

# **BIO MEDICAL IMAGE PROCESSING – BRAIN TUMOR CLASSIFICATION**

*Submitted by*

**Team AI Triad**

Anitha Madhusudhan, Ashitha Menon, Kavya Brungi

## Contents

<b>1. Introduction.....</b>	<b>4</b>
<b>2. Problem Statement.....</b>	<b>5</b>
<b>3. Dataset, Image Pre-processing, and Augmentation .....</b>	<b>5</b>
<b>3.1. Image Database .....</b>	<b>5</b>
<b>3.2. Image Pre-Processing: .....</b>	<b>5</b>
<b>3.3. Data Augmentation .....</b>	<b>6</b>
<b>4. Architecture.....</b>	<b>8</b>
<b>4.1. Convolutional Neural Networks .....</b>	<b>8</b>
4.1.1. Hyper – Parameter Tuning .....	9
<b>4.2. VGG -16 Architecture using Transfer Learning.....</b>	<b>9</b>
4.2.1. Hyper – Parameter Tuning .....	10
<b>4.3. ResNet -18 Architecture using Transfer Learning .....</b>	<b>10</b>
4.3.1. Hyper-parameter Tuning .....	11
<b>5. Comparative Analysis of Explored Models .....</b>	<b>13</b>
<b>6. User Interface Web application .....</b>	<b>13</b>
<b>7. Conclusion .....</b>	<b>14</b>
<b>8. References.....</b>	<b>15</b>

### **Executive Summary**

The detection, segmentation, and classification of brain tumors from an MRI Image is a herculean task if conducted manually. The Brain is a complex organ; it is challenging to have a complete understanding of the nature of the tumor. The utmost focus should be on the reduction of human judgment in the diagnosis of diseases since slight mistakes can cause loss of life. The diagnosis and the treatment of the disease depend on the type of tumor. In this study, improvement in the accuracy of classifying the correct type of brain tumor (gliomas, meningioma, and pituitary) is emphasized, for which Deep Learning techniques are used. Specifically used is Residual Networks (ResNet- 18) Architecture based on Convolutional Neural Networks to classify the model accurately. The proposed technique's experimental results have been evaluated and validated for performance and quality analysis on magnetic resonance brain images, based on accuracy. The experimental results achieved 92% accuracy, demonstrating the effectiveness of the proposed technique.

*Key words: Brain tumor, Classification, Deep learning, Magnetic resonance images, Convolutional Neural Network, Transfer learning, ResNet-18 Architecture, Validation Accuracy.*

## 1. Introduction

Brain tumors are the consequence of abnormal growths and uncontrolled cell division in the brain. The skull which encompasses the brain is very restricted and any growth in such restricted space can cause problems. However, brain tumors are quite common and can catapult into an aggressive disease in a smaller period, which can decrease the patients' life expectancy to a considerable degree. The growth of a tumor in the brain can vary greatly from person to person, based on the growth rate, or location, and many more reasons.

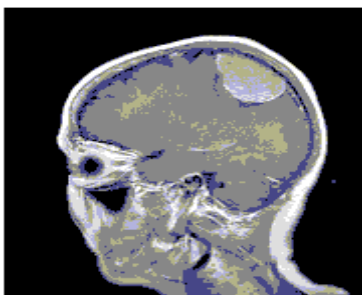
The study conducted is on malignant (cancerous) and primary (that begins in the brain) brain tumors. Primary brain tumors originate in the brain itself or tissues close to it, such as in the brain-covering membranes (meninges), cranial nerves, pituitary gland, or pineal gland. Some types of primary brain tumors such as Meningioma, Glioma, and Pituitary tumors are more common than the others. Brain Stem Gliomas are located at the base of the brain. Pituitary Tumors are located at the centre of the brain behind and above the nose, near the pituitary gland. Meningioma tumors grow from the meninges, the layers of tissue covering the brain and spinal cord. As they grow, meningiomas compress adjacent brain tissue. The study focuses on the classification of tumors into the above mentioned three types and a healthy brain (no tumor).

Magnetic Resonance Imaging (MRI) is a medical imaging technique and is mainly used to detect brain tumors. MRI image gives more information about a given medical image than the CT or ultrasound image. MRI image provides detailed information about brain structure and anomaly detection in brain tissue. The MRI images are taken from three different directions. These views are called **sagittal, axial, and coronal**. Brain Tumor segmentation techniques are a critical component in tumor detection.

