ResNet-50 based deep neural network using transfer learning for brain tumor classification

Conference Paper *in* AIP Conference Proceedings · May 2022

DOI: 10.1063/5.0082328

5 authors, including:

Madona B Sahaai
VISTAS

D. Ravikumar
DMI College of Engineering (Autonomous)

ResNet-50 based Deep Neural Network using Transfer Learning for Brain Tumor Classification

Madona B Sahaai[1,a)], G R Jothilakshmi[2,b)], Raghavendra Prasath[3,c)], Saurav Singh[4,d)]

[1]Research Scholar, [2]Associate Professor, [3,4]UG scholars
Vels Institute of Science Technology & Advanced Studies

[a)]*madona.se@velsuniv.ac.in*, [b)]*jothi.se@velsuniv.ac.in*, [c)]*arps10@icloud.com*, [d)]*singhsaurav1927@gmail.com*

***Abstract* –** Brain tumour is one of the most complicated diseases to treat in modern medicine. In the early stages of tumour development, the radiologist's primary concern is often an accurate and efficient study. Deep Learning has become a great tool for doctors and scientists to act decisively and on time with tumor patients. A training model that has accomplish considerable result in image detection and classification is the Deep Residual Network (ResNet) utilizing CNNs. The advancement of deep learning will assist radiologists in tumor diagnostics without the use of harmful procedures. With better understanding of MRI images, as well as increase in training speeds and accuracy, deep learning can open new doors for the medical research community. In this model, an accuracy of 95.3% is achieved across various classes of brain tumor datasets. We study the outcomes of multi class classification of brain tumour using Transfer Learning utilising pre-trained ResNet50 model using CNN architecture in this paper.

Keywords – deep neural networks (DNN), convolutional neural network, auto-encoders, brain-tumors, machine learning

INTRODUCTION

In medical image processing, Brain Tumor Segmentation is a critical duty. Early detection of brain cancers improves treatment options and enhances the patient's survival percentage. In the field of medicine, brain image segmentation is critical in surgical preparation and treatment planning. Image segmentation is the most important procedure for assisting in the identification, and visualization of regions of interest in any medical image. Deep convolutional networks are incredibly good at extracting features . Secondary tumors can be made up of a wide variety of tissue types, depending on the primary tumor site. Its ability to successfully detect and segment a range of brain tumors is demonstrated by its usage of many datasets with diverse tumor sizes, intensities, and locations . Deep convolution networks based on the U type ResNet may also obtain superior results for the tumor region. In terms of overall accuracy, sensitivity, and specificity, SE-ResNet-101, a deep CNN-based architectural technique, outperformed the other two recent competing brain tumor classification systems. Brain tumours are the most deadly illness, with a life expectancy of only a few months in the most advanced stages. Misdiagnosis of brain tumours will result in ineffective medical intervention, lowering patients' chances of survival. The correct identification of a tumour is critical for developing an effective treatment strategy to cure and prolong the lives of individuals with brain tumours. The effectiveness of the tumor prediction is found to be more accurate using the CNN process. The biopsy test has a number of risks, includes infection from tumour and brain haemorrhage, seizures, severe migraine, stroke, coma, and even death. The standard practice for diagnosing the severity of a brain tumour is by classification, which is based on a biopsy exam. Biopsies are used to classify brain tumors and are not normally performed prior to conclusive brain surgery. As a result, technologies for detecting and predicting the grade of cancers based on MRI data have become required. CNNs and computer-aided tumour detection systems have offered success stories and made significant advances in the field of medicine. In comparison to previous standards of neural network layers, deep convolutional layers extract crucial and robust characteristics from the input space automatically. [10] In this paper, a ppattern Recognition and size prediction of mammogram images is done using advanced image processing methodologies. [11] A detailed Study of hardware implementation of the microcalcification detection using embedded systems is done in this paper. [12] The brain is the most complex organ of the human body, consisting of nerve cells and tissues that regulate the most basic functions of the body such as breathing, muscle coordination, and our senses

PROPOSED SYSTEM

The proposed system defines the processes of tumor classification from the MRI images. They are described in the figure below:

