

BRAIN TUMOR MRI MEDICAL IMAGES CLASSIFICATION WITH DATA AUGMENTATION BY TRANSFER LEARNING OF VGG16

SABAA AHMED YAHYA AL-GALAL*, IMAD FAKHRI TAHA ALSHAIKHLI,
M. M. ABDULRAZZAQ, RAINI HASSAN

Dept. of Computer Science, International Islamic University Malaysia Jalan
Gombak, 53100, Selangor, Kuala Lumpur, Malaysia

*Corresponding Author: aljalalsaba@gmail.com

Abstract

The ability to estimate conclusions without direct human input in healthcare systems via computer algorithms is known as Artificial intelligence (AI) in healthcare. Deep learning (DL) approaches are already being employed or exploited for healthcare purposes, and in the case of medical images analysis, DL paradigms opened a world of opportunities. This paper describes creating a DL model based on transfer learning of VGG16 that can correctly classify MRI images as either (tumorous) or (non-tumorous). In addition, the model employed data augmentation in order to balance the dataset and increase the number of images. The dataset comes from the brain tumour classification project, which contains publicly available tumorous and non-tumorous images. The result showed that the model performed better with the augmented dataset, with its validation accuracy reaching ~100 %.

Keywords: A brain tumour, Classification, Medical images, MRI, Transfer learning, VGG16.

1. Introduction

A brain tumour is the result of abnormal and uncontrolled development of brain cells. The National Brain Tumour Society reports that ~700,000 people in the United States of America (USA) suffer from brain tumours, and that number will increase by 85,000 in 2021 [1]. A brain tumour is the world's 10th most common cause of death; ~3460 children under 15 were also diagnosed this year with a brain or central nervous system (CNS) tumour [2]. The early detection and classification of brain tumours is an important research domain in medical imaging, because it aids in the selection of the best treatment choice to save the patients' lives [3].

Medical imaging techniques such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans are commonly used by oncologists for early evaluations of brain tumours [4-6]. Both techniques are commonly utilised to generate highly detailed images of brain structure. Applying artificial intelligence (AI) algorithms on medical images could help doctors improve early cancer diagnosis accuracy. A variety of machine learning and deep learning (DL) algorithms have been employed for the classification, segmentation, and recognition of brain tumours, such as support vector machine (SVM), artificial neural network (ANN), and convolutional neural network (CNN). This research focuses on deep convolution neural networks (CNN), which is expected to make substantial advancements in medical image analysis [7]. According to the literature CNN is a state-of-the-art medical image analysis technique that has been effectively utilised in a variety of fields [8]. CNN are commonly employed to analyse brain images. Brain disorders, brain tissue segmentation, and anatomy were all classified in several research. Image analytics is usually done using (CNN). Generally, there are two main classifications for brain tumour medical images; the first is a binary classification for normal and abnormal categories, and the second is multi-classification for grading abnormal brain tumour images [9, 10].

This paper intends to classify brain MRI images into normal and abnormal classification using transfer learning of VGG16, then compare the results of the augmented and non-augmented datasets. The main contribution is identifying the difference between the accuracy of augmented and none-augmented dataset using best approach according to the literature. This paper is organised per the following: section two contains related works and recent articles, section three explains the methodology, dataset, and data augmentation, the fourth section details the result and discussion, and the final section concludes the work and suggest future research directions.

2. Related Works

The use of machine learning (ML) and deep learning (DL) algorithms to detect and classify brain tumour images using multiple imaging modalities, particularly MRI, has gained popularity recently. Traditional ML algorithms involve many steps before classification, mainly pre-processing and feature extraction. Feature extraction is a crucial stage in classic ML approaches since classification accuracy is dependent on the extracted features [11, 12]. In deep learning (DL) algorithms, extracted features are performed automatically, resulting in significantly improved performance. However, researchers could face obstacles when using DL models because it requires big data to train the deep learning algorithm, which could be overcome using data

augmentation, and as long as it is done appropriately, it can serve as an effective way to enlarge and balance datasets [13].

