image data in unit 16 format, a vector containing a tumor boundary, with discrete dot coordinates, and a ground truth in a binary mask image.

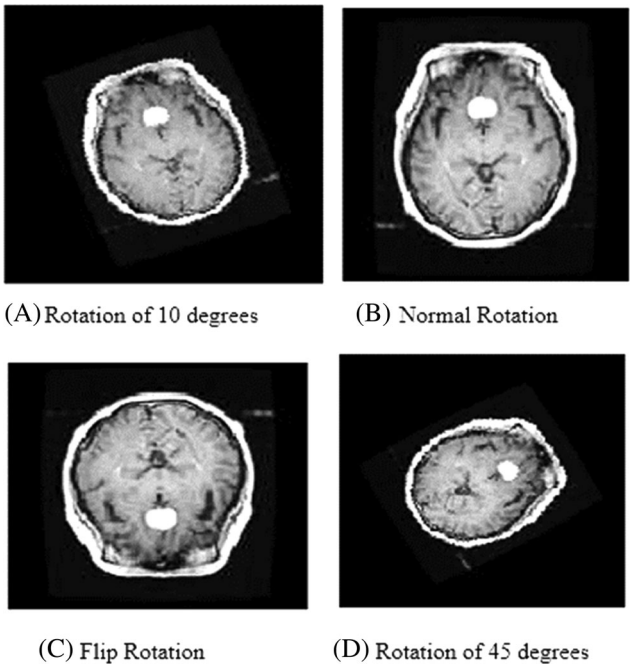


FIGURE 9 Augmented image data in different forms

MRI is an imaging technique employed in medical sciences to analyze both anatomy and physiological processes of some parts or the entire human body. The scanners associated with this kind of imaging technique have sophisticated radio waves, magnetic fields and gradients, which enable capturing of images of human organs.

The CNN architecture takes MRI images as an input unit in our experiments; hence, we only use the image data from the ".mat" files after extraction and resizing, to fit the model requirements. The CNN architecture performs automatic feature extraction using convolutional layers without human intervention. CNN generally performs better in comparison to the classic image processing methods.

The experimental setup shown in Figure 10 contains both the training and the testing processes. The enhancement and normalization are carried out in both the training and the testing. Then, feature extraction is carried out using the pretrained network, namely AlexNet, VGG16, GoogLeNet architectures. Finally, the cross-entropy loss function is computed to differentiate the input classes with the help of training data labels. This process is repeated for testing images as well, to identify the classes, whether it is glioma, meningioma, or pituitary.

Various experiments are conducted to evaluate the performance of the system. In the first experimental scenario, the images are randomly split as 70% for training and 30% for testing. In the other experimental scenario, 80% of the images are used for training and 20% for testing. Both experimental scenarios provide promising results, but the second experimental scenario provides results which are slightly better than the first experimental scenario.

In this article, we also evaluate different optimizers, namely stochastic gradient descent with momentum, Adam, and RMSprops. GoogLeNet using Adam optimizer provides better results compared to other networks. The experiments are conducted on both original and augmented data independently. After freezing some of the

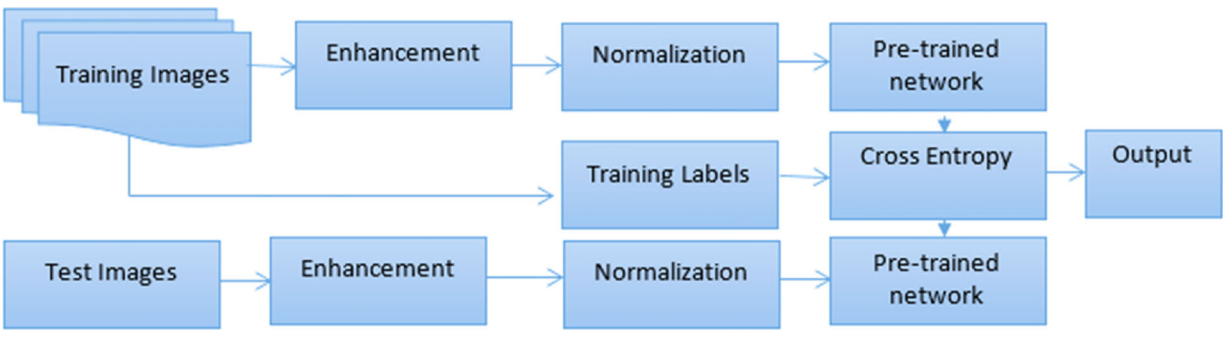


FIGURE 10 Training and testing processes in experimental setup

layers in all three pretrained networks independently, the impact of the remaining layers is also analyzed.

The AlexNet, VGG16, and GoogLNet are experimentally evaluated based on the performance metrics such as accuracy, sensitivity, specificity and F-score. The performance metrics are stated below.

True positive (Tp) = Number of identified positive samples that are truly positive.
False positive (Fp) = Number of identified positive samples that actually are negative.
True negative (Tn) = Number of identified negative samples that are truly negative.
False negative (Fn) = Number of identified negative samples that actually are positive.