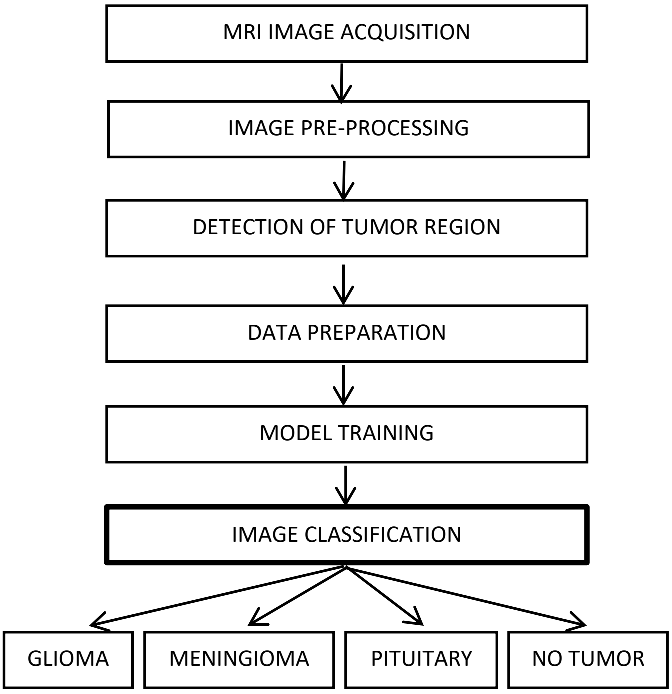


FIGURE 1. Proposed System Block Diagram

CHOICE OF BRAIN IMAGES

All the MRI Images used in this study were obtained Jun Cheng's Brain Tumor Dataset which employs state-of-the-art methods for acquiring multimodal MRI scans. This brain tumour dataset includes 3064 T1-weighted contrast-enhanced images from 233 individuals with three different types of brain tumours: meningioma (708 slices), glioma (1426 slices), and pituitary tumour (930 slices). We divided the whole dataset into four subsets and archived them in four.zip files, each having 766 slices, according to the repository's file size constraint. Also included are the 5-fold cross-validation indices. This information is presented in MATLAB data format (.mat file). A struct containing the following fields for an image is stored in each file: cjdata.label: 1 for meningioma, 2 for glioma, 3 for pituitary tumour, cjdata.label: 1 for meningioma, 2 for glioma, 3 for pituitary tumour, cjdata.label: 1 for meningioma, cjdata.image: image data, cjdata.tumor: tumour data Border: a vector that stores the coordinates of discrete spots along the tumor's edge. As a result, we may utilise it to make a binary picture of the tumour mask. Cjdata tumour Mask is a binary picture with 1s representing the tumour area.

The dataset was acquired from: https://figshare.com/articles/dataset/brain_tumor_dataset/1512427?file=7953679 , https://github.com/chengjun583/brainTumorRetrieval

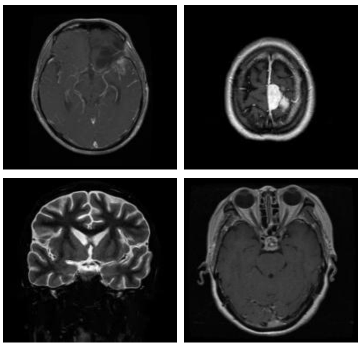


FIGURE 2. Input images of MRI Scans

IMAGE PREPROCESSING

The aim of the pre-processing stage is to optimise image data by suppressing unnecessary distortions and improving certain essential image features for subsequent processing. Image enhancement is a pre-processing method used to transform an initial image into a more desirable version. The basic goal of medical imaging analysis is to clean and increase the contrast of MRI images. The MRI pictures were created using several modalities, which resulted in artefacts and incorrect intensity levels.
The challenges in medical image processing include fusion of multimodality images, classification of image highlights, image reclamation and accurate section of highlights.