The process of assigning one or more item labels to a category is known as image classification, which is one of the most critical tasks in computer vision and pattern recognition [10, 14]. Most recent studies used the DL methods for the classification of brain tumour medical images. Sultan proposed a network consisting of 16 layers for multiclass classification of brain tumour MR images, and they reported achieving 96.13% and 98.7% accuracy for two datasets, respectively [15]. Irmak used three datasets, one for binary classification and the other two representing different grades and types of brain tumour MRI medical images; the result showed that brain tumour (binary-classification) detection is 99.33% accurate, and they reported 92.66% and 98.14 for multi-classification datasets, respectively [16].

Transfer learning is the approach used to learn different datasets using the knowledge gained via an already trained model [17]. If the dataset size required is insufficient, the network parameters can keep avoiding overfitting. When the objective size is insufficient, transfer learning methods are commonly used. There are popular architectures for this, such as AlexNet [18], GoogleNet [19], VGG19 [20], and ResNet [21]. Many researchers used transfer learning strategies in DL networks. Swati al. used transfer learning and fine-tuning on CNN (VGG19) [22], while Bakr et al. implemented transfer learning on (VGG16) [23, 24]. Deepak and Ameer used transfer learning on GoogleNet [25], while Gonella and Binaghi used V-Net to implement transfer learning [26]. There are many other attempts reported in the literature [27-30]. There are some researchers who have implemented the same dataset used on this research, Khan et al. [31] implement different algorithms using transfer learning including Inception-V3, ResNet50 and VGG16 and the validation accuracy was reported 75%, 89% and 96% respectively. Sane [32] used random tree and KNN segmentation on the same dataset, the accuracy reported is 96 %.

3. Methodology

This section explains the dataset and the proposed transfer learning model.

3.1. The dataset

The dataset has been downloaded from the Kaggle website [33], which contains 253 tumorous and non-tumorous MRI brain images. The tumorous folder contains 155 images in a folder called “yes”, while the non-tumorous images contain 98 images in a folder called “no”. Some of these images are shown in Fig. 1.

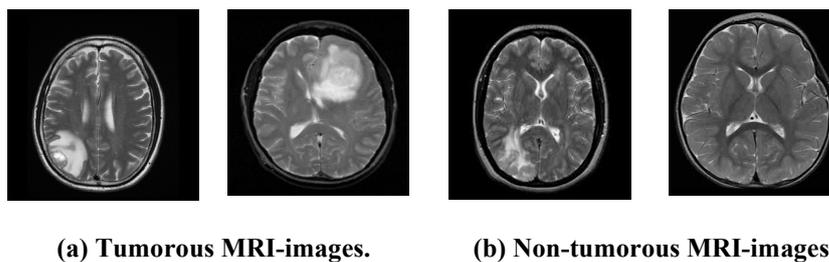


Fig. 1. Brain MRI images from the dataset.

3.2. Proposed model

Transfer learning builds on past knowledge instead of starting from scratch. We utilised a pre-trained VGG-16 convolutional neural network model that was fine-tuned by freezing parts of the layers to avoid overfitting because our dataset is considered small in DL networks. In order to increase the dataset's size, the first step was data augmentation, which will be detailed in the upcoming subsections. The flowchart of the proposed work is shown on Fig. 2.

