Capstone Project

Kanchi Tank

AWS Machine Learning Engineer Nanodegree

July 19, 2022

Brain Tumor Detection Using CNN

Definition

Project Overview

Advancement in technology has resulted in the generation of a prodigious amount of data from everywhere. Due to the increasing amounts of electronic data in the healthcare, life sciences, and bioscience industry, medical doctors and physicians are facing problems in analyzing the data using traditional diagnosing systems. Nevertheless, machine learning and deep learning techniques have aided doctors and experts in detecting deadly diseases in their early stages.

The aim of this project is to build a Brain Tumor detection model using a Convolutional Neural Network using the Brain MRI Images Dataset.

Problem Statement

A brain tumor is the growth of abnormal cells in the brain tissues [4]. It can be cancerous and non-cancerous. The brain tumor is the most dangerous disease and can be diagnosed efficiently and reliably with the help of technology. We can use automated techniques on MRI Images for accurate detection of brain tumors. The brain is one of the most complicated organs inside the human body. It works with a plethora of cells. Myriads of approaches for the efficient diagnosis of brain tumors have been proposed by numerous practitioners and researchers for adequate tumor detection [8].

The goal of this project is to build a Convolutional Neural Network model to accurately predict brain tumors from brain images.

Evaluation Metrics

1. Accuracy
   It is a metric that illustrates how the model functions across all classes. It is a usual metric for binary classifiers and is useful when all classes are of equal significance. It is calculated as the ratio between the number of correctly classified data instances to the total number of data instances [1].

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$$

2. F1 Score

   This metric is elucidated as the harmonic mean of precision and recall. It becomes 1 only when precision and recall are both 1. Similarly, it becomes high when both precision and recall are high [5].

$$\text{F1 Score} = \frac{True\ Positive}{True\ Positive + \frac{1}{2}(False\ Positive + False\ Negative)}$$

3. AUC-ROC Curve

   The Receiver Operator Characteristic (ROC) curve metric is mainly for binary classification problems. It is a probability curve that plots True Positive Rate against False Positive Rate at various threshold values.

   The Area Under the Curve (AUC) measures the 2-dimensional area present underneath the entire ROC curve [2].

Analysis

Data Exploration

The dataset that is used in this project is Brain MRI Images for Brain Tumor Detection. This dataset was obtained from Kaggle. The dataset encompasses two folders named "yes" and "no." These folders collectively contain 253 brain MRI images. The images are grayscale and come in different sizes, i.e., they have distinct widths, heights, and the number of channels. The folder "yes" consists of 155 tumorous brain MRI images, and the folder "no" consists of 98 non-tumorous brain MRI images.

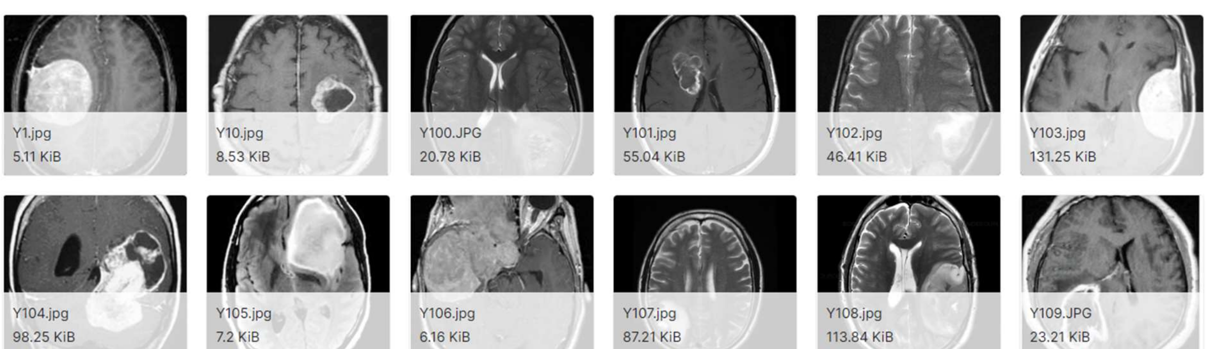

Fig. 1. Tumorous Brain MRI Images

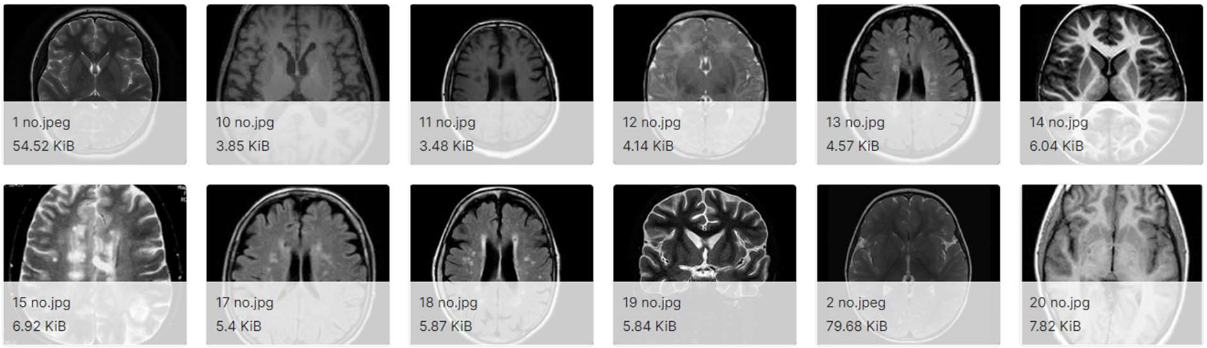

Fig. 2. Non-tumorous Brain MRI Images

Exploratory Data Analysis

The Brain MRI Images for Brain Tumor Detection dataset consists of two classes of images: tumorous and non-tumorous. To visualize the count of data points/ images, we have plotted a bar plot for it. Fig. 3 shows the count of brain MRI images in the original dataset.