To create a huge amount of data for CNN architectures and avoid over-fitting, data augmentation techniques like as rotation and flipping are used. The following steps are followed in the pre-processing stage:

The original image is paired with the sharpened images for further enhancement.When MRI scanned images are stored in a database, they are converted to grey scale image sizes of 255 x 255. Since these photographs have been processed to eliminate noise, the visual quality of the noise images has been affected. The high pass filter for edge detection and sharpening is responsible for the image's high resolution and lack of noise.

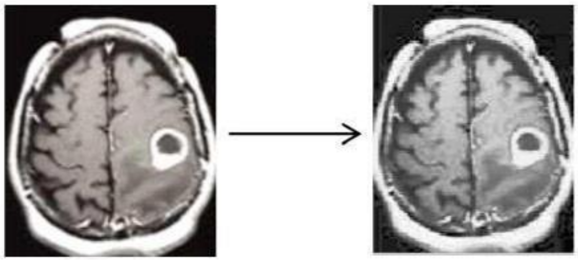


FIGURE 3. Histogram Equalized Image

Histogram equalisation improves the contrast of a grey-scaled image by mapping new pixel values to the histogram. It's a technique for broadening the dynamic range of an image's histogram. Histogram balancing assigns force estimations to pixels in the information Image, resulting in a consistent distribution of forces in the yield Image and increased image separation.
Because of the noise in the photos, ResNets are unable to distinguish them. Binarization will be used to eliminate the background noise created in photographs. Each of the three channels (R, G, and B) in a colour picture has a value ranging from 0 to 255. One of the most essential elements of binarization is the conversion of grey-scale photos to black and white (0 to 1) pictures. Binarization smoothes and simplifies the outlines of various items in the image. This function extraction aids the model's learning process.

$$(x, y) = \begin{cases} 1 & f(x, y) \geq T \\ 0 & f(x, y) \leq T \end{cases}$$

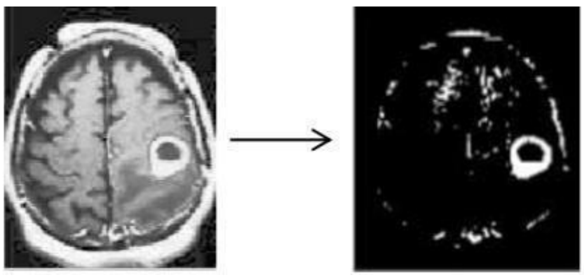


FIGURE 4. Binarized image formed from grey scale

Morphological approaches evaluate a picture using an organising component, which is a little form or arrangement. The organising component may be found in all possible places of the image, and it is contrasted with the pixels in the surrounding neighbourhood. Enlargement, dissolution, opening, and closing are the four basic processes. Only growth and disintegration are used in the suggested work.

DETECTION OF TUMOR

Neural Nets: Deep learning is subdivided into neural networks, which takes inspiration from human brain. They take data and train themselves to spot patterns, and then forecast the results form a new set of data. Neuron networks are made up of layers of neurons. The network's primary processing units are these neurons. The input layer accepts the data, whereas the output layer forecasts the eventual result. Between the visible and hidden layers are the hidden layers, which do the majority of the calculations necessary by our network. Our brain tumour photos have a resolution of 128 by 128 pixels, for a total of 16,384 pixels. Each neuron in the first layer receives a pixel as input.

Transfer Learning: Transfer learning enables neural networks to be trained with substantially less data. Through transfer learning, which are basically transferring the knowledge that a model has learned from a previous task to our current one. The idea is that the two jobs aren't completely disjointed, so we can use the network parameters that the model has learnt over time without having to undertake the training ourselves. Transfer learning has repeatedly enhanced model accuracy while reducing training time with less data as well.

Activation Function: In binary classification, the sigmoid function ranges from 0 to 1, and it is used to forecast probability as an output, whereas the Softmax function is used for multi-class classification. In binary classification using the feed forward technique, the tanh function spans from -1 to 1 and is regarded superior than the sigmoid function. For a non-linear operation, we make use of ReLU (Rectified Linear Unit). The goal of ReLU is to add non-linearity to our ConvNet. Because the data we want our ConvNet to learn in the actual world is non-negative linear numbers. Here the output is $f(x) = \max(0, x)$. In addition to ReLU, additional nonlinear functions like as tanh and sigmoid can be employed. The majority of data scientists choose ReLU since it outperforms the other two in terms of performance.