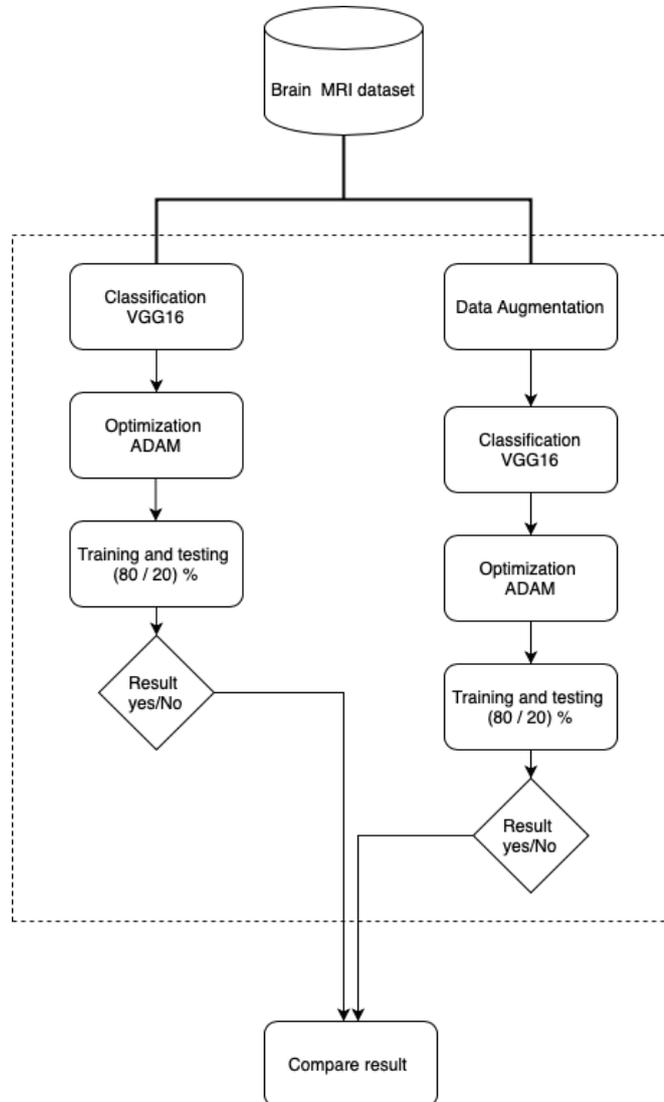


Fig. 2. flowchart of the work.

3.2.1. Dataset augmentation

A large dataset is required for the DL model to succeed. The model's efficiency can be improved by increasing the amount of data available. Researchers can use data

augmentation to improve the range of data available to the learning models without actively acquiring more data. The tumorous images have been augmented seven times, resulting in 1085 images, while the non-tumorous images have been augmented ten times, resulting in 980 images. The reason that non-tumorous images have been augmented more is to balance the dataset. The total number of the augmented dataset is 2065 images. Some of the used operations include rotation, shifting, horizontal, and vertical flips. Figure 3 shows examples of the augmented dataset. The dataset was divided into three sections for training (80%), validation (10%), and testing (10%). The training data is used to learn the model, while the validation data is used to evaluate the model and tune the parameters. The test data will be used to evaluate the proposed model in the end.

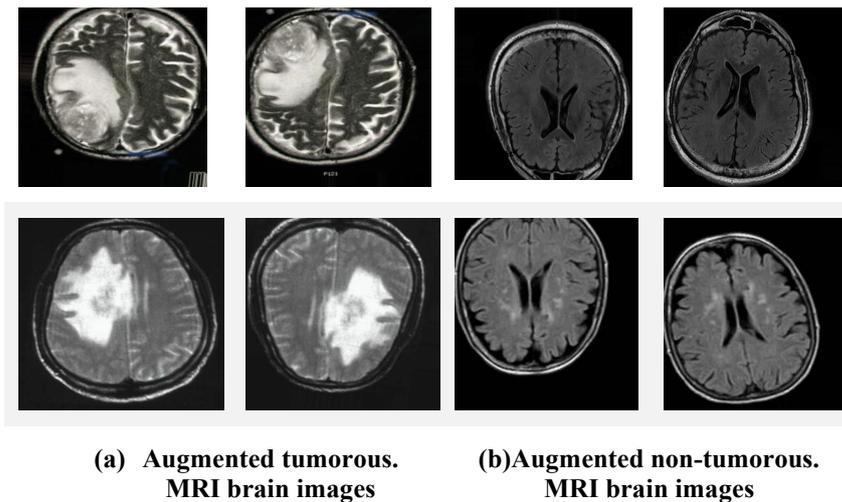


Fig. 3. Brain MRI images from the augmented dataset.

3.2.2. The Model Architecture

The model consists of 16-layers, encompassing convolutional layers, pooling layers and fully connected layers, per Fig. 4. The main layers and functions in this architecture are explained in the following subsections.