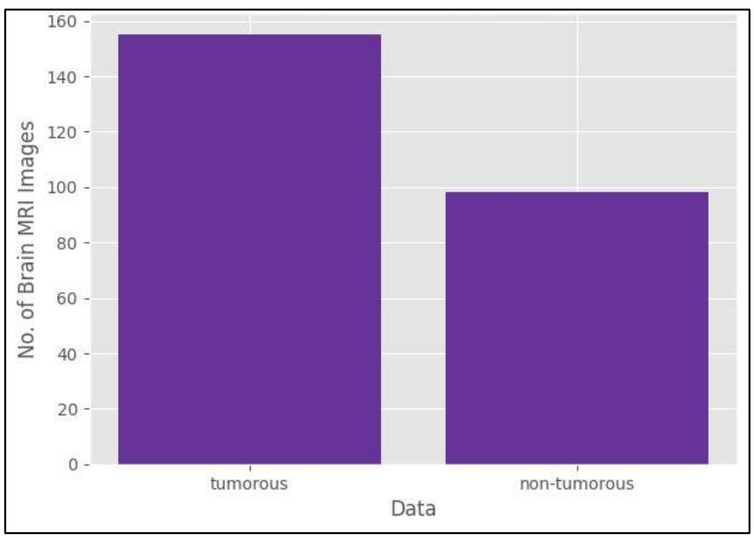

Fig. 3. Count of Brain MRI Images

The plot shows that the data is not balanced, and the amount of data is quite less. Hence, we have performed data augmentation to increase the amount of data and balance it. Fig. 4 shows the count of the brain MRI images after data augmentation. We can infer from the plot that the new amount of data is 2065 images, where 1085 images belong to the tumorous class and 980 images belong to the non-tumorous class.

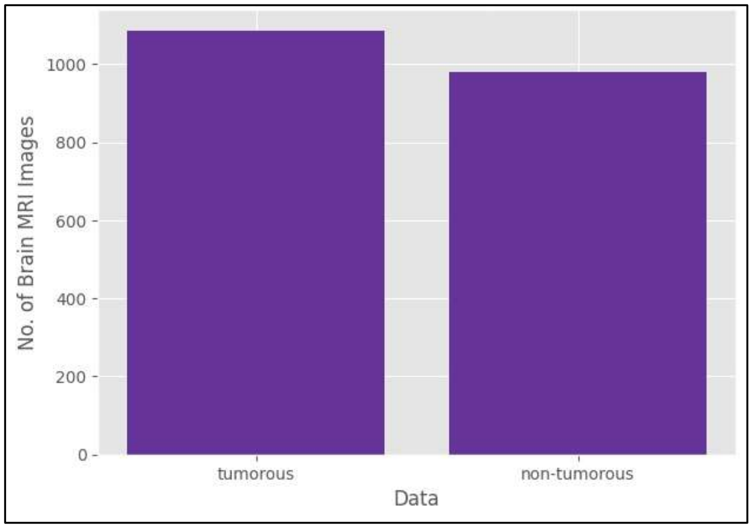

Fig. 4. Count of Brain MRI Images after Data Augmentation

Algorithms and Techniques

A Convolutional Neural Network (CNN) model is used in this problem. CNN is a robust algorithm for image and video processing. It is currently one of the best algorithms for the automated processing of images. Applications involving object detection, image recognition, image segmentation, etc., are some of the tasks that CNN specializes in.

Convolutional Neural Network aspires to process data with a grid structure [6, 7]. CNN consists of three-layer types: the Convolutional Layer, Pooling Layer, and Fully Connected Layer. The convolutional layer is the most essential and the primary layer in the CNN architecture [7]. Next, to reduce the dimensionality of the feature map, the pooling layer comes into play. Lastly, the outputs from the final pooling layer or the convolutional layer are fed as the input to the fully connected layer after flattening.