Accuracy is the ability to discriminate the different classes correctly. To estimate the accuracy of a network, the proportion of true positive and true negative among all samples is computed, that is,

$$\text{Accuracy} = (Tp + Tn)/(Tp + Tn + Fp + Fn) \quad (3)$$

Sensitivity is the ability to see the unique independent class samples properly. To estimate it, we should compute the proportion of true positive in unique independent samples, that is,

$$\text{Sensitivity} = Tp/(Tp + Fn) \quad (4)$$

Specificity is the ability to identify the negative samples properly. To estimate it, we must compute the proportion of true negative in normal cases, that is,

$$\text{Specificity} = Tn/(Tn + Fp) \quad (5)$$

3.1 | Performance of AlexNet with skip connection

We have taken individual alternate layers from AlexNet and obtained the performance metrics as presented in Table 5. From Table 5, we can see that the increment of convolutional layers leads to accuracy improvements, that is, Conv_1 gives 50% accuracy but conv_3, conv_4 give 63.5% and 95%, respectively, which are better than conv_1; so the features learned from the conv_5 layer are better than all other layers. The overlapping feature maps are eliminated by the dropout layers, so it provides better performance.

The performances of the intermediate layers are analyzed after training the pretrained AlexNet architecture (trained on million images) on the brain tumor dataset. It is carried out by extracting and utilizing the features learnt from the individual intermediate layers for brain tumor classification. The features learnt from the intermediate convolutional layers such as conv_1, conv_3, conv_5 and two fully connected layers are separately used, by freezing the remaining layers for brain tumor classification, using the machine learning classifier. Here, the support vector machine classifier is used to extract the image features. It is understood that taking features from conv_3 layer includes the mid-level features learnt from the lower levels after performing successive convolution, that is, conv_1 to conv_3.

From Table 5, after freezing some of the layers in the AlexNet, the performance of layers such as, conv_1 to conv_3, conv_1 to conv_ 5, conv_1 to FC_6 and conv_1 to FC_7 is analyzed. It is observed that, when the number of convolutions is increased, the results of the fully connected layer FC_6 have higher performance metrics, that is, accuracy of 96.1%, sensitivity of 95.2%, and specificity of 98.2%. Also, it has a higher precision of 95% and an F-score of 94.5%. In conv_1, the network extracts low-level features such as edges, textures, and at the output of conv_3, certain mid-level features are extracted. Finally, at the end of conv_5, high-level features such as the entire region of interest are extracted. Therefore, the first FC_6 layer is able to produce better performance after the extraction of the features from the lower level to a higher level.

Figure 11 shows the training progress in which the training accuracy is getting 100% when the number of epochs is 200 and the loss is reducing up to 0.003. Here,

TABLE 5 Performance of AlexNet with skip connection

Layers	Accuracy (%)	Sensitivity (%)	Specificity (%)	F-score (%)	Precision (%)
Conv_1	50	55	78	59	56
Conv_3	63.5	70.2	84	72.5	72
Conv_5	95	94.2	98	93	94
FC_6	96.1	95.2	98.2	94.5	95
FC_7	92	91	96	92.2	91

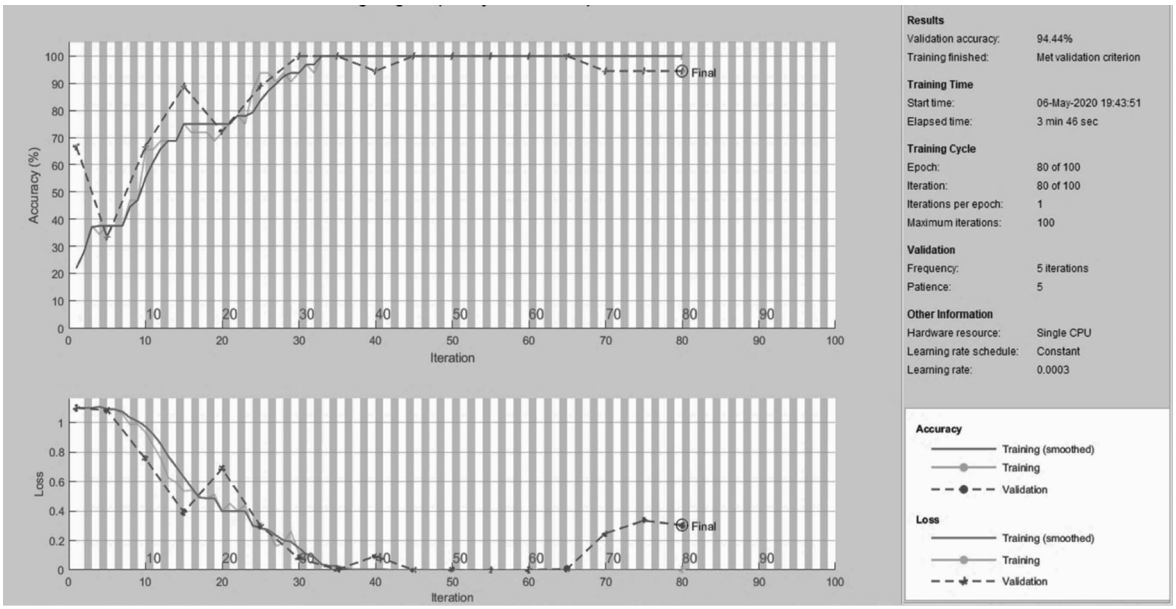
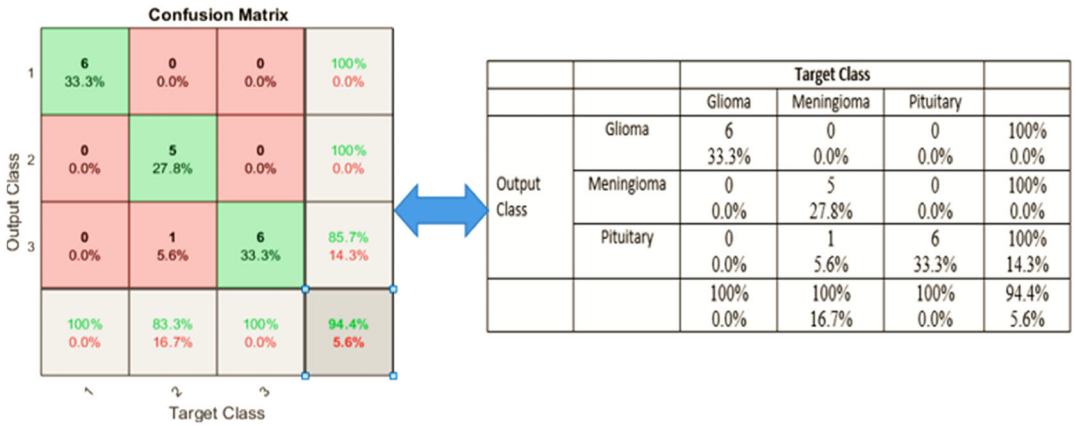


