Comparative Study of Different Pre-trained Deep Learning Models for Brain Tumor Detection

Shiva KalyanSunder Diwakaruni[1*], Aryan Karnati[1], Shashank Reddy Boyapally[1], Dr. Supreethi K.P.[2]

[1]Student,
Department of Computer Science and Engineering,
Jawaharlal Nehru Technological University, Hyderabad, India

[2]Professor,
Department of Computer Science and Engineering,
Jawaharlal Nehru Technological University, Hyderabad, India

Abstract:

Brain tumor is the mass growth of abnormal cells in brain some of which may lead to cancer. Early detection and treatment can save the human life preventing the further growth of cells. The traditional method to detect brain tumor is from Magnetic Resonance Imaging(MRI) scans. Deep learning algorithms enabled computational models consist of multiple processing layers that represent data with multiple levels of abstraction and improved the means of recognition, prediction, and diagnosis of many life threatening diseases. Accurate analysis of MRI scans helps in the fast prediction of brain tumor and also helps the radiologist in making quick decisions. In the proposed work, concept of transfer learning is adopted and different pre-trained deep convolutional neural networks like ResNet, DenseNet, etc., are applied in detecting the presence of brain tumor using MRI dataset and their performance is analyzed.

Keywords: Magnetic Resonance Imaging(MRI), Brain Tumor, Convolutional Neural Networks(CNN), Transfer Learning, Keras.

I.INTRODUCTION

The brain is a most important organ in the human body which controls the entire functionality of other organs and helps in decision making. It is primarily the control center of the central nervous system and is responsible for performing the daily voluntary and involuntary activities in the human body.A brain tumor is a mass or growth of abnormal cells in your brain.It affects person stability and mental health. Many different types of brain tumors exist.[1]Some are noncancerous (benign), and some brain tumors are cancerous (malignant).In the United States in 2015, approximately 166,039 people were living with brain or other central nervous system tumors. Over 2018, it was projected that there would be 23,880 new cases of brain tumors and 16,830 deaths in 2018 as a result, [2] accounting for 1.4 percent of all cancers and 2.8 percent of all cancer deaths. Brain tumors are the ninth most common cancer in the, UK [3] (around 10,600 people were diagnosed in 2013), and it is the eighth most common cause of cancer death (around 5,200 people died in 2012).

MRI scan is the first suggested method for the diagnosis, pre-surgical planning and post-therapeutic monitoring of brain tumors. Magnetic resonance imaging (MRI) is a medical imaging technique that uses a magnetic field and computer-generated radio waves to create detailed images of the organs and tissues in your body[4]. Among other techniques, MRI is the most popular and risk-free. MRI provides higher contrast for soft tissues of the brain with high resolution than CT images. MRI helps in easy detection of abnormalities in the brain.

With the advances in the field of medical imaging, every organ in human body can be investigated with large number of images produced. However, manual interpretation of large amounts of MR images by radiologists becomes a laborious task. Different tumors have different effects on body, some of which may be dangerous. Accurate detection and analysis is required to increase the life span of a person.[5]Computer-aided systems with semi-automatic or automatic analysis of images has become an important research topic.

Brain tumor detection aims to detect the location and active brain tumor. This is done by comparing abnormal tissue with normal tissues. The detection is used for treatment planning methods, disease monitoring and disease progression. Advanced technologies like deep learning can be used to detect brain tumors accurately which are cost effective and time saving. [6]Some studies reveal that convolutional neural networks(CNN) perform better in detecting brain tumor because they do not require manual segmentation and perform fully automatic classification.

One of the solution for increasing performance of deep learning is transfer learning. Transfer learning is a part of machine learning and artificial intelligence where knowledge gained from one task is applied to a different but similar task. Transfer learning has different pre-trained models like ResNet, InceptionV3, DenseNet, etc., These models are available in Keras library.

In this study, different pre-trained models are applied to detect the brain tumor likelihood by analyzing MR images. Kaggle is the source of MRI dataset used in the experiment. The design involves some image pre-processing techniques and training the models. The performances of all models were evaluated in terms of accuracy, f1-score, precision and recall.

Section 2 presents about the related work carried out in order to study and do the experiment. Section 3 describes about the methodology involved in the study. Section 4 explains the experimental setup and the resultsof the study. Conclusion is mentioned in section 5.

II. RELATED WORK

Many techniques have been proposed for classification of brain tumors in MR images, most notably, fuzzy clustering means (FCM), support vector machine (SVM), artificial neural network (ANN), knowledge-based techniques, and expectation-maximization (EM) algorithm technique to extract important information from medical imaging modules.