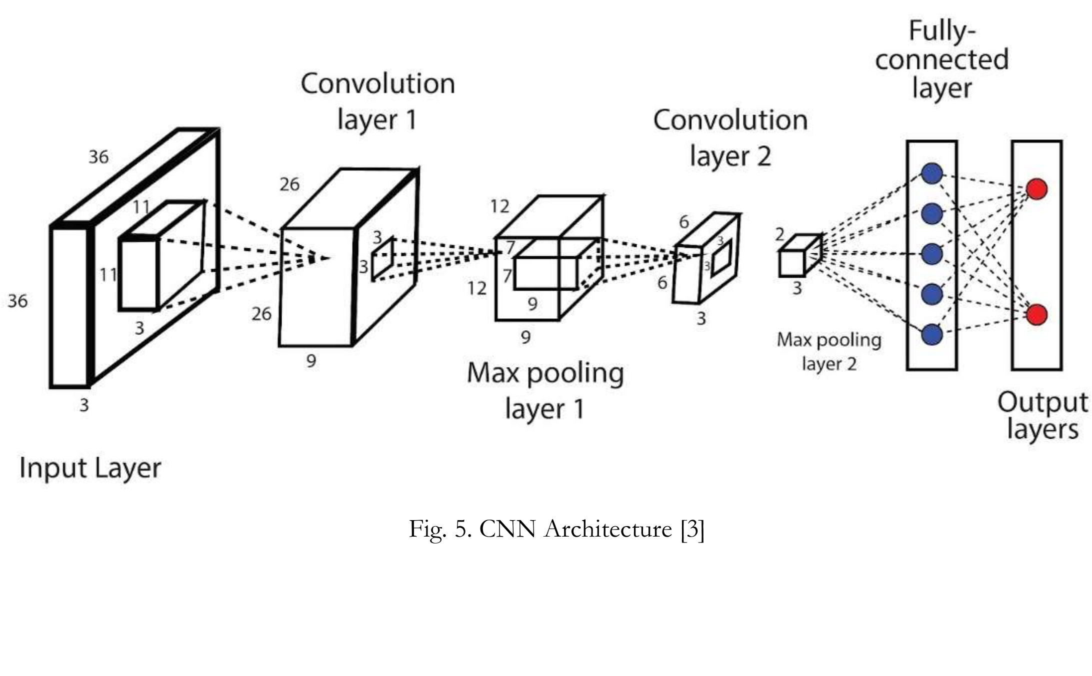

Fig. 5. CNN Architecture [3]

Benchmark

In this project, to classify brain MRI images into tumorous and non-tumorous, we have used VGG19 CNN architecture. It is a 19-layers deep convolutional neural network [9]. VGG19 is an advanced convolutional neural network with pre-trained layers. It is very deep and has been trained on millions of sundry images with intricate classification tasks. It has a substantial understanding of what defines an image in terms of structure, color, and shape [11]. Fig. 6 shows the architecture of VGG19.

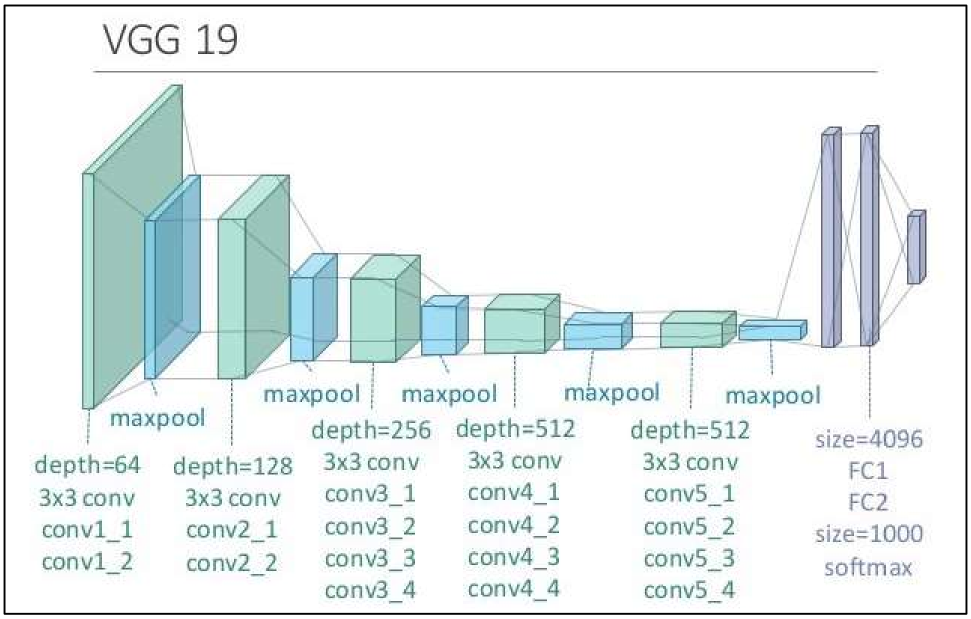

Fig. 6. VGG19 Architecture [10]

Transfer learning has been performed in this project. It is a machine learning technique in which we reuse a pre-trained model as the commencing point for a model on a new task [12]. We have built 3 models using VGG19 CNN architecture and fine-tuned them using transfer learning.

Model 1 was a frozen VGG19 CNN model that gave a test accuracy of 71.29% and training accuracy of 68.06%. Model 2 was built from incremental unfreezing and fine-tuning using transfer learning. This time the test accuracy of this model increased to 73.13% and training accuracy increased to 79.35%. Finally, model 3 (Fine-tuned CNN model) was built by unfreezing and fine-tuning the entire network. This model performed the best by giving a test accuracy of 98.06% and a training accuracy of 100%.

Methodology

Data Augmentation

Since the number of images in the dataset is less, we have performed data augmentation. It is a set of techniques to artificially expand the amount of data by spawning new data points from existing data. It includes making minor modifications to data or using deep learning methods to render new data points.