FIGURE 11 Training progress of AlexNet with skip connection



		Target Class			
		Glioma	Meningioma	Pituitary	
Output Class	Glioma	6 33.3%	0 0.0%	0 0.0%	100% 0.0%
	Meningioma	0 0.0%	5 27.8%	0 0.0%	100% 0.0%
	Pituitary	0 0.0%	1 5.6%	6 33.3%	100% 14.3%
		100% 0.0%	100% 16.7%	100% 0.0%	94.4% 5.6%

FIGURE 12 Confusion matrix for AlexNet with skip connection

the learning rate is 0.0003 and the patch size is 32. Validation accuracy is also progressing in a promising manner. There is no quick loss variation.

The numerical computation in the confusion matrix is obtained as follows:

True positive (Tp) = 6; true negative (Tn) = 5 + 0 + 1 + 6 = 12, false positive (Fp) = 0 and false negative (Fn) = 0.
The true positive = Tp/sample total; 6/18 = 33.33%. In our case, the number of trained samples has been 18 gliomas out of 20 and 17 out of meningioma; 5/17 = 27.8% and 6 from pituitary 6/18 = 33.33%.

From Figure 12, the digits 1, 2 and 3 represent the different classes of brain tumors, namely glioma, meningioma and pituitary. The performance of this brain tumor classification system is visualized by the confusion matrix as shown in Figure 12. Here, the brain tumor is classified into three classes. Therefore, the confusion matrix has a 3 × 3 matrix. From Figure 12, in the experiment, among the 60 test samples, 6 glioma test samples are identified as glioma, 3 samples are identified as meningioma, and 4 as pituitary. Likewise, for meningioma classes: 5 samples are accurately identified as meningioma, 2 samples as glioma, and 2 as pituitary. For pituitary class: 6 samples are identified as pituitary, 2 samples as glioma, and 3 samples are identified as meningioma, respectively. So, the overall accuracy for the fine-tuned AlexNet is 94.4%.

We also analyze the area under the region of convergence on both original and augmented image data. The receiving operating characteristics (ROC) is the graphical

plot used to show the diagnostic ability of the binary classifiers. For the AlexNet, on the original data, the ROC value is 95.6 and on the augmented image data, the ROC value is obtained as 97.6. It shows that the augmented image data makes the AlexNet effective, having learned more features for better classification, as shown in Figure 13.

3.2 | Performance of VGG16 with skip connection

We have analyzed the effect of each layer in VGG16 architecture. The accuracy difference between the first and last convolutional layers is 32.5%, which reveals that the first convolutional layers learning depends on the visible variations in the image, while the final convolutional layers learning depends on the image visible and invisible variations, very deeply as it gains higher accuracy.

From Table 6, after freezing some of the layers in the VGG16 architecture, the performances of convolutional layers from conv_1 to conv_13, as well as fully connected layers are examined. Each convolutional layer is learning different kind of features from the brain image. The increment of the number of convolutional layers increases the classification performance. Therefore, the final convolutional layer, that is, conv_13 provides a good performance compared to all other convolutional layers. The final fully connected layers have higher performance compared to all other layers in the VGG16 architecture, which has an accuracy of 97.2%.