Sachdeva et al. [7] have presented ANN and PCA-ANN methods for multiclass brain tumor classification, segmentation, and feature extraction.

In the research work of Fau et al. [8], the Convolutional Neural Network (CNN) was implemented for the detection of meningioma, glioma and pituitary tumor respectively from MRI images. In this paper, techniques like data acquisition, data preprocessing, pre-model, model optimization and hyper parameter tuning are applied.

Agn et al. [9]proposed a deep learning-based model that adjusts the parameters in brain tumor radiotherapy planning to minimize the risk for healthy tissues.

Sharma et al. [10] have presented a technique utilizing texture-primitive features with artificial neural network (ANN) as segmentation and classifier tool in the classification of brain tumor from MR images.

In the paper of Saddam et al. [11], an automated segmentation algorithm for brain tumor using deep convolutional neural networks (DCNN) is proposed. Deep networks tend to have a lot of parameters thus overfitting is almost always an issue especially when data are sparse. The Patch-based training method is used for the model. The proposed algorithm includes preprocessing in which images are normalized and bias field corrected, and post-processing where small false positives are removed using morphological operators.

Transfer learning has been used for analyzing different medical problems. Jain et al. [12]used a pre-trained VGG-16 network to diagnose Alzheimer's disease via MR images. In the research,Deniz et al. [13] used VGG-16 and AlexNet model for breast cancer diagnosis and then they classified tumors using SVM. Kaur and Gandhi [14]conducted a study that showed that transfer learning is more effective in brain tumor classification than traditional machine learning.

III. METHODS AND METHODOLOGIES

In this study, classification approach for brain MR images was presented using transfer learning of deep convolutional neural networks. A dataset of 3672 MR images was used from Kaggle. The images are first preprocessed before the learning process. The workflow involved in the process is given in Fig.1.

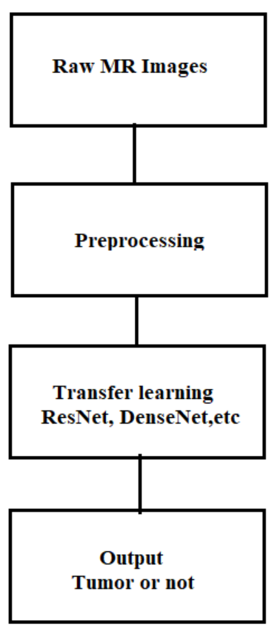

Fig.1. Flowchart of the study

3.1 Dataset

The dataset used in the study consists of 3762 brain MR images. The dataset is taken from Kaggle open source website. 2100 of these are images without tumor and 1662 images are with tumor. The graph containing the tumor (Class-1) and non-tumor (Class-0) image distributions within the dataset is shown in Fig.2.

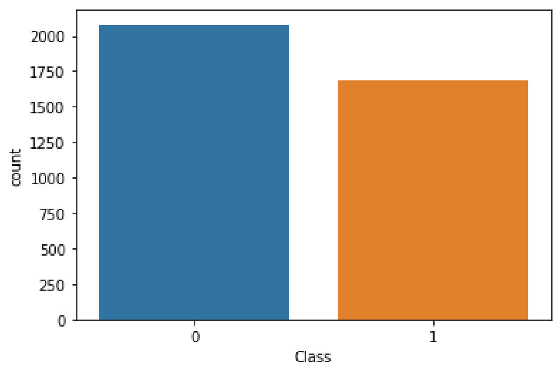

Fig.2.Image Distribution of dataset

3.2 Data Preprocessing

During the preprocessing stage, all the raw images are cropped according to the boundaries and polar points. The boundaries of the raw images were determined and the part outside was discarded. This makes data processing easier. Images of different width and length values were resized to 224*224-pixel size after cropping.Then this data is processed into a single numpy array using different pre-trained model's preprocess input function which scales the pixel values between -1 and 1.

Deep learning models generally require large amounts of data.Hence, different data augmentation techniques like flipping, rotations and translation are used to avoid overfitting of large-capacity learners and memorizing training sets. Original images are flipped horizontally and vertically. Vertical flip is performed by rotating an image by 180 degrees and then a horizontal flip on that. The images are rotated at right angles to preserve the image dimensions. Translation is doneby moving the images along the X, Y and both directions.

3.3 Transfer Learning

Transfer learning has a great potential for computer-aided detection of medical problems. Transfer learning is to transfer the weights of a network previously trained with large amounts of data to another model created to solve a similar problem. This method is important if there is small amount of data. Transfer learning works by taking a pre-trained model, repurposing the learned knowledge and loading it into our environment and finally fine tuning it to achieve higher accuracy. The calculated weights of the pre-trained model are transferred to the new model and only the classifiers in the last part of the new model are trained. The pre-trained models in the transfer learning are originally applied on ImageNet dataset.