Moreover, data augmentation can also solve the data imbalance issue. In our case, 61% (155 images) belong to the tumorous class, and 39% (98 images) belong to the non-tumorous class, creating a data imbalance. So, to solve the data imbalance issue, we have generated nine new images for every image belonging to the "no" class, and six images for every image belonging to the "yes" class.

Data Preprocessing

The preprocessing of data consists of the following steps:

1. To compute the extreme points along the contour of the brain in the image, apply a cropping strategy using OpenCV. Crop the part of the image that represents only the brain.
2. Resize the images as they come in distinct sizes from the dataset. Resize all our images to (240, 240, 3) to feed them as input to the Convolutional Neural Network.
3. Apply normalization techniques as we want to scale the pixel values in the range 0-1.
4. Append the label to y and the image to X.
5. Shuffle X and y as the data is ordered, meaning the first part of the array belongs to one class and the second part belongs to the other, and we don't desire that.
6. Split X and y into training, validation, and test set by 70%, 15%, and 15%.

Implementation

Firstly, the brain MRI images were collected from Kaggle, and data preparation was done by importing necessary libraries and packages. This was followed by the exploratory data analysis where we found out the count of tumorous and non-tumorous images in the dataset. Next, data augmentation was performed to increase the number of images in the dataset. After creating a new directory of augmented data, data preprocessing was done. Data preprocessing involved image cropping, and image loading. Next, data was then split into training, test, and validation sets, following the 70:15:15 ratio. We got 759 tumorous training images, 686 non-tumorous training images, 163 tumorous valid images, 147 non-tumorous valid images, 163 tumorous testing images, and 147 non-tumorous testing images. This was followed by the model building where we built 3 models of VGG19 architecture using transfer learning. Evaluation metrics were calculated after each model build. Finally, predictions were made using the best model, i.e., model 3 (fine-tuned CNN).

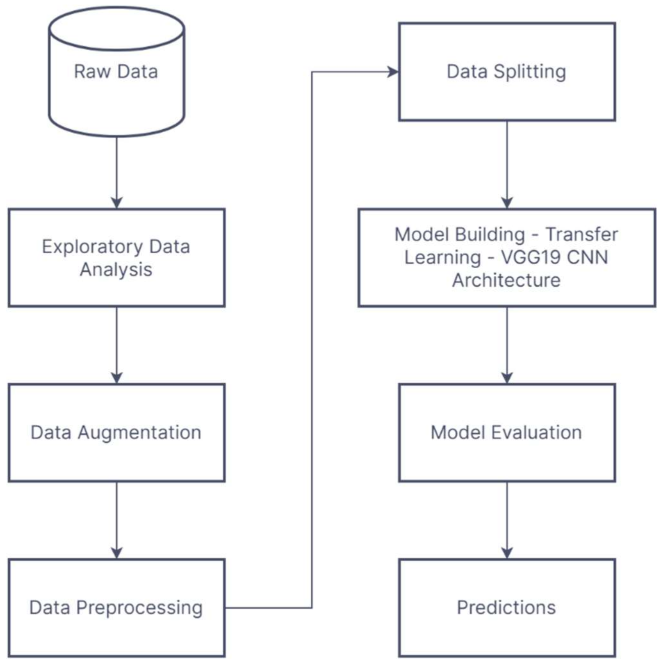

Fig. 7. Methodology

Results

Model Evaluation and Validation

Model 1: Frozen CNN

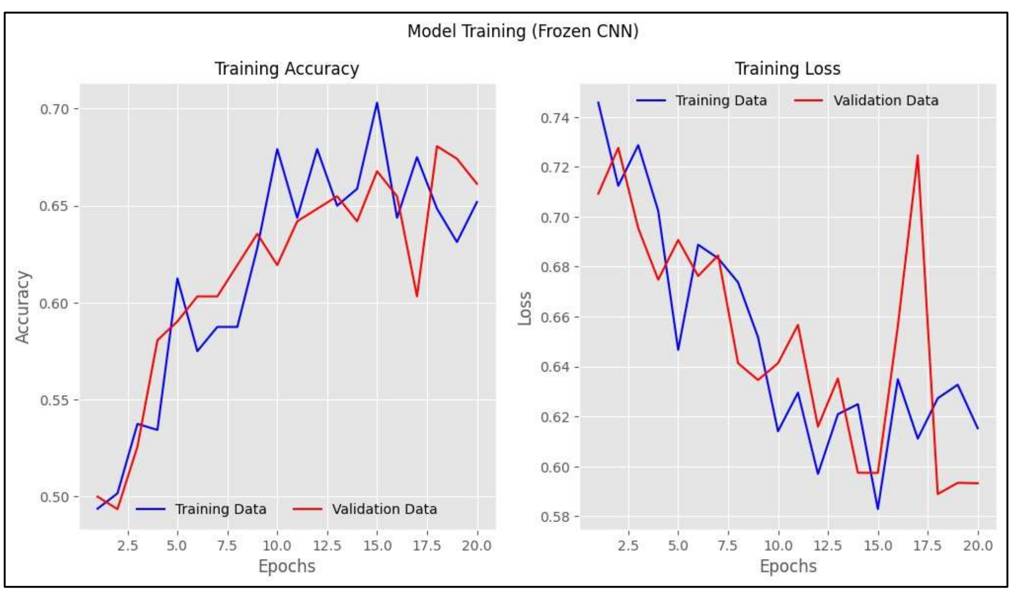

Fig. 8. Training Accuracy and Loss plot for model 1

```
Validation loss:       0.5933
Validation accuracy: 0.6613
Test loss:             0.5527
Test accuracy:         0.7129
```