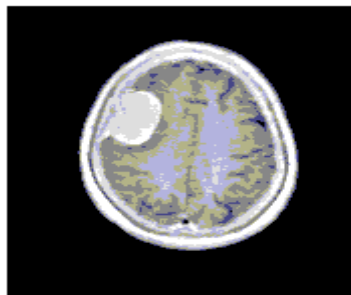
**Sagittal View:** Cross-section along the midline of the brain for the side view (left/right).

**Axial View:** Horizontal cross-section along the centre of the brain for the top/bottom view along the plane.

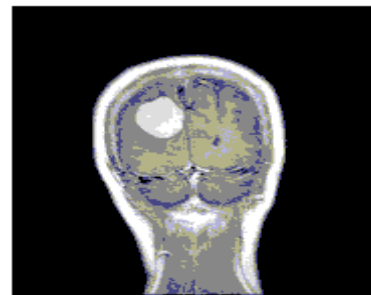
**Coronal View:** Cross-section along the centre of the brain for the front/back view along the plane.



Sagittal view



Axial view



Coronal view

An MRI uses magnetic fields, not x-rays, to produce detailed images of the body. MRIs create more detailed pictures than CT scans. In general, diagnosing a brain tumor usually begins with magnetic resonance imaging (MRI)

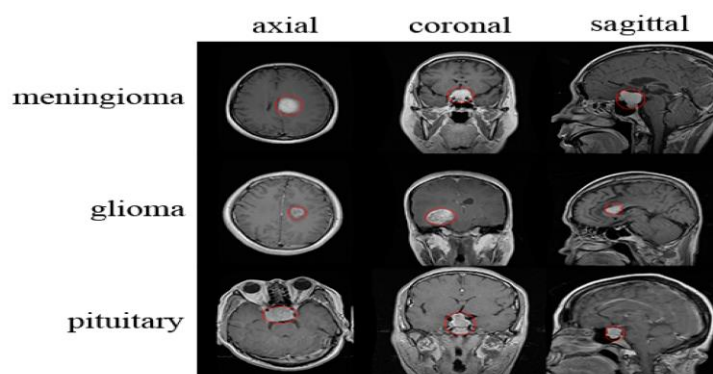
## 2. Problem Statement

The data generated by MRI can be huge and it can be a daunting task to manually classify the tumor in a shorter period. Here the proposed system is automatic brain tumor detection and classification which will reduce the visual inspection process and can be considered more accurate. The study aims to build a model which to a reasonable accuracy can classify the images into three common types of brain tumors (gliomas, meningioma tumors, and pituitary tumors) and no tumor. The study is conducted using a region-based method for the classification of tumors in MRI images, based on Convolutional Neural Networks (CNNs) exploiting Residual Networks (ResNet-18).

## 3. Dataset, Image Pre-processing, and Augmentation

### 3.1. Image Database

The dataset contains MRI images of three types of the tumor as well as images(segmented) of a healthy brain. The dataset of the images is split into training and testing folders with a subdivision in each for the respective tumor classes. The training dataset has 900 files of glioma tumor, 915 files of meningioma tumor, 874 files of pituitary tumor, and 487 files of no tumor. Every tumor data contains all the 3-cross-section view of images.



**Image 1.1:** Representation of normalized magnetic resonance imaging (MRI) images showing different types of tumors in different planes. In the images, the tumor is marked with a red outline.

*Source : mdpi.com - Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network*

The dataset images were already converted to JPEG(.jpg) format from the Dicom (.dcm) format before the pre-processing and augmentation techniques were applied to the images.

### 3.2. Image Pre-Processing:

Image Pre-processing is the usage of algorithms to process images to improve the data, known as features, of the images by suppressing unnecessary distortions or noise. It can enhance the important features to classify accurately and build a model from the improved model.

The main aim in brain image examination is the level of accuracy and time taken in the identification of the tumor if detected. In the study, the steps taken to process the images include grayscale conversion, resizing, and denoise (negative image) which are explained further.