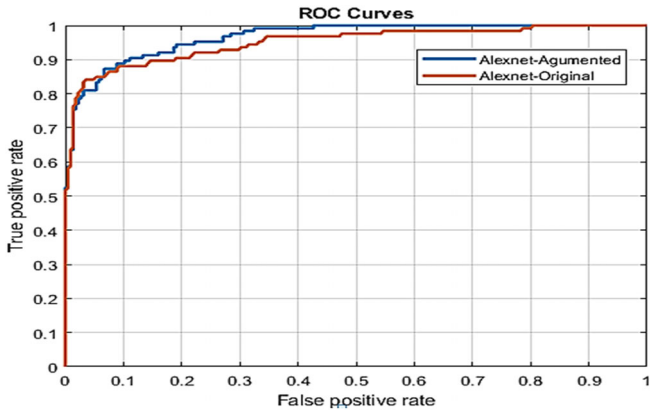


FIGURE 13 Receiving operating characteristics (ROC) performance for AlexNet with skip connection on original and augmented image data

Figure 14 shows the training progress in which the training accuracy is getting 100% when the number of epochs is 200 and the loss is reducing up to 0.003. Validation accuracy is also progressing in a promising manner but it is better than AlexNet network performance.

The numerical computation in the confusion matrix is obtained as follows:

True positive $(Tp) = 4$; true negative $(Tn) = 4 + 0 + 0 + 4 = 8$, false positive $(Fp) = 0$ and false negative $(Fn) = 0$.

The true positive = Tp/sample total; $4/12 = 33.33\%$. In our case, the number of trained samples has been 18 glioma out of 20 and 17 out of meningioma; $4/12 = 33.3.8\%$ and from pituitary; $4/12 = 33.3\%$.

From Figure 15, the digits 1, 2 and 3 represent the different classes of brain tumors, namely glioma, meningioma and pituitary. In the experiment, among the 60 test samples, 6 glioma test samples are identified as glioma, 3 samples are identified as meningioma, and 4 as pituitary. Likewise, for meningioma classes: 5 samples are accurately identified as meningioma, 2 samples as glioma, and 2 as pituitary. For pituitary class: 6 samples are identified as pituitary, 2 samples as glioma, and 3 samples are identified as meningioma, respectively. Therefore, the overall accuracy for the fine-tuned VGG network is 100.00%.

TABLE 6 Performance of VGG16 with skip connection

Layers	Accuracy (%)	Sensitivity (%)	Specificity (%)	F-score (%)	Precision (%)
Conv_1	50	55	78	59	56
Conv_3	63.5	70.2	84	72.5	72
Conv_5	65.3	73.2	85	73.5	74.5
Conv_7	68.5	75	88	76.7	77.2
Conv_9	72.5	78.5	88.4	79.2	81.2
Conv_11	77.3	82	91.5	82.2	84
Conv_13	82.5	88.2	94	89.5	91
FC_1	96.5	96.2	95.2	92.5	92.5
FC_2	98.9	96.9	97	94.2	94.8

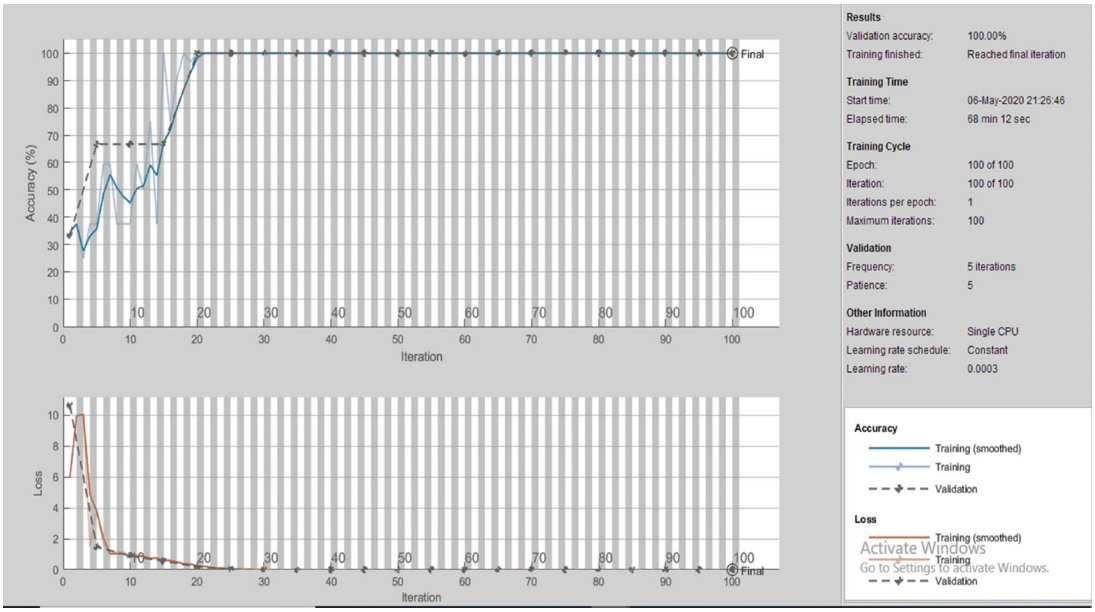


FIGURE 14 Training progress for VGG16 with skip connection

FIGURE 15 Confusion matrix for VGG16 with skip connection

		Target Class			
		Glioma	Meningioma	Pituitary	
Output Class	Glioma	4 33.3%	0 0.0%	0 0.0%	100% 0.0%
	Meningioma	0 0.0%	4 33.3%	0 0.0%	100% 0.0%
	Pituitary	0 0.0%	0 0.0%	4 33.3%	100% 100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%

For VGG16 architecture, from Figure 16, on the original data, the ROC value is 96.9, and on the augmented image data, the ROC value is obtained as 98.1. It shows that the augmented image data makes the VGG16 architecture effective, having learned more features for better classification, which is slightly better than AlexNet and lower than GoogLeNet.