Fig. 9. Test, validation accuracy and loss of model 1

```
F1 Score: 0.7101
```

Fig. 10. F1 Score of model 1

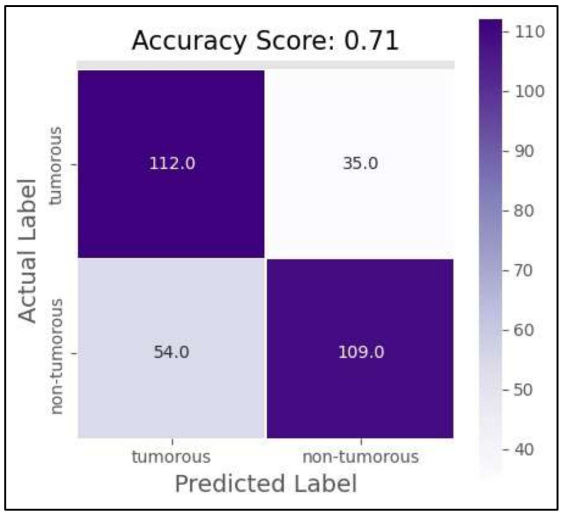

Fig. 11. Confusion Matrix of model 1

Model 2: Incremental unfreezing and fine-tuning

```
Validation loss:       0.47262707352638245
Validation accuracy: 0.7354838848114014
Test loss:             0.48020148277282715
Test accuracy:         0.7612903118133545
```

Fig. 12. Test, validation accuracy and loss of model 2

```
F1 Score: 0.8063
```

Fig. 13. F1 Score of model 2

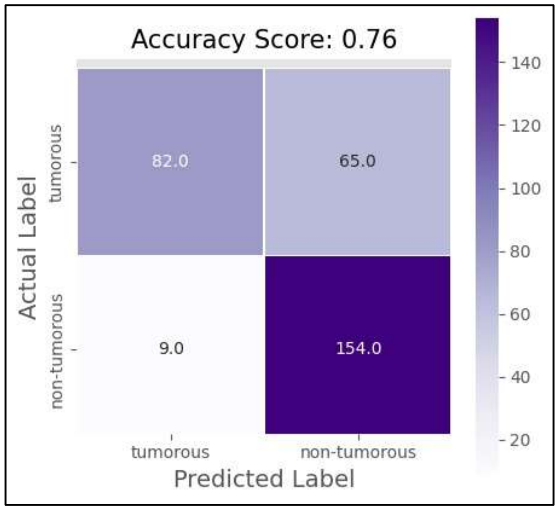

Fig. 14. Confusion Matrix of model 2

Model 3: Unfreezing and fine-tuning the entire network

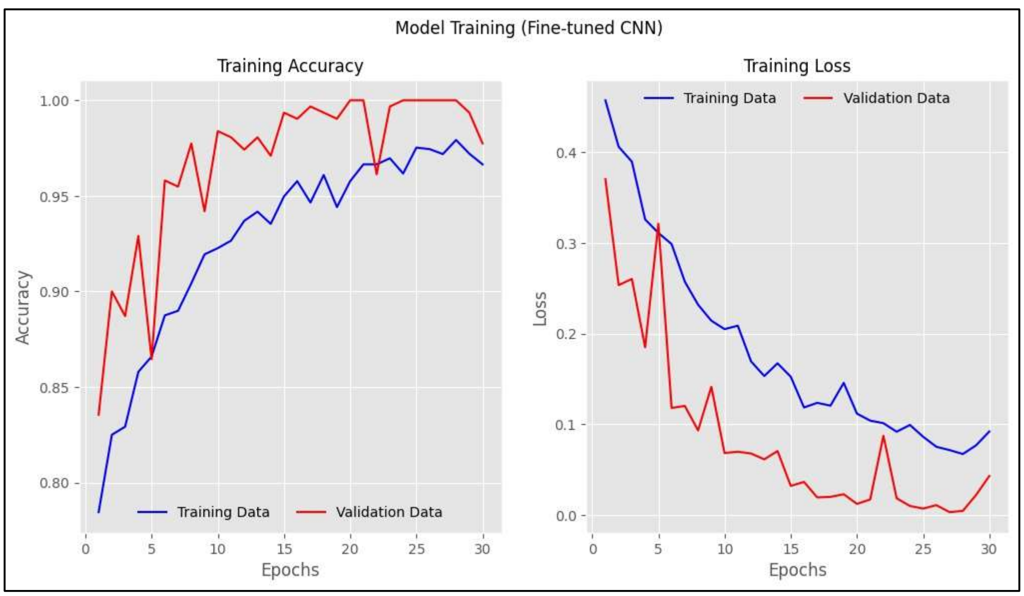

Fig. 15. Training Accuracy and Loss plot for model 3

```
Validation loss     :0.0430
Validation accuracy :0.9774
Test loss           :0.0553
Test accuracy       :0.9806
```