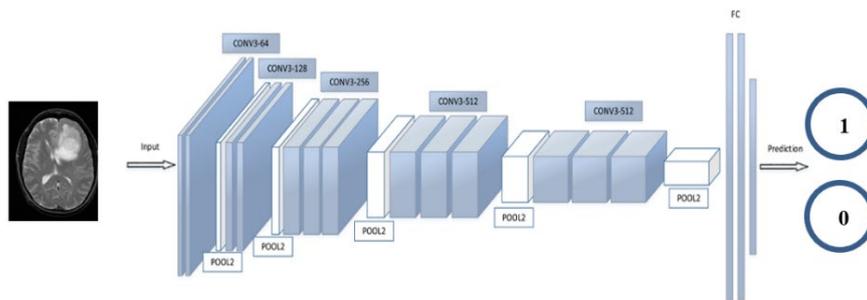


Fig. 4. Transfer learning of CNN (VGG16) architectures; (1) represents tumorous image and (0) represent non-tumorous image [34].

1) Conv-layers

Convolutional layers are the fundamental components of CNN; they use tiny input data to learn features by preserving pixel relationships [34]. It is a mathematical process with two inputs: the kernel applied on an image matrix, as shown in Fig. 5.

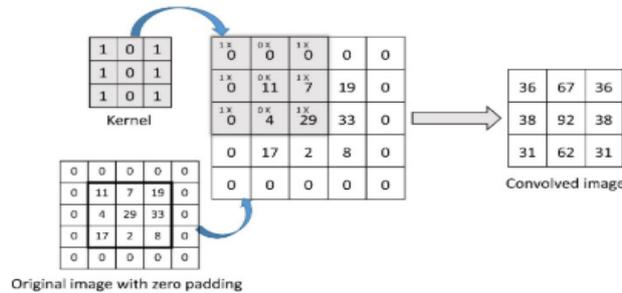


Fig. 5. Conv-layer example of input (3×3) and kernel (3×3) [35].

2) Max pooling

In this layer, each image’s dimensionality is reduced while critical information is preserved [35]. For example, Max Pooling returns the maximum value of the part of the image projected by the kernel. Another example of the pooling layer is average pooling, which returns the average value, shown in Fig. 6.

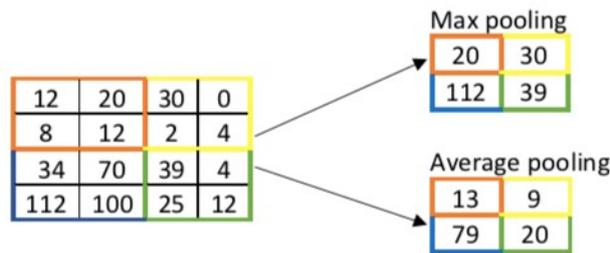


Fig. 6. Example of pooling layer (Max-Average).

3) Batch normalisation

Its purpose is to normalise the previous layers' output. Batch normalisation improves learning efficiency and can also be used as a regularisation approach to preventing model overfitting.

4) Flatten layer

In a flatten layer, data are converted from a 2-dimensional (2-D) array to flatten the one-dimension (1-D) array connected to the final classification layers (fully connected layers), as illustrated in Fig. 7.

5) Dense layers

The dense layer is fully connected. Neurons in these layers have complete links to all activations in the previous layer. The activations of these neurons can be calculated by multiplying the matrix with an offset bias. The last dense layer has two outputs as shown in the table A-1.

6) Activation function

As shown in Fig. 8, there are various activation functions; however, we will focus on Rectified Linear Units (ReLU) in this paper. The ReLU function is the most extensively utilised activation function in recent neural networks [36].

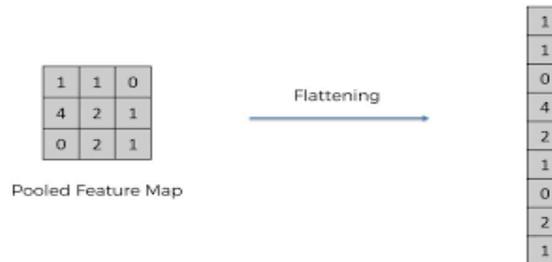


Fig. 7. Example of flatten layer.