An image is nothing more than a two-dimensional array of numbers (or pixels) ranging between 0 and 255. It is defined by the mathematical function  $f(x, y)$  where  $x$  and  $y$  are the two coordinates horizontally and vertically. The value of  $f(x, y)$  at any point is giving the pixel value at that point of an image.

Numerous standards have been developed over the years to describe how an image can be stored within a file. However, once the image is loaded into memory, it typically takes one of two forms: indexed or RGB (red, green, blue). An RGB image is a three-dimensional byte array that explicitly stores a colour value for each pixel. RGB image arrays are made up of width, height, and three channels of colour information i.e. Red, Blue, and Green. The dataset used for this study contains MRI Images which are already in a greyscale form. Greyscale image pixel values lie between 0 to 255 and range from black to white with different shades of grey. Light intensity at each pixel in a greyscale image lies in a single band of the electromagnetic spectrum.

The images represent the input layer of the CNN network, they are resized to 150 X 150 dimensions / Pixels.

The only denoising filter used in the images is negative imaging which is a technique to reverse the different shades found in the greyscale image which originally ranges from 0 – 255 with 0 as white and 255 as black. In this technique, each pixel value of the input image is subtracted from the total levels in the image (255) and mapped onto the output image. The result makes the blacker shades greyer and the whiter shades in the images darker which emphasizes the features but keeps it intact without removing any important information. Brain images are complex in nature and removal of noise an important procedure for accurate identification of the tumor. In this study, the effort is made to keep accuracy in the examination of the brain images without removing any feature.

### **3.3. Data Augmentation**

Image Augmentation is used to increase or artificially expand the dataset to avoid overfitting the predicted model. Various parameters like flipping, rotation, scaling, translation is used to increase the data count of the sample set by 3 or 4 times.

In the study, both linear and affine transformations are used.

A flip (mirror effect) is done by reversing the pixels horizontally or vertically and is a type of linear transformation. For Flipping the image horizontally, the pixel situated at coordinate  $(x, y)$  will be situated at coordinate  $(width - x - 1, y)$  in the new image. An existing image is flipped to expand the dataset which will help the machine to identify and classify the tumor from different types of views.

Applying an affine transformation to a uniformly distorted image can correct for a range of perspective distortions by transforming the measurements from the ideal coordinates to those used.

An affine transformation is an important class of linear 2-D geometric transformations which maps variables (e.g. pixel intensity values located at position  $(x1,y1)$  in an input image) into new variables (e.g.  $(x2,y2)$  in an output image) by applying a linear combination of translation, rotation, scaling and/or shearing (i.e. non-uniform scaling in some directions) operations.

In the study, the images have been rotated at a 45-degree angle. Scaling operations of the zoom out on a scale of (1.5,1.5) and zoom in to a scale of (0.7,0.7) were performed.

The translate operator performs a geometric transformation which maps the position of each picture element in an input image into a new position in an output image, where the translation, an image element located at  $(x1, y1)$  in the original is shifted to a new position  $(x2, y2)$  in the corresponding output image by displacing it through a user-specified translation  $(\beta1, \beta2)$ . The translation used for the operation is displaced to a position at (25,25).

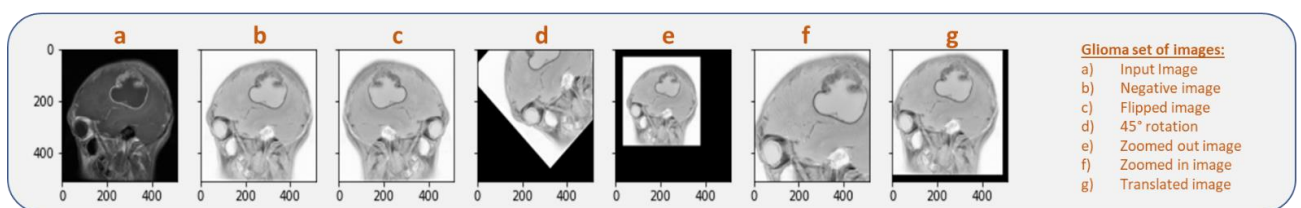


Image 3.3.1. Images representing Glioma Tumor through the application of pre- processing and augmentation techniques.