Fig. 16. Test, validation accuracy and loss of model 3

F1 Score: 0.9812

Fig. 17. F1 Score of model 3

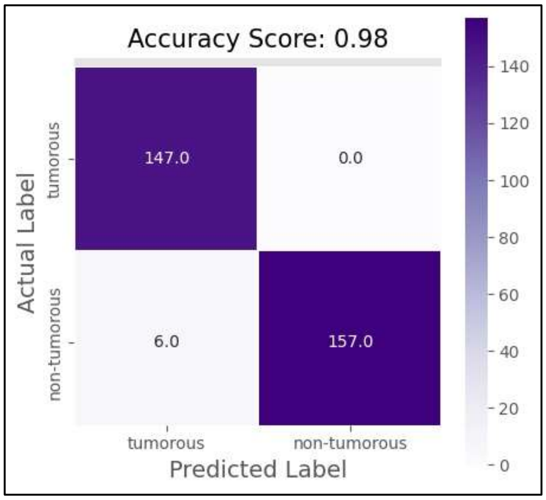

Fig. 18. Confusion Matrix of model 3

AUC-ROC Curve

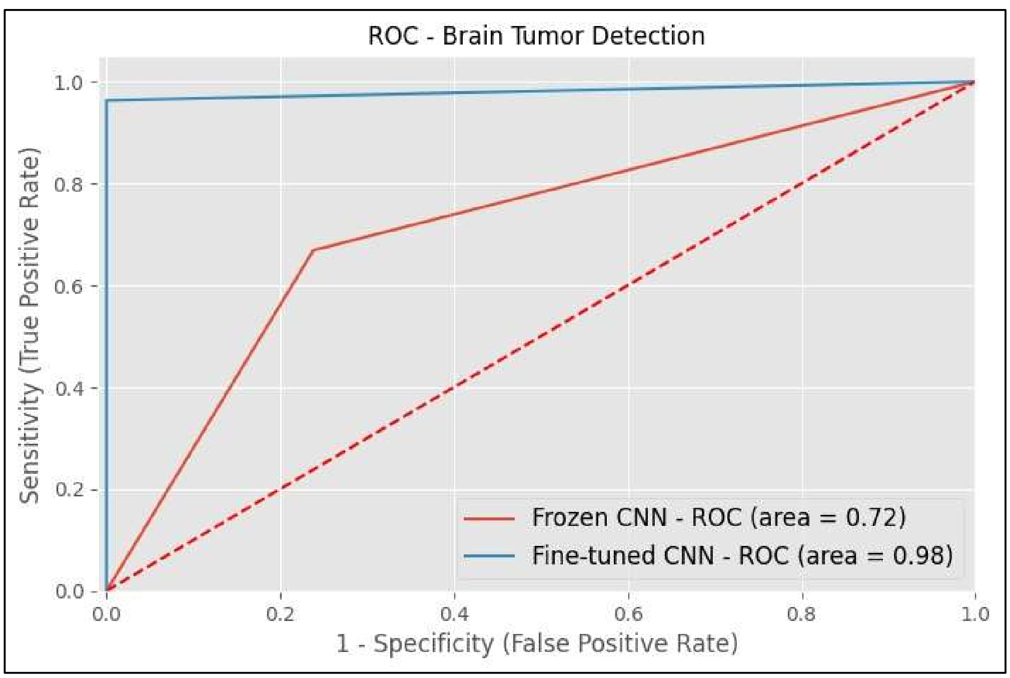

Fig. 19. AUC-ROC curve for model 1 and 3

Performance Evaluation

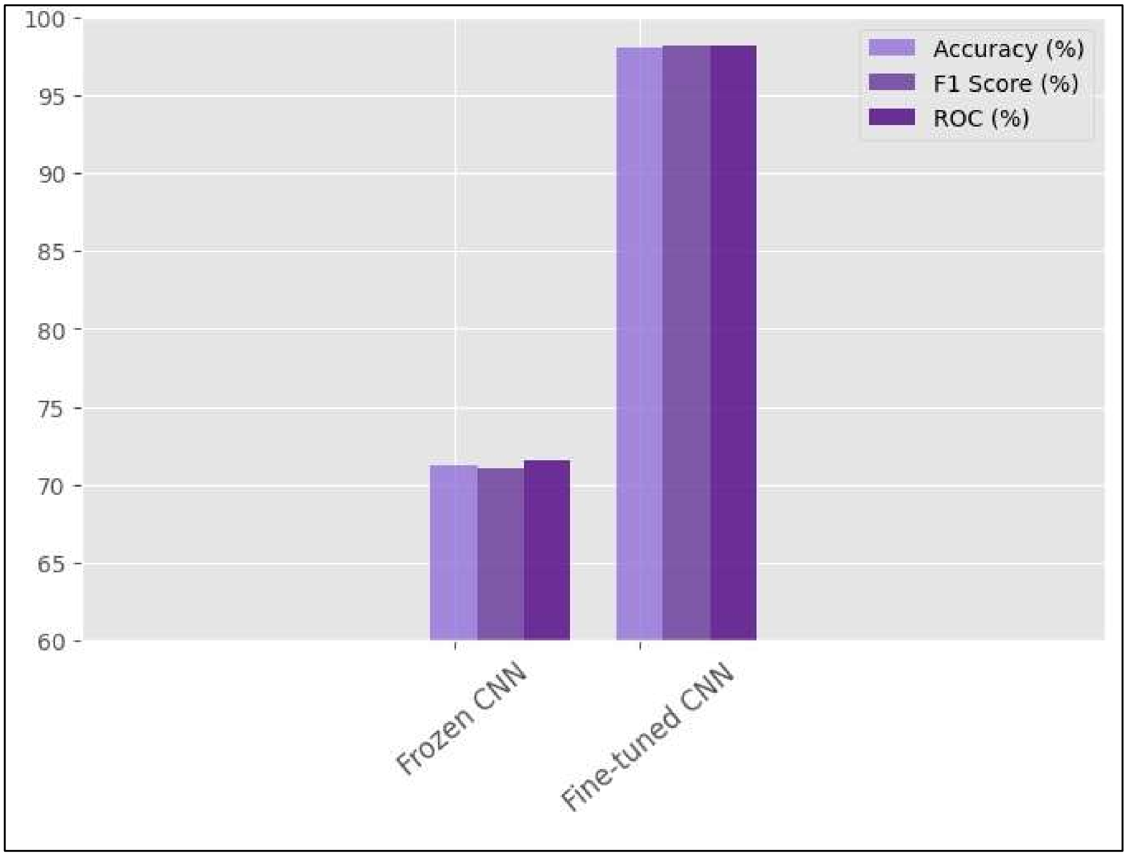

Fig. 20. Performance Evaluation Bar Plot for model 1 and 3

Prediction Results

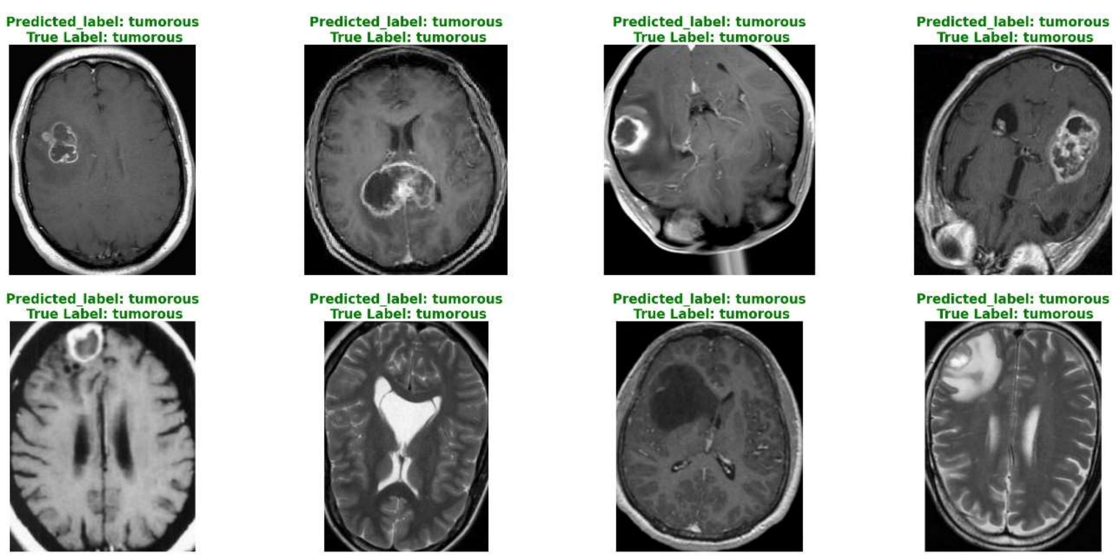

Fig. 21. Prediction Results 1

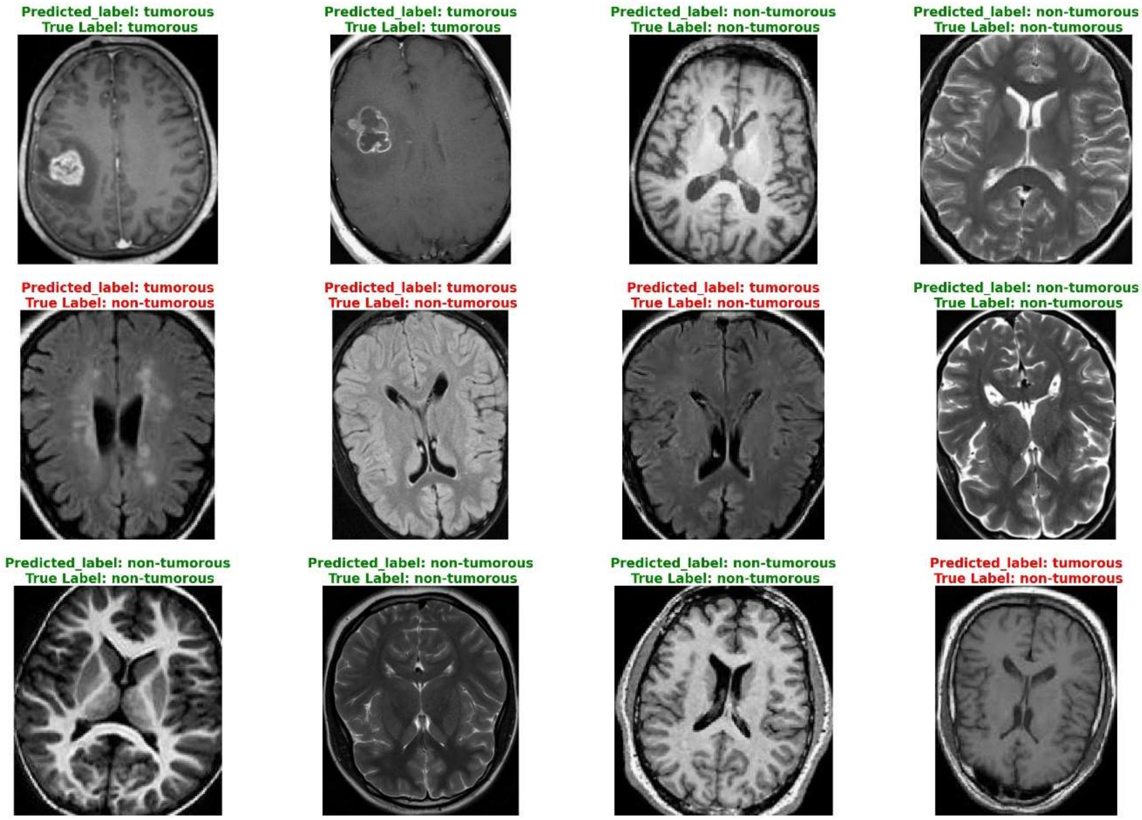

Fig. 22. Prediction Results 2

Conclusion

Disease prediction has the potential to become a stakeholder's asset. These include health insurance companies, government, etc. It identifies patients at risk of hazardous disease, which then helps the physicians and clinicians take suitable actions to avoid or minimize the risk. In this project, we have built a VGG19 CNN model using transfer learning and the best model was built after fine-tuning. Our model gave a test accuracy of 98.06% and a training accuracy of 100%. The F1 Score of this model is 98.12% and the AUC is 98.16%. This model can be deployed in a web application for the prediction of brain tumors. This web app will be extremely useful for doctors and physicians in accurately predicting brain tumors from MRI images of patients.