ImageNet is formally a project aimed at (manually) labeling and categorizing images into almost 22,000 separate object categories for the purpose of computer vision research.

ImageNet in the context of deep learning and Convolutional Neural Networks refers to the ImageNet Large Scale Visual Recognition Challenge or ILSVRC for short. The goal of this image classification challenge is to train a model that can correctly classify an input image into 1,000 separate object categories.

In this study, brain MRI classification was done using pre-trained ResNet, DenseNet, InceptionV3, VGG16, etc., and their different versions. These models also demonstrate a strong ability to generalize the images outside the ImageNet dataset via transferlearning, such as feature extraction and fine-tuning.

3.3.1 ResNet or Residual Network

In order to solve a complex problem, we stack some additional layers in the deep neural networks which results in improved accuracy and performance. But it has been found that there is a maximum threshold depth in neural networks. If we add more layers, the performance eventually degrades beyond this limit. This problem has been alleviated using ResNet which was introduced by He et al. in 2015 and the model became the winner of ImageNet competition with a 3.57% error rate. The Residual block of ResNet is shown in Fig.3.

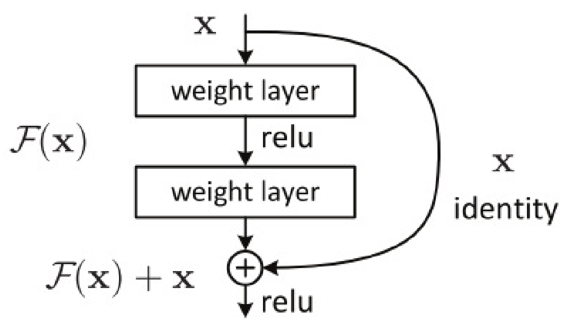

Fig.3. Residual block

In ResNet, there is a direct connection called skip connection which skips some layers in between. Due to this skip connection, [15]the output is not the same now. Without skip connection, the output is H(x)=f(wx+b) or H(x)=f(x). With skip connection, it is H(x)=f(x)+x. The skip connections solve the problem of vanishing gradient descent by allowing the model to learn the identity functions using residual blocks.

Different versions of ResNet are ResNet50, ResNet50V2, ResNet152V2, etc., where number denoting the number of weighted layers. The architecture of ResNet50 is shown in Fig.4.

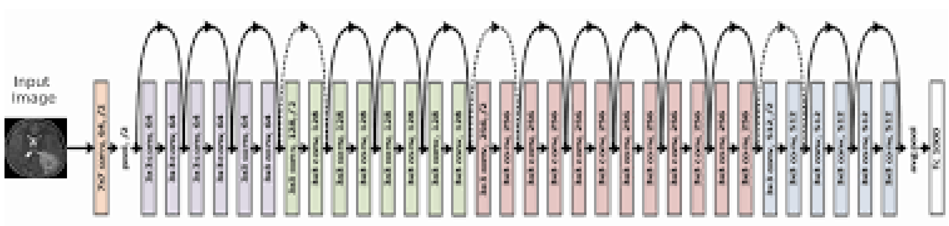

Fig.4. Resnet50 Architecture

3.3.2 DenseNet

A DenseNet is a type of convolutional neural network invented by Cornwell University, Tsinghua University and Facebook AI Research (FAIR). It uses dense connections between layers, through Dense Blocks, where we connect all layers (with matching feature-map sizes) directly with each other. In this model, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers. [16] Each layer is receiving a “collective knowledge” from all preceding layers. Dense block is shown in Fig.5.

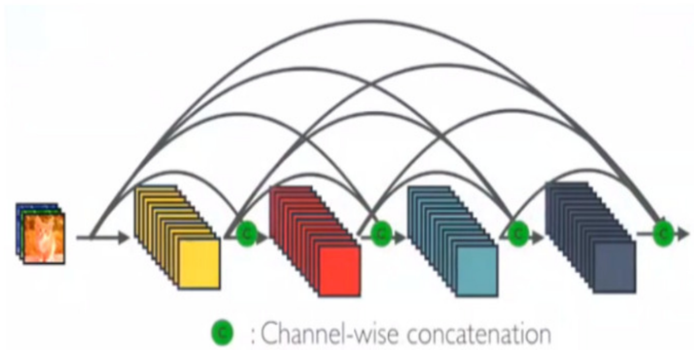

Fig.5. Dense Block

DenseNet starts with a basic convolution and pooling layer. Then there is a dense block followed by a transition layer, another dense block followed by a transition layer, another dense block followed by a transition layer, and finally a dense block followed by a classification layer. There are different versions of DenseNet models namely DenseNet121, DenseNet169, DenseNet201 where the number denotes the number of trainable weight layers. The architecture of DenseNet is shown in Fig.6.