3.3 | Performance of GoogLeNet with skip connection

For the GoogLeNet, the performance of the different inception network layers is analyzed. From Table 7, the performances of layers such as inception 3, inception 4, inception 5 sublayers are examined. When the inception modules are increased, the features learning increases and it leads to improvement in accuracy. The last inception module 5b produces better performance compared to other inception modules. Therefore, the learning of features in the last inception module is better than others. The accuracy of final inception module is 98.7, which is better than AlexNet and VGG16 architectures.

From Table 7, we can see that the learning variations between the inception blocks are different, in which the first inception block gives accuracy of 80% and the performance increases for consecutive inception module and the final modules gives accuracy of 97%. The difference is 16%, which is less compared to VGG16 and AlexNet architectures.

GoogLeNet training progress and validation performance are higher than both AlexNet and VGG16 architectures. Figure 17 shows loss is reducing up to

0.0005. Here, the learning rate is 0.0003 for final epoch 100.

Figure 18 shows the confusion matrix for GoogLeNet with skip connection. In the experiment, among the 60 test samples, 6 glioma test samples are identified as glioma, 3 samples are identified as meningioma, and 4 as pituitary. Likewise, for meningioma classes, 5 samples are accurately identified as meningioma, 2 samples as glioma, and 2 as pituitary. For pituitary class, 6 samples are identified as pituitary, 2 samples as glioma, and 3 samples are identified as meningioma, respectively. So, the overall accuracy for the fine-tuned GoogLeNet is 98.5%. In Figure 18, NaN means not a number.

For GoogLeNet, from Figure 19, on the original data, the ROC value is 97.9 and on the augmented image data, the ROC value is obtained as 98.6, which is better compared to the other two networks, AlexNet and VGG16 architectures.

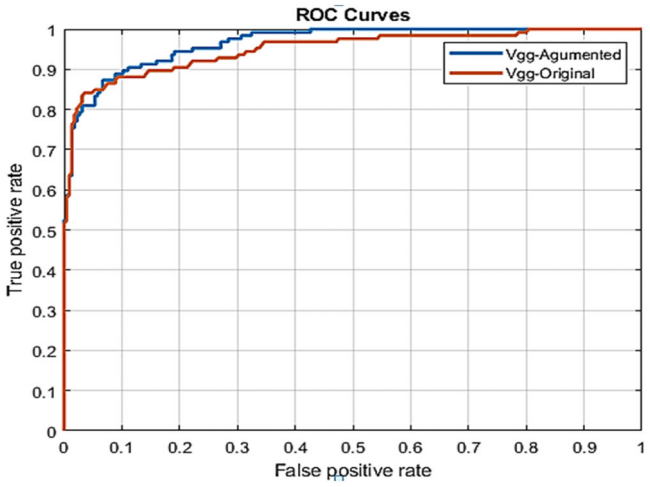


FIGURE 16 Receiving operating characteristics (ROC) for VGG16 with skip connection on original and augmented image data

4 | DISCUSSION WITH RELATED WORK

When compared with the existing methods,[30-32] the proposed skip-forward models perform well in terms of all the performance metrics. In the previous work, they have got the accuracy for all the three CNN architectures, AlexNet, VGG, GoogLeNet as 97.39, 98.69, and 98.04, respectively. However, in our case, we got 97.60, 98.99, 98.50, respectively, because we have implemented the frequency-domain information enhancement in the preprocessing and changing the potential parameters to fine-tune the network, which helps in improving the classification performance.

In References 30–32, the three CNN architectures were fine-tuned all by changing the final fully connected layer of the network to 3 for three different classes. Also, the experiments were carried out by freezing the layers for different networks. Consequently, there is a chance of getting vanishing moments and overfitting the feature maps.

However, in this article, a TL model is presented via CNN architectures with skip or shot connections based on the features learning capabilities of all the three pretrained networks. We have analyzed the layer-by-layer outcome and the feature map for all intermediate layers, and introduced the shot connections wherever needed. This leads to better performance. Also, we have employed a robust frequency-domain information enhancement technique to make image information more visible for feature extraction in the preprocessing. These two modifications have led to good results in the system performance.

Table 8 shows the performance metrics of the three CNN architectures with skip connections employed in this article, in comparison to the existing models in References 30–32, while Figure 20 shows the performances of all the TL models.