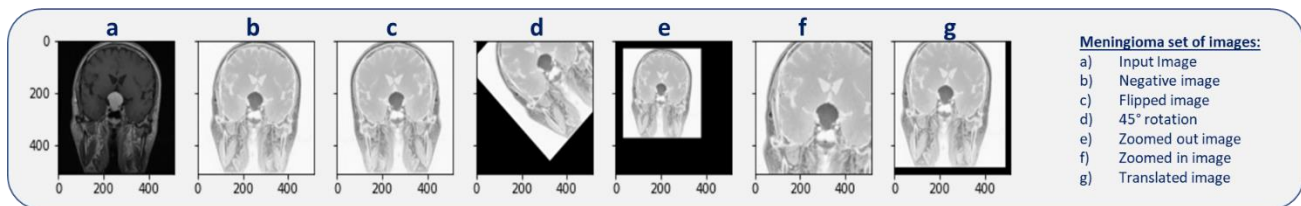


Image 3.3.2. Images representing meningioma tumor through the application of pre-processing and augmentation techniques

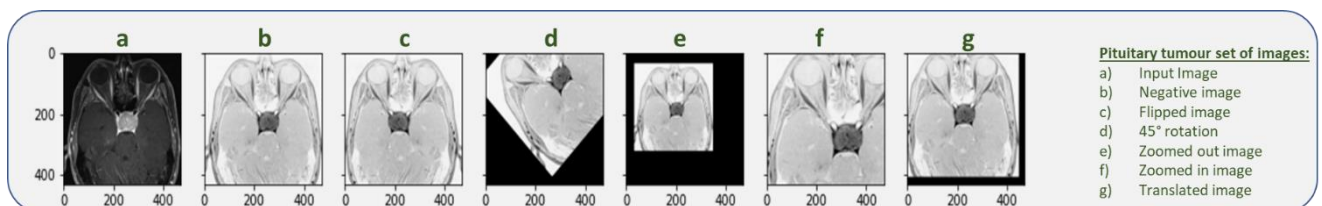


Image 3.3.3. Images representing pituitary tumor through the application of pre- processing and augmentation techniques

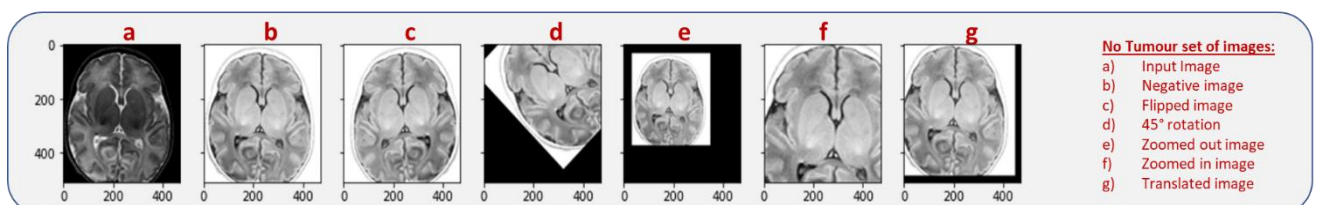


Image 3.3.4. Images representing no tumor brain through the application of pre- processing and augmentation techniques

After image augmentation techniques were performed, the image dataset significantly increased from 3,176 images to 11,116 images.

#### 4. Architecture

Brain Tumor classification is particularly important to develop an effective treatment plan. If the classification procedure is automated, especially the MRI Images, it will significantly reduce the time for diagnosis and time have taken to develop the plan of treatment. The classification accuracy depends on the extracted features which depend on the expert competencies. Deep Learning methods include neural networks to classify images and Convolutional Neural Networks (CNN's) are the most useful in various applications including classification.

In the primary research conducted before the study, various networks were explored, and the most viable methods were considered to build the prediction model. The networks on which experiments for the study conducted were, CNN model without transfer learning, Transfer learning models on VGG-16 Architecture and ResNet-18 Architecture. The model which gave the least loss and the highest validation accuracy was finally considered.

##### 4.1. Convolutional Neural Networks

CNN's are usually composed of several layers, each involving linear as well as nonlinear operators, that are learned jointly, in an end-to-end manner, to solve a particular task. CNN architectures have been designed to learn spatial hierarchies of features by building multiple blocks: convolution layers with a set of filters, pooling layers, and fully connected layers (FCLs).