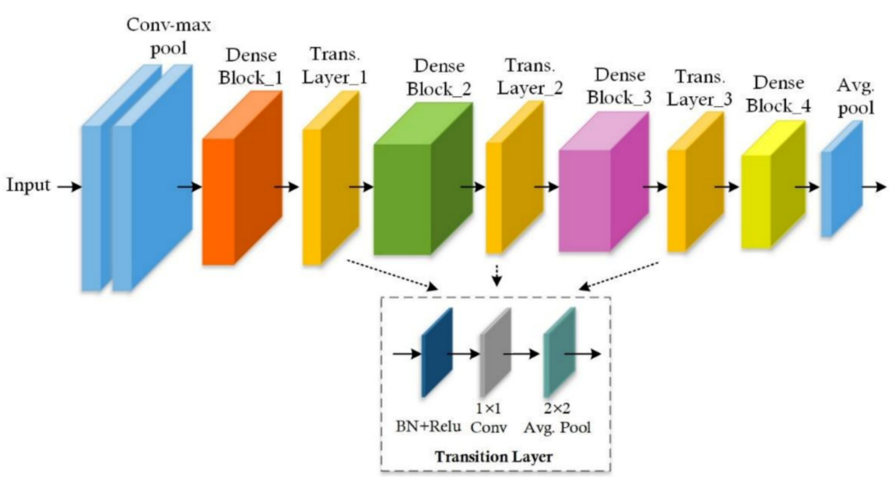

Fig.6. DenseNet Architecture

3.3.3 InceptionV3

InceptionV3 is a convolutional neural network for assisting in image analysis and object detection, and got its start as a module for GoogleNet. It is the third edition of Google's Inception Convolutional Neural Network, originally introduced during the ImageNet Recognition Challenge. The model was proposed in the paper “Rethinking the Inception

Architecture for Computer Vision", published in 2015 by Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, and Jonathon Shlens.

InceptionV3 was designed making minor changes in the previous versions. It consists of 42 layers. In addition to the previous versions, there are label smoothing, normalization for auxiliary classifiers, 7x7 convolution, and RMSProp optimizer [17]. The architecture of the InceptionV3 model is shown in Fig.7.

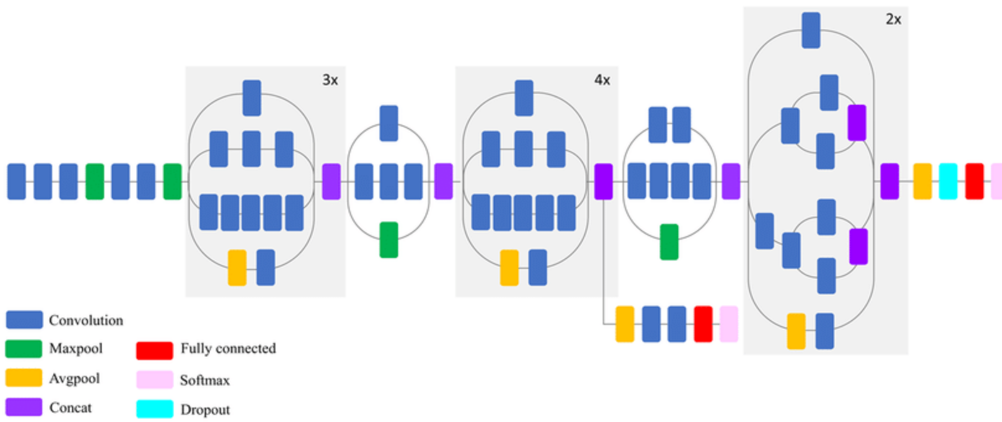

Fig.7. InceptionV3 architecture

3.3.4 MobileNet

MobileNet is a class of CNN open-sourced by Google. MobileNet model is designed to be used in mobile applications and is TensorFlow's first Mobile Computer Vision model. They can be built upon for classification, detection, embeddings, and segmentation.MobileNet uses depth-wise separable convolutions. A depth-wise separable convolution is made from two operations i.e., depth-wise convolution and pointwise convolution [18].