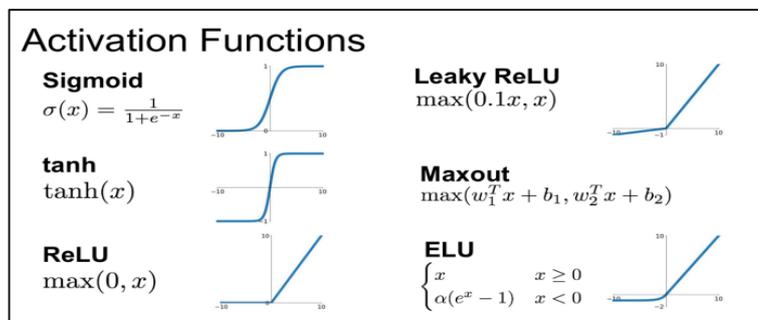


Fig. 8. Activation functions illustration [36].

7) Optimisation

CNN models frequently use the optimiser (Adam). As the implementation is simple and effective in computation and memory requirements, it is suitable for many data and/or parameters.

4. Results and Discussion

The dataset has been augmented for two reasons; first, when the dataset comprises an equal number of samples from each class, classification accuracy is an effective measure for characterising performance. However, the dataset used in this paper is uneven, which means that it needs a more thorough assessment of the proposed system using additional performance indicators. Therefore, we used data augmentation techniques in order to balance the dataset. The second reason is that DL neural network performs better in a larger dataset. The experiments were performed using Google-Colab-GPU. The dataset has been divided into 80%

training, 10 % validation, and 10% testing, and the results showed that both train and validation accuracy increased throughout the training time to reach ~100% accuracy, per Fig. 9. Table1-A in appendix A shows the layers and number of parameters in the VGG16 transfer learning model.

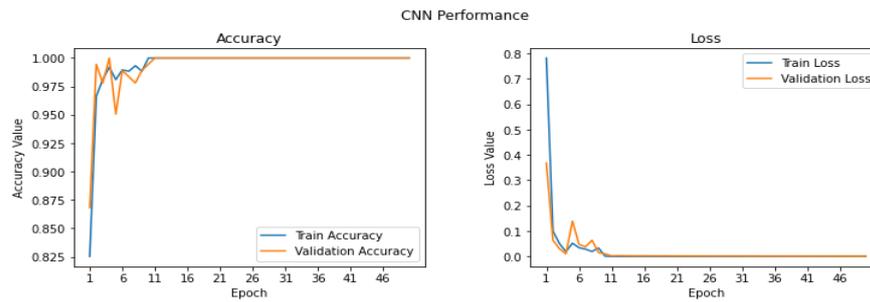


Fig. 9. Training progress: The Accuracy and the loss history represent each set validation and training history of the augmented dataset.

The same model (Transfer learning of VGG16) with the same experiment environment (Google-Colab-GPU) has been employed on the dataset without augmentation to study and compare the accuracy changes for both datasets. This experiment showed that train and validation accuracy kept increasing throughout the training iterations; however, the validation accuracy for this experiment is ~93.75 %, and the testing accuracy is 94.44%, per Fig. 10.

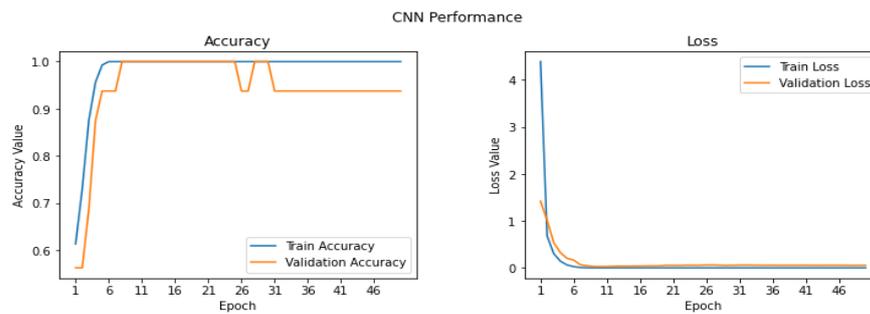


Fig. 10. Training progress: The Accuracy and the loss history represent each set validation and training history of the dataset without augmentation.

Per the above discussion, it can be concluded that data augmentation is an excellent method to improve the model's accuracy and balance the dataset. Table 1 illustrates a comparison between the proposed model and previous works.

Table 1. Test model Specifications and test conditions.