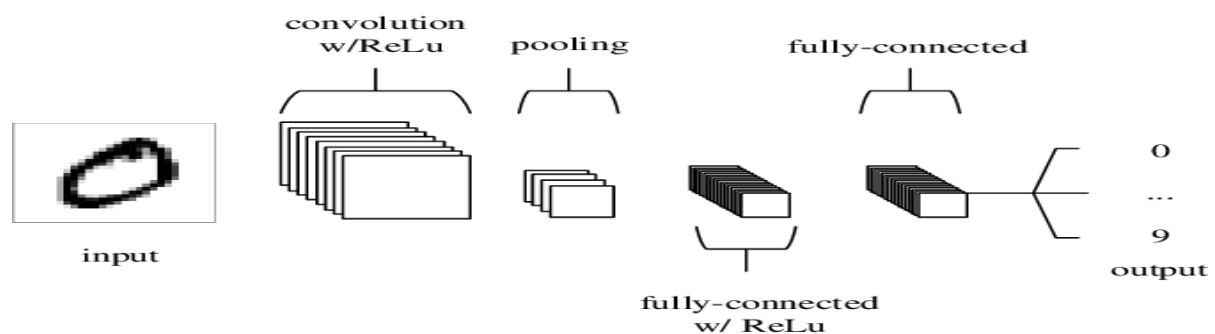


Fig 4.1. – CNN Architecture:

Source – researchgate.net

To construct the CNN Architecture, 5 layers were built : The first layer is the 5 x 5 convolutional layer with 64 output channels(filters) and a stride of 2 with a dropout rate of 0.25 followed by two, 3×3 maximum pooling layers with a stride of 2 and 128 output channels. The dropout rate is kept the same as 0.25 in the second layer. The fourth and the fifth layer consists of 2 x 2 maximum pooling layer with a stride of 2 and 128 and 256 output channels, respectively. The dropout rate is slightly increased to 0.3 in the fourth and fifth layer. The layers are then flattened to form the fully connected layer. ReLU (Rectifier Linear Unit) is the activation function used with a dropout rate of 0.5.



#### 4.1.1. Hyper – Parameter Tuning

The Hyper – parameters used for compiling the model is Adam optimizer with default values of learning rate as 0.001, decay as 0, and accuracy as metrics. The dataset has been trained for 10 epochs for a batch size of 16 with the loss function as categorical cross-entropy.

The validation accuracy obtained in the experiment was 29%, which became the main reason to turn to transfer learning models.

#### 4.2. VGG -16 Architecture using Transfer Learning

Training the networks with transfer learning is usually much faster. Transfer learning uses the gained knowledge that solves one problem and applied it to solve different related problems by using a trained model to learn a different set of data.

CNN's training becomes more complicated and can lead to overfitting because of the sample size of medical datasets. Also, applying deep pre-trained CNNs based on transfer learning in medical imaging needs to adjust the hyper-parameters and learning parameters of the models to achieve a good result.

The VGG network architecture is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth. Reducing volume size is handled by max pooling. It always uses 3 x 3 filters with stride of 1 in convolution layer and uses same padding in pooling layers 2 x 2 with a stride of 2. Two fully connected layers, each with 4,096 nodes are then followed by a softmax classifier. The “16” stand for the number of weight layers in the network.

	Layer	Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	224 x 224 x 3	-	-	-
1	2 X Convolution	64	224 x 224 x 64	3x3	1	relu
	Max Pooling	64	112 x 112 x 64	3x3	2	relu
3	2 X Convolution	128	112 x 112 x 128	3x3	1	relu
	Max Pooling	128	56 x 56 x 128	3x3	2	relu
5	2 X Convolution	256	56 x 56 x 256	3x3	1	relu
	Max Pooling	256	28 x 28 x 256	3x3	2	relu
7	3 X Convolution	512	28 x 28 x 512	3x3	1	relu
	Max Pooling	512	14 x 14 x 512	3x3	2	relu
10	3 X Convolution	512	14 x 14 x 512	3x3	1	relu
	Max Pooling	512	7 x 7 x 512	3x3	2	relu
13	FC	-	25088	-	-	relu
14	FC	-	4096	-	-	relu
15	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

Fig 4.2. VGG-16 Architecture

*Source – researchgate.net*

#### 4.2.1. Hyper – Parameter Tuning

The hyper-parameters used for compiling the model is SGD optimizer with default values of learning rate as 0.0001, decay as 0, and accuracy as metrics. The dataset has been trained for 20 epochs with the loss function as categorical cross-entropy. For callbacks, checkpointing is set to earlystopping and to monitor the validation accuracy of the model.