The main difference between MobileNet architecture and a traditional CNN instead of a single 3x3 convolution layer followed by the batch norm and ReLU. Mobile Nets split the convolution into a 3x3 depth-wise convolution and a 1x1 pointwise convolution, as shown in the Fig.8.

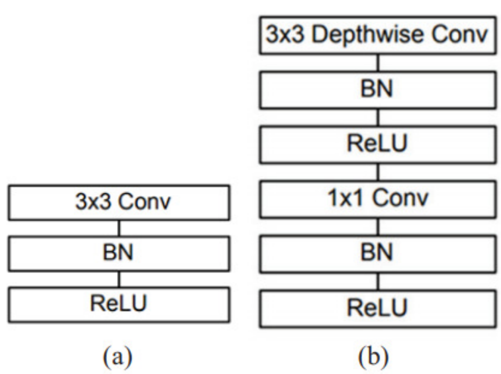

Fig.8.(a) Standard Convolutional layer (b)Depth-wise separable convolution layer

MobileNetV2 is based on an inverted residual structure where the residual connections are between the bottleneck layers. It is 53 layers deep. In this model, there are two types of blocksas shown in below Fig.9. One is residual block with stride of 1. Another one is block with stride of 2 for downsizing. There are 3 layers for both types of blocks. The firstlayer is 1×1convolutionwithReLU6. The secondlayer is the depth-wiseconvolution. The thirdlayer is another 1×1convolutionbut without any non-linearity.

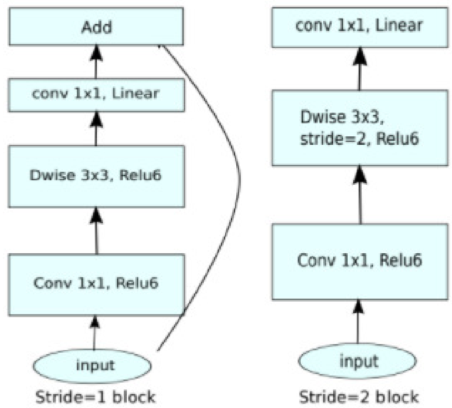

Fig.9. MobileNetV2 Architecture

3.3.5 VGG-16

VGG network architecture is a CNN model introduced in 2014 by Simonyan and Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition". VGG-16 is a special VGG type with 16 weighted layers. Its layers are convolution, maxpooling, activation, and fully connected layer [19]. The architecture of VGG-16 is shown in the Fig.10.

Fig.10. VGG-16 architecture

There are 21 layers in the architecture: 13 convolution, 5 pooling and 3 dense layers. 16 of these layers are weighted layers. The model achieves 92.7% top-5 test accuracy on ImageNet dataset which contains 14 million images belonging to 1000 classes.

IV. EXPERIMENTAL STUDY

4.1 Implementation

The proposed models are implemented in python by using various powerful deep learning libraries and frameworks like Keras and TensorFlow respectively. It also uses numpy, sklearn, etc., libraries. The model is implemented using Google Colab with Tesla K80 GPU accelerator. The raw data is preprocessed and the processedimage dataset is split into 80%

training data i.e., 3009 images and remaining 20% as test data i.e., 753 images. The training data is given to the pre-trained models for training with 200 epochs. In training the models, Stochastic Gradient optimizer is used. It adapts the learning rate to obtain optimization in the result. After the training, testing and prediction are done to get the F1 score, accuracy and precision.

4.2 Evaluation Parameters

Confusion matrix also known as error matrix is a table that visualizes the performance of an algorithm. Confusion matrix parameters are considered as evaluation parameters to evaluate the experiment results. There are four cases in confusion matrix True positive (TP), False Positive (FP), True Negative (TN), False Negative (FN) as seen in Fig.11.

TP and TN are the cases in which both prediction and actual values are positive and negative respectively. FP is a case in which prediction made is positive but the actual value is negative and FN is a case of negative prediction and actual positive. Based on these four parameters, F1-score, recall, precision and accuracy are calculated as in Eq. (1), Eq. (2), Eq. (3) and Eq. (4) respectively.