Ref.	Algorithm	Accuracy
Bakr et al. [24]	DCNN(VGG16)	96%
Sane [32]	KNN-RF	96 %
Proposed work	TL-VGG16	100 %

5. Conclusions

This research provides a brain tumour MRI medical images classification method that is accurate and automatic, with minimal pre-processing. Deep transfer learning was used to extract features from brain MRI images in the proposed system. In addition, the dataset was augmented and balanced to improve its performance. Relative to other works, the system reported the highest categorisation accuracy. Also, train and validation accuracy measures were employed to assess the system's performance and robustness, and as a result, the proposed model could be implemented to augment brain tumour early detection. Future research will focus on multi-class datasets segmentation and classification.

Acknowledgement

This project has been funded by Research Initiative Grant (KICT-RG) Research Project KICT-RG20-003-0003.

Abbreviations

AI	Artificial Intelligence
CNN	Convolutional Neural Network
DL	Deep Learning
ML	Machine Learning
MRI	Magnetic Resonance Imaging

References

1. National Brain Tumor Society: Quick brain tumor facts (2021). Retrieved July 21, 2020, from <https://braintumor.org/brain-tumor-information/brain-tumor-facts/>.
2. Brain Tumor: Statistics (2021). *Cancer. Net*. Retrieved September 2021, Retrieved July 21, 2021, from <https://www.cancer.net/cancer-types/brain-tumor/statistics>.
3. Brunese, L.; Mercaldo, F.; Reginelli, A.; and Santonella, A. (2020). An ensemble learning approach for brain cancer detection exploiting radiomic features. *Computer Methods and Programs in Biomedicine*, 185, 105134.
4. Aiello, M.; Cavaliere, C.; D'Albore, A.; and Salvatore, M. (2019). The challenges of diagnostic imaging in the era of big data. *Journal of Clinical Medicine*, 8(3), 316-327.
5. Işin, A.; Direkoğlu, C.; and Şah, M. (2016). Review of mri-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102, 317-324.
6. Kong, Y.; Gao, J.; Xu, Y.; Pan, Y.; Wang, J.; and Liu, J. (2019). Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. *Neurocomputing*, 324, 63-68.
7. Wachinger, C.; Reuter, M.; and Klein, T. (2018). DeepNAT: Deep convolutional neural network for segmenting neuroanatomy. *NeuroImage*, 170, 434-445.
8. Alqudah, A.M.; Alquran, H.H.; Abuqasmieh, I.A.; Alqudah, A.; and Al-Sharu, W. (2020). Brain tumor classification using deep learning technique - A comparison between cropped, uncropped, and segmented lesion images with

- different sizes. *International Journal of Advanced Trends in Computer Science and Engineering*, 8(6), 3684-3691.
9. Mengash, H.A.; and Mahmoud, H.A.H (2021). Brain cancer tumor classification from motion-corrected MRI images using convolutional neural network. *Computers, Materials and Continua*, 68(2), 1551-1563.
 10. Kader, I.A.E.; Xu, G.; Shuai, Z.; Saminu, S.; Javaid, I.; and Ahmad, I.S. (2021). Differential deep convolutional neural network model for brain tumor classification. *Brain Sciences*, 11(3), 352-368.
 11. Kang, J.; Ullah, Z.; and Gwak, J. (2021). Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors*, 21(6), 1-21.
 12. Al-Galal, S.A.Y.; Alshaikhli, I.F.T.; and Abdulrazzaq, M.M. (2021). MRI brain tumor medical images analysis using deep learning techniques: a systematic review. *Health and Technology*, 11(7), 267-282.
 13. Khan, A.R.; Khan, S.; Harouni, M.; Abbasi, R.; Iqbal, S.; and Mehmood, Z. (2021). Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microscopy Research and Technique*, 84(7), 1389-1399.
 14. Amin, J.; Sharif, M.; Gul, N.; Yasmin, M.; and Shad, S.A. (2020). Brain tumor classification based on DWT fusion of MRI sequences using convolutional neural network. *Pattern Recognition Letters*, 129, 115-122.
 15. Sultan, H.H.; Salem, N.M.; and Al-Atabany, W. (2019). Multi-Classification of Brain Tumor Images Using Deep Neural Network. *IEEE Access*, 7, 69215-69225.
 16. Irmak, E. (2021). Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, 9, 1-22.
 17. Pan, S.J.; and Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.
 18. Krizhevsky, A.; Sutskever, I.; and Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2, 1097-1105.
 19. Lin, M.; Chen, Q.; and Yan, S. (2014). Network in network. *2nd International Conference on Learning Representations, ICLR 2014 -Conference Track Proceedings*, United States, 1-10.
 20. Simonyan, K.; and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, University of Oxford. London, 1-14.
 21. He, K.; Zhang, X.; Ren, S.; and Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 770-778.
 22. Swati, Z.N.K.; Zhao, Q.; Kabir, M.; Ali, F.; Ali, Z., Ahmed, S.; and Lu, J. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 75, 34-46.