The validation accuracy improved in this experiment to 43%. The next experiment was conducted with ResNet-18 Architecture to check whether there will be a further improvement in the validation accuracy.

#### 4.3. ResNet -18 Architecture using Transfer Learning

ResNet which is a short form of Residual Networks is a network of 18 convolutional layers. The architecture of ResNet- 18 is briefly explained as follows-

Increasing the depth of the network for accuracy is imperative for better accuracy in a deep learning model. The aim of this study is also to increase the accuracy of the classification of brain tumors. In each network architecture, layers are built upon for increasing the depth of the network. However, by going on stacking layers the network depth may not be achieved. Deep networks are hard to train because of the notorious vanishing gradient problem — as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient infinitely small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly. The skip connection introduced by ResNet is used to pass on information from the higher layers to the lower layers. It helps the gradients to flow smoothly from layer to layer.

The below figure shows the *residual block* of ResNet, where the solid line carrying the layer input to the addition operator is called a *residual connection* (or *shortcut connection*). With residual blocks, inputs can forward propagate faster through the residual connections across layers.

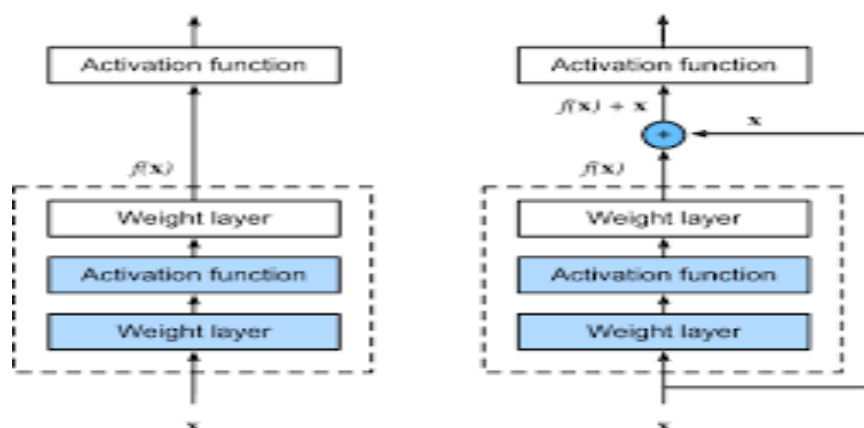


Fig 4.3 – A regular block (left) and a residual block (right).

Source - [https://d2l.ai/chapter\\_convolutional-modern/resnet.html](https://d2l.ai/chapter_convolutional-modern/resnet.html)

The first two layers of ResNet - 18 are the  $7 \times 7$  convolutional layer with 64 output channels and a stride of 2 which is followed by the  $3 \times 3$  maximum pooling layer with a stride of 2. Each convolutional layer is followed by a batch normalization layer and a ReLU activation function.

ResNet-18 uses four modules made up of residual blocks, each of which uses several residual blocks with the same number of output channels. The number of channels in the first module is the same as the number of input channels. Since a maximum pooling layer with a stride of 2 has already been used, it is not necessary to reduce the height and width. In the first residual block for each of the subsequent modules, the number of channels is doubled compared with that of the previous module, and the height and width are halved.

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	$7 \times 7, 64$ , stride 2
conv2_x	$56 \times 56 \times 64$	$3 \times 3$ max pool, stride 2 $\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	$7 \times 7$ average pool
fully connected	1000	$512 \times 1000$ fully connections
softmax	1000	

Fig 4.4. – ResNet-18 architecture

Source – *researchgate.net*

#### 4.3.1. Hyper-parameter Tuning

In the study, for compiling the model Stochastic Gradient Optimizer is used (SGD) with default values of learning rate as 0.01, decay as 0, accuracy as metrics. The dataset has been trained for 20 epochs for a batch size of 16. For fine-tuning the hyperparameters, 20 ablation experiments have been performed, with the loss function as categorical cross-entropy.