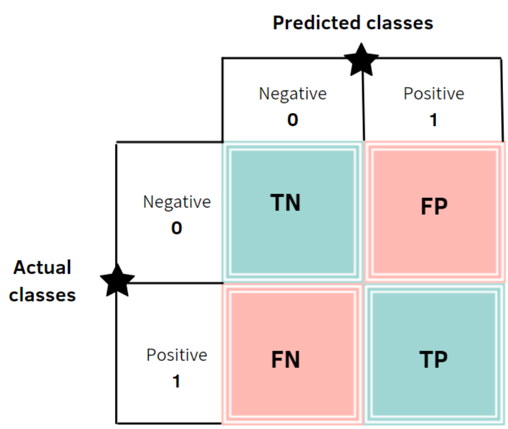

Fig.11. Confusion Matrix

$$F1\ Score = 2 * (Recall * Precision) / (Recall + Precision) \quad (1)$$

$$Recall = TP/(TP + FN) \quad (2)$$

$$Precision = TP/(TP + FP) \quad (3)$$

$$Accuracy = (TP + TN)/ (TP + FP + FN + TN) \quad (4)$$

F1 score is high if there is a balance between precision and recall.

4.3 Findings and Discussion

In this study, the dataset consisting of brain MR images was classified using different pre-trained models. All other external parameters such as learning rate optimization, batch size, and number of epochs were the same for each model applied.

The accuracy and the loss graphs on epoch basis are given for top five models. Besides, the performance of each model was compared in terms of the specified metrics.

ResNet50 model: The accuracy and loss graphs obtained are given in Fig.12.

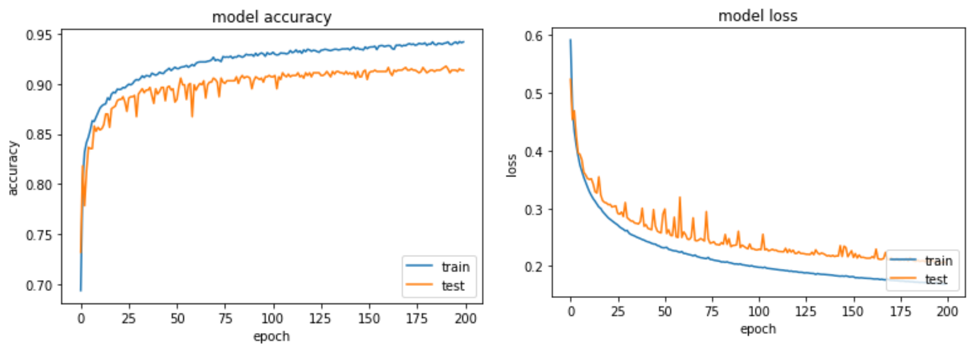

Fig.12.Accuracy and loss graphs of ResNet50 model

DenseNet169 model: The accuracy and loss graphs obtained are given in Fig.13.

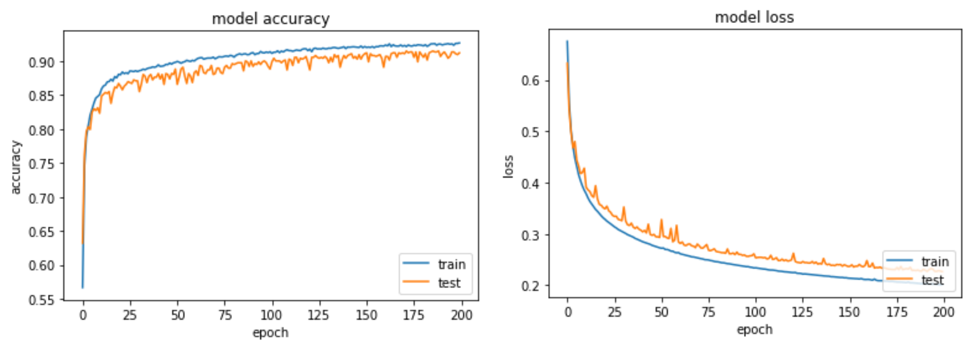

Fig.13. Accuracy and loss graphs of DenseNet169 model

InceptionV3 model: The accuracy and loss graphs obtained are given in Fig.14.

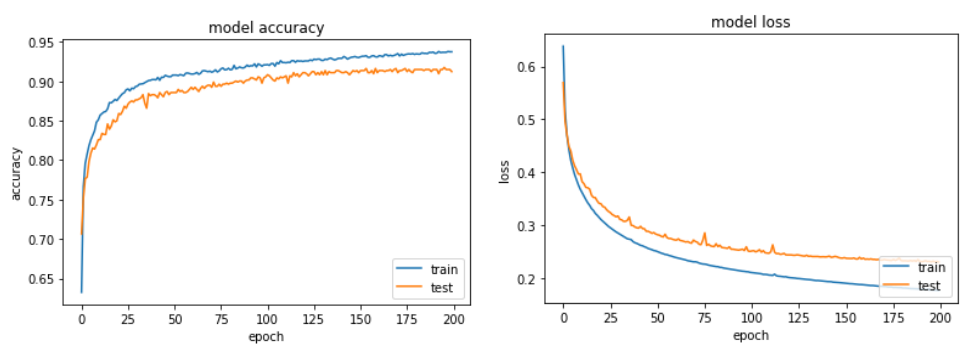

Fig.14. Accuracy and loss graphs of InceptionV3 model

MobileNetV2 model: The accuracy and loss graphs obtained are given in Fig.15.