DATA PREPARATION

We employ a state-of-the-art convolutional neural network, ResNet50, pre-trained on the brain tumor dataset by Jun Cheng, because CNNs have proven to be particularly successful in representation learning, since they extract features using convolutional filters and train the parameters using back-propagation. To suit the input criteria of the ResNet50, the picture was scaled to 224x224 pixels.

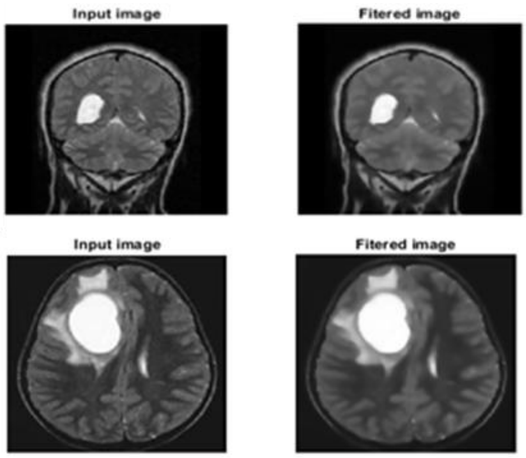


FIGURE 5. ResNet 50 Architecture

TRAINING MODEL FOR CLASSIFICATION

Residual Neural Network (ResNet): ResNet lets you to train hundreds, if not thousands, of layers while still getting outstanding results. Because of its great representational abilities, it has enhanced many computer vision applications outside picture classification, such as object identification and facial identification. According to the universal approximation theorem, a feed forward network with a single layer may reflect any function given enough power. On the other side, the layer might be rather vast, and the network might be subject to data over-fitting. The well-known vanishing gradient problem makes it difficult to train deep networks: when a gradient is back-propagated to older layers, repeated multiplication causes the gradient to become indefinitely tiny. When a result, as the network grows in size, the output gets saturated, if not degraded fast.

There were numerous ways to dealing with the vanishing gradient problem before ResNet, such as putting an auxiliary loss in an intermediate layer as additional monitoring, but none of them seemed to entirely solve the problem. ResNet's central concept is to introduce an identity shortcut relation that skips one or more layers. The authors argue that layer stacking does not reduce network efficiency because we may easily stack identity mappings (a non-functional layer) on top of the existing network and get the same result. This means that the training error of the deeper model need not be greater than that of its shallower counterparts.

To combat this problem a new architecture was introduced that takes advantage of the benefits of shortcut links by connecting all layers directly. Each layer's input consists of the function maps of all previous layers, and its output is transferred to each subsequent layer in this novel architecture. Depth- concatenation is used to combine the function maps. While ResNet has shown to be useful in a wide range of applications, one key drawback is that training a deeper network takes weeks, making it practically impractical in real-world applications. Huang et al. offered a paradoxical strategy of randomly lowering layers during training while experimenting for the whole network to overcome this issue. Because the authors employed the residual block as a building component for their network, when a certain residual block is engaged during testing, the input travels through both the identity shortcut and the weight layers, however when it is deactivated, the input only flows via the identity shortcut. Each layer has a chance of surviving through training and is dropped at random.

The Residual Network was used to tackle the problem of the vanishing/exploding gradient in this design. In this network, we employ a technique known as skip connections. The skip link skips a few steps of preparation and goes straight to the manufacturing. We allow the network to match the residual mapping rather than enabling layers to learn the underlying mapping.
In the place of using H(x) as the primary mapping we utilize F(x):

$$F(x) := H(x) - x, \text{ which gives } H(x) := F(x) + x.$$

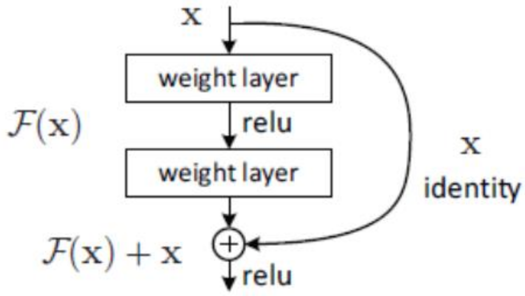


FIGURE 6. ResNet residual learning

The benefit of including this sort of skip link is that any layer that degrades design performance will be skipped by regularisation. As a result, vanishing gradients do not pose a difficulty for training very deep neural networks.