The above was considered for tuning the hyper-parameters after Adam optimizer and decay gave a comparatively lesser accuracy.

The study uses 30% of the data for validation. For callbacks, checkpointing is set up for checking the Receiving Operative Characteristic (ROC) curve and Area Under Curve(AUC) score and to save the network weights only when there is an improvement in classification accuracy on the validation dataset (monitor= 'val\_accuracy' and mode= 'auto'). The entire

model including the weights are saved to the file 'AKA.h5' and it is loaded only when there is a classification accuracy improvement on the validation dataset.

Callbacks at early stopping were set up in the previous architectures of CNN and VGG-16. In the experimental study, it was observed that the validation accuracy was comparatively lesser.

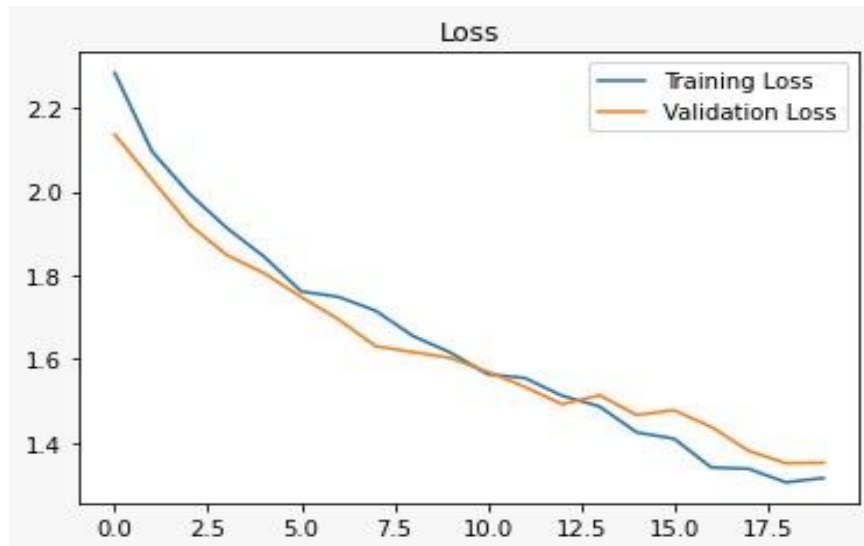


Figure 4.3.1. Figure represents the decrease in the loss function.

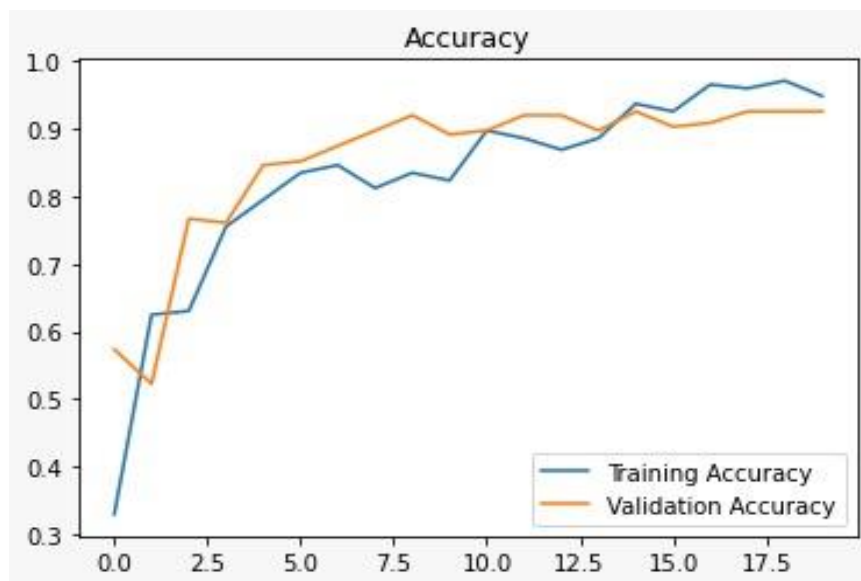


Figure 4.3.2. Figure represents increase in the training and validation accuracy

The model was tested on 90 images. The metric used is Accuracy which uses a confusion matrix to predict the output of the image loaded. In the study, the rate of the tumor classified accurately using the test metric is 90%.