TABLE 7 Performance of GoogLeNet with skip connection

Layers	Accuracy (%)	Sensitivity (%)	Specificity (%)	F-score (%)	Precision (%)
Inception 3a	80	85	88	89	86
Inception 3b	91.5	90.2	94	92.5	92
Inception 4a	89.8	90	93.4	90	90
Inception 4b	91.1	90.2	95.2	91.5	90.7
Inception 4c	92	91	96	92.2	91
Inception 4d	94	92.5	94.1	94.2	94
Inception 4e	95.5	94	95	95.5	94.5
Inception 5a	97	96.2	96.7	97.1	96
Inception 5b	98.7	97	97.8	98.2	97.7

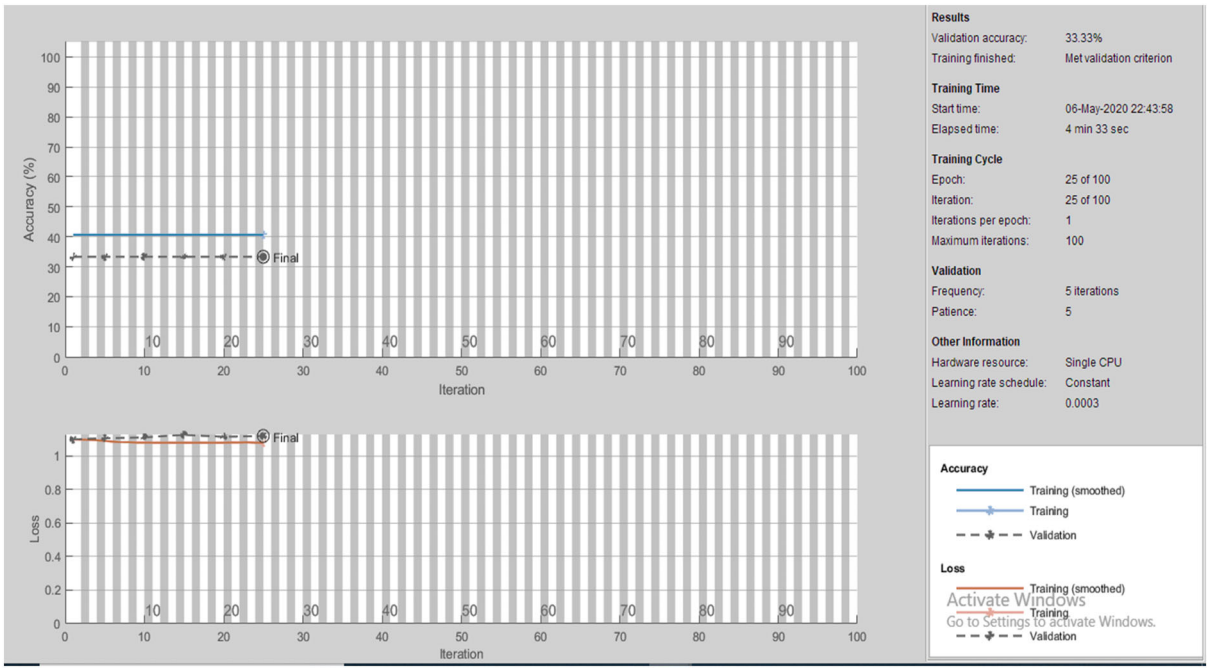
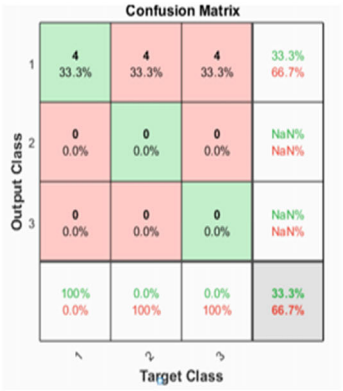


FIGURE 17 Training progress for GoogLeNet with skip connection

FIGURE 18 Confusion matrix for GoogLeNet with skip connection



		Target Class			
		Glioma	Meningioma	Pituitary	
Output Class	Glioma	4 33.3%	4 33.3%	4 33.3%	33.3% 67.7%
	Meningioma	0 0.0%	4 33.3%	0 0.0%	NaN NaN
	Pituitary	0 0.0%	0 0.0%	4 33.3%	NaN NaN
		0.0% 100%	0.0% 100%	0.0% 100%	33.3% 67.7%

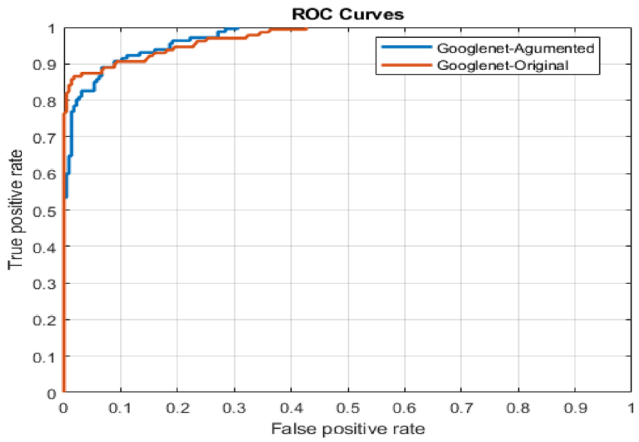


FIGURE 19 Receiving operating characteristics (ROC) performance for GoogLeNet with skip connection on original and augmented image data

From Table 8 and Figure 20, the VGG16 with shot connection performs well, that is accuracy of 98.92% compared to the other two architectures, AlexNet and GoogLeNet with shot connection. Also, when compared to the existing models,[30-32] the VGG16 with shot connection gives a 0.33% improvement in terms of accuracy.