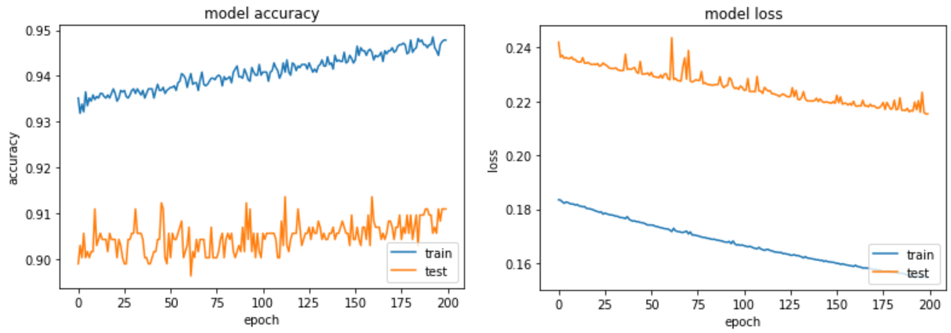

Fig.15. Accuracy and loss graphs of MobileNetV2 model

VGG16 model: The accuracy and loss graphs obtained are given in Fig.16.

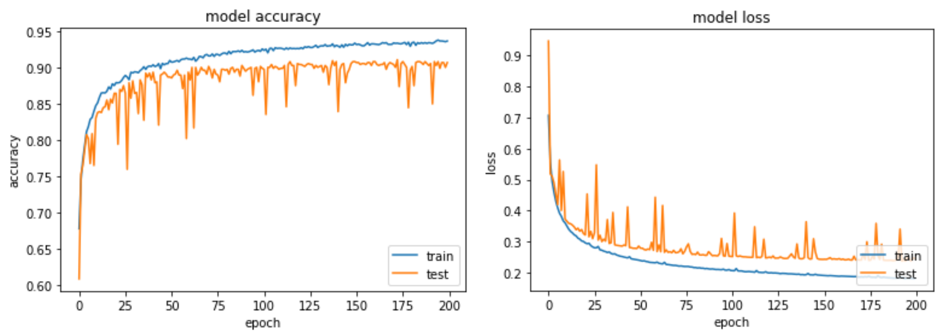

Fig.16. Accuracy and loss graphs of VGG16 model

The performance results of the all the pre-trained models for the same number of epochs are given in Table 1.

Table 1. Performance Results (*: Top-5 models)

Model	**Accuracy**	**Loss**	**Precision**	**Recall**	**F1-score**
ResNet50*	91.37	20.57	91.43	91.37	91.35
DenseNet169*	91.24	22.69	91.29	91.23	91.21
DenseNet201*	91.24	22.73	91.26	91.23	91.22
InceptionV3*	91.24	23.1	91.29	91.23	91.22
MobileNetV2*	91.1	21.54	89.77	89.77	89.77
ResNet152V2	90.97	23.93	90.99	90.97	90.96
VGG16	90.7	24.52	91.03	90.7	90.65
ResNet101V2	89.64	25.41	89.81	89.64	89.65
DenseNet121	89.38	24.38	89.45	89.37	89.34
InceptionResNetV2	89.38	27.52	89.61	89.37	89.33
ResNet50V2	88.98	26.8	89.08	88.98	88.94

XceptionNet	88.84	29.19	89.04	88.84	88.8
EfficientNetB4	87.65	30.84	87.77	87.65	87.61
NasNetMobile	84.99	34.65	85.16	84.99	84.93

V. CONCLUSION

In the proposed work, different pre-trained models have been explored which were not used before in brain tumor detection like XceptionNet, EfficientNet and NasNet. The accuracies of these models are not so high as compared to others. In the study, confusion matrix is used to compare the models. From the above performance results Table 1, the models ResNet50, DenseNet169, DenseNet201, InceptionV3 and MobileNetV2 have accuracies 91.37, 91.24, 91.24, 91.24 and 91.1 respectively. Since the models have very less difference between the accuracies, we compared the models based on the number of true positives and false positives. From Fig. 17, if we compare ResNet50 and DenseNet201 confusion matrices, the former model has less false positives than the latter one. Apart from that, ResNet50 model has the lowest loss. Therefore, ResNet50 model performed comparatively better on given dataset than other models.