These skip connections, like LSTM, make use of parametric gates. The amount of data that travels via the skip connection is determined by these gates. This design, on the other hand, has not proven to be more accurate than ResNet architecture.
When training an excessive amount of deep CNN, ResNet was proposed to overcome the gradient vanishing and feature map vanishing problems. The ResNet works because identity connections across non-adjacent layers have no effect on the ideal mapping we wish to create.

The identity connection, on the other hand, makes back propagation more fluid since the gradients can transit through an extra short cut channel.

While choosing the right hyper-parameters, it is a priority for transfer learning to utilize a low learning rate to take advantage of the weights of the pre-trained model. This, as well as selecting the optimizer such RMSprop or SGD or Adam will have an impact on the epochs rate to successfully train the model.

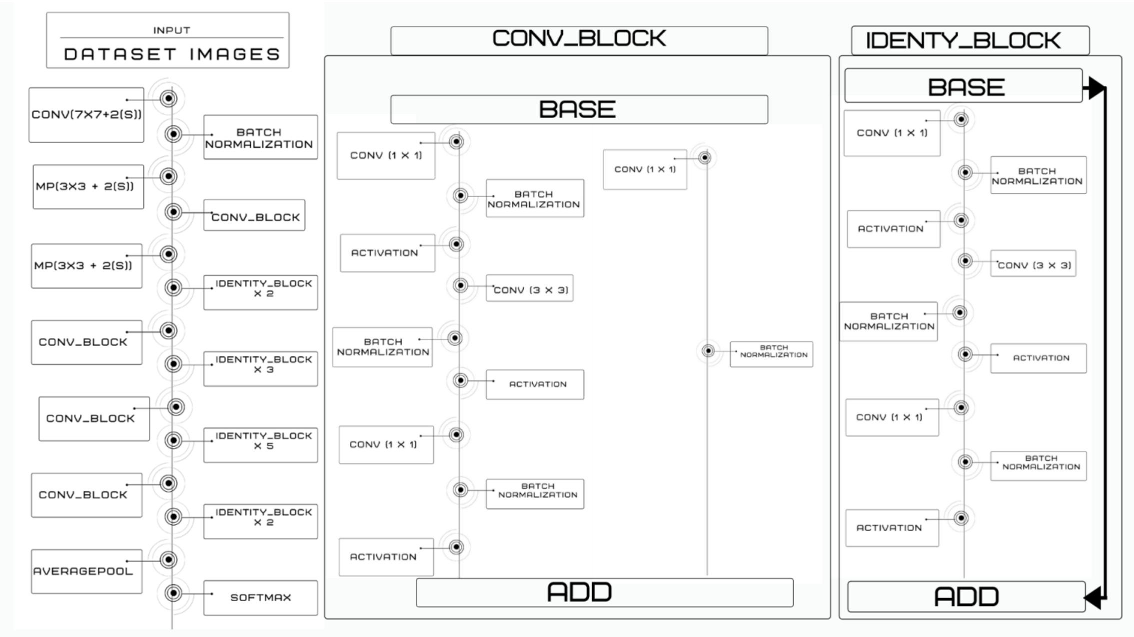


FIGURE 7. ResNet 50 Architecture

EVALUATION MEASUREMENTS

Classification Accuracy is considered to be a final metric for examining the results delivered by a few techniques used in the dataset in the writing. The four metrics used to assess a strategy based on the characteristics of the confusion matrix are accuracy, specificity, precision, recall, and F1-score.

There are four main terms:

True Positives: Situations where we predicted it to be POSITIVE and the output was YES.
True Negatives: Situations where expected it to be negative and the output was also NO.
False Positives: Situations where we expected to get a negative result but received positive.
False Negatives: Circumstances where we expected positive but resulted with negative.