Our TL model via CNN with skip connection has the advantages of avoiding the vanishing gradients and time complexity commonly with TL networks. In addition, our TL model via CNN with skip connection has led to the better classifications of brain tumor, with higher accuracy than the existing models. Moreover, the use of frequency-domain information enhancement technique in the preprocessing has led to better image clarity and has helped improve performance of brain tumor classifications. On the other hand, the major shortcoming of our TL

Accuracy %/models		AlexNet	VGG16	GoogLeNet
31	No skip connection	93.33	96.25	97.1
32		97.5	95	95
30		97.39	98.69	98.04
Our TL model	CNN with skip connection	97.6	98.92	98.3

Abbreviations: CNN, convolutional neural networks; TL, transfer learning.

TABLE 8 Comparison of the transfer learning models

FIGURE 20 Comparison of transfer learning models

models via CNN with skip connection is that only one dataset is used in the experimental validations with limited input classes.

5 | CONCLUSION

This article has presented a TL model via CNN with skip connection to detect the abnormality in the brain MRI. This article employs three CNN architectures, that is, AlexNet, VGG16 and GoogLeNet, to implement the TL model based on the MRI slices of brain tumor dataset for brain tumor classifications. The CNN architectures with skip connections are able to:

- attain highest accuracy of 98.92 for the VGG16 with skip connection in brain tumor classification, and
- reduce the vanishing gradient and time complexity commonly with TL networks.

When implementing transfer learning based on the MRI slices of brain tumor dataset, it is good for the performance of brain tumor classifications to employ the frequency-domain information enhancement technique in the preprocessing as it give better images clarity.

For future work, implementing the TL model may involve:

- use of the ensemble classifiers to improve the classification accuracy, and
- extension to a multiclass classifier for fine-tuning and freezing.

The TL model can be implemented in other applications such as COVID-19 MRI images, skin cancer and other types of medical image applications.

AUTHOR CONTRIBUTIONS

Saleh Alaraimi: Conceptualization of the work and simulations analysis. **Kenneth E. Okedu**: Literature review, data analysis and writing of the paper. **Richard Holden**: Data analysis of the work. **Omair Uthmani**: Literature review of the work. **Hugo Tianfield**: Supervision, data analysis and writing of the paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in (repository name, eg, "figshare") at http://doi.org/10.6084/m9.figshare.1512427.v5, Reference 34.

ORCID

Saleh Alaraimi https://orcid.org/0000-0002-3695-8961
Kenneth E. Okedu https://orcid.org/0000-0002-9635-1029