References

[1] Accuracy, Precision, and Recall in Deep Learning | Paperspace Blog. (2022). Retrieved 2 July 2022, from https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy/#:~:text=Accuracy%20is%20a%20metric%20that,the%20total%20number%20of%20predictions

[2] AUC-ROC Curve in Machine Learning Clearly Explained - Analytics Vidhya. (2022). Retrieved 1 July 2022, from https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/

[3] Balodi, T. (2022). Convolutional Neural Network with Python Code Explanation | Convolutional Layer | Max Pooling in CNN. Retrieved 1 July 2022, from https://www.analyticssteps.com/blogs/convolutional-neural-network-cnn-graphical-visualization-code-explanation

[4] Brain Tumor: Symptoms, Signs & Causes. (2022). Retrieved 2 July 2022, from https://my.clevelandclinic.org/health/diseases/6149-brain-cancer-brain-tumor

[5] Confusion Matrix, Accuracy, Precision, Recall, F1 Score. (2022). Retrieved 30 June 2022, from https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd#:~:text=F1%20Score%20becomes%201%20only,0.857%20%2B%200.75)%20%3D%200.799

[6] Danukusumo, K., Pranowo, & Maslim, M. (2017). Indonesia ancient temple classification using convolutional neural network. 2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC). doi: 10.1109/iccerec.2017.8226709