23. Badža, M.M.; and Barjaktarović, M.C. (2020). Classification of brain tumors from mri images using a convolutional neural network. *Applied Sciences*, 10(6), 1015-1036.
24. Siddiaue, M.A.B.; Sakib, Rahman, S.; Khan, M.M.R.; Tanzeem, A.K.; Chowdhury, M.; and Yasmin, N. (2020). Deep convolutional neural networks model-based brain tumor detection in brain MRI images. *4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, Palladam, India, 909-914.
25. Deepak, S.; and Ameer, P.M. (2020). Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. *Computers in Biology and Medicine*, 125, 1-10.
26. Gonella, G.; Binaghi, E.; Nocera, P.; and Mordacchini, C. (2019). Investigating the behaviour of machine learning techniques to segment brain metastases in radiation therapy planning. *Applied Sciences*, 9(16), 1-21.
27. Lu, S.; Lu, Z.; and Zhang, Y.-D. (2019). Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science*, 30, 41-47.
28. Rehman, A.; Naz, S.; Razzak, M.I.; Akram, F.; and Imran, M. (2019). A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits, Systems, and Signal Processing*, 39(2), 757-775.
29. Saba, T.; Mohamed, A.S.; El-Affendi, M.; Amin, J.; and Sharif, M. (2020). Brain tumor detection using fusion of hand crafted and deep learning features. *Cognitive Systems Research*, 59, 221-230.
30. Talo, M.; Baloglu, U.B.; Yıldırım, Ö.; and Acharya, U.R. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images. *Cognitive Systems Research*, 54, 176-188.
31. Khan, H.A.; Jue, W.; Mushtaq, M.; and Mushtaq, M.U. (2020). Brain tumor classification in MRI image using convolutional neural network. *Mathematical Biosciences and Engineering*, 17(5), 6203-6216.
32. Sane, U. (2021). Advanced boost brain tumor classification with random tree & knn segmentation. *International Journal of Advance Scientific Research And Engineering Trends*, 6(2), 11-17.
33. Chakrabarty, N. (2019). Brain MRI images dataset for brain tumor detection, Kaggle,. Retrieved July 21, 2021, from <https://www.kaggle.com>.
34. Trakoolwilaiwan, T.; Behboodi, B.; Lee, J.; Kim, K.; and Choi, J.-W. (2017). Convolutional neural network for high-accuracy functional near- infrared spectroscopy in a brain- computer interface. *Neurophoton*, 5(1). Retrieved July 21, 2021, from <https://www.doi.org>.
35. Zhang, Y.D.; Hou, X.X.; Chen, Y., Chen, H.; Yang, M.; Yang, J.; and Wang, S.H. (2018). Voxelwise detection of cerebral microbleed in CADASIL patients by leaky rectified linear unit and early stopping. *Multimedia Tools and Applications*, 77(17), 21825-21845.
36. Wang, S.; Jiang, Y.; Hou, X.; Cheng, H.; and Du, S. (2017). Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling. *IEEE Access*, 5, 16576-16583.

*Appendix A***Table.A-1. Layer and Number of Parameters in VGG16**

Layer (type)	Output Shape	Param #
input_1 (Input Layer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
global_average_pooling2d(G1)	(None, 512)	0
dense (Dense)	(None, 1024)	525312
dense_1 (Dense)	(None, 1024)	1049600
dense_2 (Dense)	(None, 512)	524800
dense_3 (Dense)	(None, 2)	1026
=====		
Total params: 16,815,426		
Trainable params: 16,815,426		
Non-trainable params: 0		