## 5. Comparative Analysis of Explored Models

To summarize the findings from the experiments, the data has been consolidated and given below in a tabular format.

Model	Layers	Training Accuracy	Validation Accuracy
CNN without Transfer Learning	05	69%	29%
VGG – 16 using Transfer Learning	16	54%	43%
ResNet – 18 using Transfer Learning	18	94%	96%

Table 5.1. Comparative analysis of Architectures.

***Based on the findings on the training and test accuracy from the above three experiments, the pre-trained CNN network architecture adopted in the study is based on the ResNet-18.***

## 6. User Interface Web application

The results of the study have been bound to build an application with customized user interfaces. Since the model has been built on Python, Dash has been used to build the bio-analytical tool. For this study, the dash web application has been hosted locally.

Dash is a productive Python framework for building web analytic applications which has been an extension of Flask, Plotly and React. Dash is ideal for building data visualization apps with highly custom user interfaces.

The dash application is built with the core components of dash accomplished with html and bootstrap components. Previously explained model is incorporated in the dash to create a user-interface for the model.

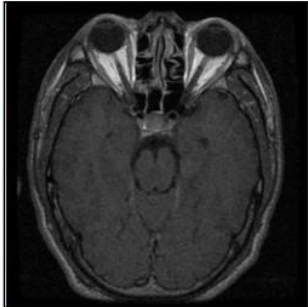
Below are the briefly steps followed in the dash web analytic application which was created to deploy the model locally:

1. Defined the dash bootstrap external sheet (dbc.themes.BOOTSTRAP) for the dash application.
2. Built the complete html structure with the respective callback on top of the external sheet.
3. Html structure also includes the input methods for uploading the test image which will be passed on to the model.
4. Ran the predefined pre-processing steps and passed on the image to the saved model (AKA.h5).
5. With the help of the callback the output is generated on the web application.
6. Output contains detection and classification of the type of the tumor. And, a brief explanation about the classified tumor and the probability distribution. The probability distribution is the weightage of each class of tumor generated for the given image.

## Bio-medical Image Processing: Brain Tumor Classification

Upload Files here

Upload a brain MRI Image which has a good resolution and wait for the results!



A tumor has been detected which may be a Pituitary tumor

Pituitary tumors are abnormal growths that develop in your pituitary gland. Most pituitary tumors are noncancerous (benign) growths that remain in your pituitary gland or surrounding tissues.

Tumor Classes	Glioma	Menengioma	Pituitary
Probability in %	27.11%	18.31%	51.16%

Probability distribution: Probability distribution is the weightage for each class generated for the given image. With the combination of the model accuracy and probability rate, doctor\clinician can make a more accurate diagnosis.

Disclaimer: The user should upload .jpg image which has a high resolution and has not been edited in any form.

This is a pilot project and is not suggested to use it for diagnosis without prior approval!

### 7. Conclusion

Brain tumour detection and classification has been always a labour-intensive process where even a slight error by the Doctors can cost human lives. With the advent of newer technologies and integration of data science with medical science, Deep learning has paved way to automate the entire process by applying artificial intelligence. The model will be a big step forward in personalizing treatments and ensuring better patient outcomes. In this study, we have compared various architectures and have built a model exploiting the ResNet-18 Architecture based on Convolutional Neural Networks with an accuracy of 92% as an aid for classifying the three major types of brain tumor; Gliomas, Meningioma and Pituitary. We believe this study would play a key role in increasing the confidence of the end users (Doctors, Surgeons) when diagnosing the tumor and assisting to create personalized treatments, which may lead to increased life expectancy of the patient.

## 8. **References :**

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7256397/> - Transfer Learning Using Convolutional Neural Network Architectures for Brain Tumor Classification from MRI Images
2. [https://d2l.ai/chapter\\_convolutional-modern/resnet.html](https://d2l.ai/chapter_convolutional-modern/resnet.html)
3. <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>
4. <https://towardsdatascience.com/building-a-brain-tumor-classification-app-e9a0eb9f068>
5. <https://medium.com/plotly/introducing-dash-5ecf7191b503>
6. Learning OpenCV 3 Computer Vision with Python Second Edition; Minichino Joe, Howse Joseph, (2015)