REFERENCES

1. Alafaci C, Granata F, Cutugno M, Caffo M, Caruso G, Salpietro FM. Modern neuroimaging techniques in the diagnosis of brain tumor. *Clinical Management and Evolving Novel Therapeutic Strategies for Patients with Brain Tumors*. London, UK: IntechOpen; 2013:55.
2. Ari A, Hanbay D. Deep learning-based brain tumor classification and detection system. *Turk J Electr Eng Co*. 2018;26(5): 2275-2286.
3. Despotović I, Goossens B, Philips W. MRI segmentation of the human brain: challenges, methods, and applications. *Comput Math Methods Med*. 2015;2015:1-23.
4. Dolgushin M, Kornienko V, Pronin I. Introduction to neuroimaging techniques in the diagnosis of brain cancer metastases. *Brain Metastases*. Cham: Springer; 2018:27-27.
5. Durmo F, Lätt J, Rydelius A, et al. Brain tumor characterization using multibiometric evaluation of MRI. *Tomography*. 2018;4(1):14.
6. Azer SA. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: a systematic review. *World J Gastrointest Oncol*. 2019;11(12): 1218.
7. Banerjee S, Mitra S, Sharma A, Shankar BU. A CADe system for gliomas in brain MRI using convolutional neural networks. *arXiv Preprint*. 2018:3(1):53-61.
8. Carass A, Wheeler MB, Cuzzocreo J, Bazin PL, Bassett SS, Prince JL. A joint registration and segmentation approach to skull stripping. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, New York; 2007.
9. Chaddad A, Tanougast C. Quantitative evaluation of robust skull stripping and tumor detection applied to axial MR images. *Brain Informatics*. 2016;3(1):53-61.
10. el Kader Isselmou A, Xu G, Zhang S, Saminu S, Javaid I. Deep learning algorithm for brain tumor detection and analysis using MR brain images. In *Proceedings of the 2019 International Conference on Intelligent Medicine and Health*; 2019.
11. Farmanfarma KK, Mohammadian M, Shahabinia Z, Hassanipour S, Salehiniya H. Brain cancer in the world: an epidemiological review. *World Cancer Res J*. 2019;6:5.
12. Fischmeister FPS, Höllinger I, Klinger N, et al. The benefits of skull stripping in the normalization of clinical fMRI data. *NeuroImage: Clinical*. 2013;3:369-380.
13. Gilanie G, Bajwa UI, Waraich MM, Habib Z, Ullah H, Nasir M. Classification of normal and abnormal brain MRI slices using Gabor texture and support vector machines. *Signal Image Video Process*. 2018;12(3):479-487.
14. Kuklisova-Murgasova M, Quaghebeur G, Rutherford MA, Hajnal JV, Schnabel JA. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Med Image Anal*. 2012;16(8):1550-1564.
15. Lin W, Tong T, Gao Q, et al. Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. *Front Neurosci*. 2018; 12:777.
16. Litjens G, Sánchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep*. 2016;6:26286.
17. Madhupriya G, Guru NM, Praveen S, Nivetha B. Brain tumor segmentation with deep learning technique. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE; 2019.
18. Mohsen H, El-Dahshan ESA, El-Horbaty ESM, Salem ABM. Classification using deep learning neural networks for brain tumors. *Future Comput Inform J*. 2018;3(1):68-71.
19. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. Cancer diagnosis using deep learning: a bibliographic review. *Cancers*. 2019; 11(9):1235.
20. Park BY, Byeon K, Park H. FuNP (fusion of neuroimaging preprocessing) pipelines: a fully automated preprocessing software for functional magnetic resonance imaging. *Front Neuroinform*. 2019;13:5.
21. Pugalenthi R, Rajakumar MP, Ramya J, Rajinikanth V. Evaluation and classification of the brain tumor MRI using machine learning technique. *J Control Eng Appl Inform*. 2019;21(4):12-21.
22. Rajasekaran KA, Gounder CC. Advanced brain tumor segmentation from MRI images. *High-Resolution Neuroimaging: Basic Physical Principles and Clinical Applications*. London, UK: IntechOpen; 2018:83.
23. Roy S, Maji P. A simple skull stripping algorithm for brain MRI. In *2015 8th International Conference on Advances in Pattern Recognition (ICAPR)*. IEEE; 2015.
24. Sobhaninia Z, Rezaei S, Noroozi A, et al. Brain tumor segmentation using deep learning by type specific sorting of images. *arXiv Preprint*. 2018;arXiv:1809.07786:1-4.
25. Suk HI, Lee, SW, Shen D. Alzheimer's disease Neuroimaging Initiative. Deep learning in diagnosis of brain disorders. In *Recent Progress in Brain and Cognitive Engineering*. Vol. 5. Dordrecht: Springer; 2015:203-213. https://doi.org/10.1007/978-94-017-7239-614.
26. Swiebocka-Wiek J. Skull stripping for MRI images using morphological operators. In *IFIP International Conference on Computer Information Systems and Industrial Management*. Springer, Cham; 2016.
27. Tandel GS, Biswas M, Kakde OG, et al. A review on a deep learning perspective in brain cancer classification. *Cancers*. 2019;11(1):111.
28. Widhiarso W, Yohannes Y, Prakarsah C. Brain tumor classification using gray level co-occurrence matrix and convolutional neural network. *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems)*. 2018;8(2):179-190.
29. Szegedy C, et al. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
30. Rehman A, Naz S, Razzak MI, Akram F, Imran M. A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits Syst Signal Process*. 2020;39(2):757-775. https://doi.org/10.1007/s00034-019-01246-3.
31. Bhanumathi V, Sangeetha R. 'CNN Based Training and Classification of MRI Brain Images', *2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019*. IEEE; 2019. https://doi.org/10.1109/ICACCS.2019.8728447.
32. Revi KR, Wilscy M. 'Pretrained Convolutional Neural Networks as Feature Extractor for Image Splicing Detection', *2018 International Conference on Circuits and Systems in Digital Enterprise Technology, ICCSDET 2018*. IEEE; 2018. https://doi.org/10.1109/ICCSDET.2018.8821242.

33. He K, Zhang X, Ren S, Sun J, et al. Deep residual learning for image recognition. Paper presented at: Proceeding of the IEEE Computer Society Conference on Computer Vision and Patterns Recognition; June 27-30, 2016; Las Vegas, NV, USA. https://doi.org/10.1109/CVRP.2016.90.
34. Cheng J. Brain tumor dataset. figshare.data; 2018. https://doi.org/10.6084/m9.figshare.1512427.v5. Accessed May 30, 2018.
35. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. Paper presented at: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; June 7-12, 2015; Boston, MA, USA. https://doi.org/10.1109/CVPR.2015.7298594.
36. Yuan L, Wei X, Shen H, Zeng LL, Hu D. Multi-center brain imaging classification using a novel 3D CNN approach. *IEEE Access*. 2018;6:49925-49934.
37. Gonzalez, T. F. (2007) *Handbook of Approximation Algorithms and Metaheuristics*. New York: Chapman and Hall/CRC, Taylor and Francis Group. https://doi.org/10.1201/9781420010749.
38. Yang Y, Yan LF, Zhang X, et al. Glioma grading on conventional MR images: a deep learning study with transfer learning. *Front Neurosci*. 2018;12:804.
39. Simonyan K, Zisserman A. 'Very deep convolutional networks for large-scale image recognition', *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*; 2015.

How to cite this article: Alaraimi S, Okedu KE, Tianfield H, Holden R, Uthmani O. Transfer learning networks with skip connections for classification of brain tumors. *Int J Imaging Syst Technol*. 2021;31:1564–1582. https://doi.org/10.1002/ima.22546