(a)

	Negative	Positive
Negative	376	24
Positive	41	312

(b)

	Negative	Positive
Negative	375	25
Positive	41	312

(c)

	Negative	Positive
Negative	373	27
Positive	39	314

(d)

	Negative	Positive
Negative	375	25
Positive	41	312

(e)

	Negative	Positive
Negative	364	36
Positive	41	312

Fig.17. Confusion matrices of top-5 models: (a) ResNet50 model (b) DenseNet169 model (c) DenseNet201 model(d) InceptionV3 model (e) MobileNetV2 model

Acknowledgement
The author gives thanks for this work being supported by Jawaharlal Nehru Technological University, Hyderabad. The author extends the thanks to the professor, Dr. Supreethi K.P., Jawaharlal Nehru Technological University, Hyderabad.

References
[1] Brain tumor by Mayo Clinic Staff. (2021) https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084
[2] Brain Cancer Stat Facts. (2021)https://seer.cancer.gov/statfacts/html/brain.html
[3] Cancer Research UK, Brain Tumor Stats.(2021)http://www.cancerresearchuk.org/cancer-info/cancerstats/types/brain/
[4] MRI information by Mayo Clinic. (2021) https://www.mayoclinic.org/tests-procedures/mri/about/pac-20384768
[5] Bauer, Stefan, Roland Wiest, Lutz-P. Nolte, and Mauricio Reyes. 2013. "A Survey of MRI-Based Medical Image Analysis for Brain Tumor Studies." Physics in Medicine & Biology 58(13): R97.
[6] Kamnitsas, Konstantinos, Christian Ledig, Virginia FJ Newcombe, Joanna P. Simpson, Andrew D. Kane, David K. Menon, Daniel Rueckert, and Ben Glocker. 2017. "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation." Medical Image Analysis 36:61–78.
[7] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Segmentation, feature extraction, and multiclass brain tumor classification," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1141–1150, 2013.
[8] Fausto Milletari Seyed-Ahmad Ahmadi Christine Kroll Annika Plate Verena Rozanski Juliana Maiostre Johannes Levin Olaf Dietrich Birgit Ertl-Wagner Kai Bötzel, Nassir Navab 2016 Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound Elsevier Inc 164 92-102.
[9] Agn, Mikael, Per Munck af Rosenschöld, Oula Puonti, Michael J. Lundemann, Laura Mancini, Anastasia Papadaki, Steffi Thust, John Ashburner, Ian Law, and Koen Van Leemput. 2019. "A Modality-Adaptive Method for Segmenting Brain Tumors and Organs-at-Risk in Radiation Therapy Planning." Medical Image Analysis 54:220–237
[10] N. Sharma, A. Ray, S. Sharma, K. Shukla, S. Pradhan, and L. Aggarwal, "Segmentation and classification of medical images using texture-primitive features: application of BAM-type artificial neural network," *Journal of Medical Physics*, vol. 33, no. 3, pp. 119–126, 2008.
[11] Saddam Hussain, Syed Muhammad Anwar, Muhammad Majid, "Brain Tumor Segmentation using Cascaded Deep Convolution Neural Network", pp 1998-2001
[12] Jain, Rachna, Nikita Jain, Akshay Aggarwal, and D. Jude Hemanth. 2019. "Convolutional Neural Network Based Alzheimer's Disease Classification from Magnetic Resonance Brain Images." Cognitive Systems Research 57:147–59. DOI: 10.1016/j.cogsys.2018.12.015.
[13] Deniz, Erkan, Abdulkadir Şengür, Zehra Kadiroğlu, Yanhui Guo, Varun Bajaj, and Ümit Budak. 2018. "Transfer Learning Based Histopathologic Image Classification for Breast Cancer Detection." Health Information Science and Systems 6(1):18.
[14] Kaur, Taranjit, and Tapan Kumar Gandhi. 2020. "Deep Convolutional Neural Networks with Transfer Learning for Automated Brain Image Classification." Machine Vision and Applications 31:1–16.
[15] ResNet by Hussain Mujtaba. (2020) https://www.mygreatlearning.com/blog/resnet/

[16] DenseNet by Sik-Ho Tsang. (2018) https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803
[17] InceptionV3 by Vihar Kurama. (2020) https://blog.paperspace.com/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet/
[18] MobileNet by Abhijeet Pujara. (2020) https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470
[19] Vgg16 by NeuroHive. (2018) https://neurohive.io/en/popular-networks/vgg16/