[7] Febrianto, D., Soesanti, I., & Nugroho, H. (2020). Convolutional Neural Network for Brain Tumor Detection. IOP Conference Series: Materials Science and Engineering, 771(1), 012031. doi: 10.1088/1757-899x/771/1/012031

[8] Kumar, S., Dhir, R., & Chaurasia, N. (2021). Brain Tumor Detection Analysis Using CNN: A Review. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). doi: 10.1109/icais50930.2021.9395920

[9] VGG-19 convolutional neural network. (2022). Retrieved 1 July 2022, from https://www.mathworks.com/help/deeplearning/ref/vgg19.html;jsessionid=61f4e9473920165c667ac9639843#:~:text=VGG%2D19%20is%20a%20convolutional,%2C%20pencil%2C%20and%20many%20animals

[10] Bansal, S. (2022). CNN Architectures: VGG, ResNet, Inception + TL. Retrieved 1 July 2022, from https://www.kaggle.com/code/shivamb/cnn-architectures-vgg-resnet-inception-tl

[11] Bardhi, M. (2022). Image Detection Using the VGG-19 Convolutional Neural Network. Retrieved 2 July 2022, from https://medium.com/mlearning-ai/image-detection-using-convolutional-neural-networks-89c9e21fffa3

[12] Baheti, P. (2022). A Newbie-Friendly Guide to Transfer Learning. Retrieved 1 July 2022, from https://www.v7labs.com/blog/transfer-learning-guide