These measurements are determined by the conditions mentioned below:

The most obvious performance indicator is accuracy. The percentage of accurately predicted events divided by the total number of predicted events is known as accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

The number of true positives divided by the total number of true positives + the number of false positives is the definition of precision.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Sensitivity is another term for recall. It's the percentage of the total number of relative relevant occurrences that were found.

$$\text{Recall} = \frac{TP}{TP + FN}$$

The F1 Score is the weighted average of Precision and recall which takes both measures into consideration:

$$\text{F1 Score} = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision + Recall}}$$

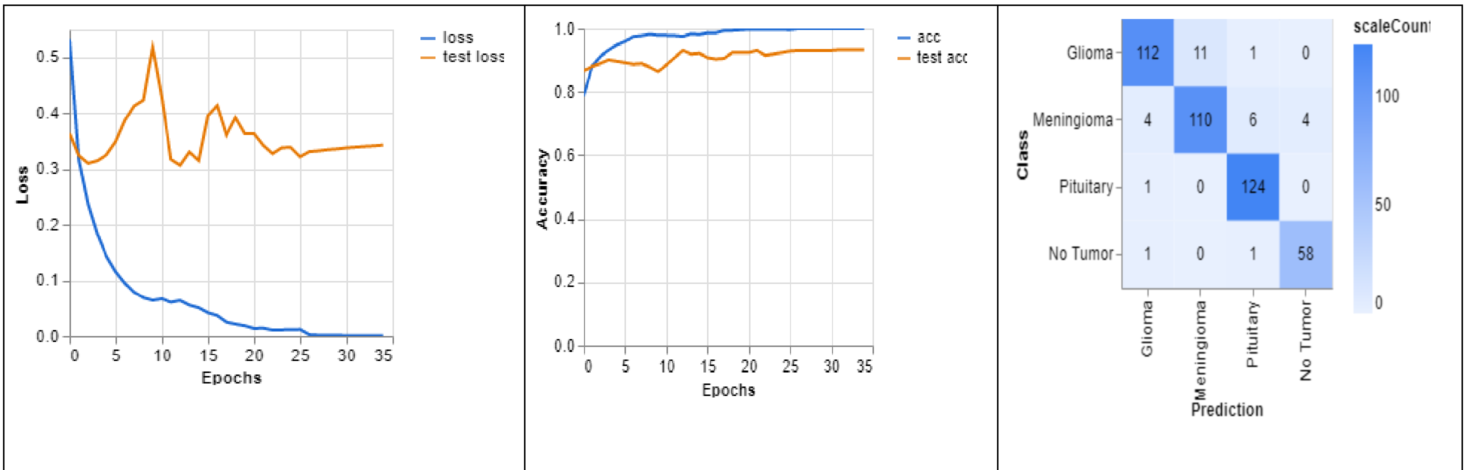
When dealing with an uneven data collection, F1 Score is preferred over accuracy since it accounts for both false positives and false negatives. F-measures are used to balance the ratio of false negatives using a weighting parameter (beta) it is given as

$$\text{F} = \text{P x R} \frac{(1+\beta)^2}{(P+R)\beta^2}$$

Sensitivity, specificity, and error rate are some of the other performance indicators that are employed. They allow us to assess the possibility of over- or under-estimation of tumour sub-regions. The error rate is the percentage of anticipated classes that are categorised wrongly by a decision model.

RESULTS

A) RESULTS FROM TRAINED MODEL



a) b) c)

FIGURE 8. a) Obtained Accuracy Results b) Obtained Loss Results c) Confusion Matrix (Multi Class)

TABLE 1. Comparison of metrics

Evaluation Metrics	DNN with auto-encoders	DNN with ResNet -50
Accuracy	93.2	95.3
Specificity	82.5	87.8
Precision	90.1	93.7
Recall	87.6	92.2
F1- Score	91.3	94.6

TABLE 2 . Validation accuracy by class

Class	Accuracy
Glioma	0.93
Meningioma	0.94
Pituitary	0.98
No Tumor	0.97

CLASSIFICATION USING MOBILE APPLICATION

The screenshots are taken from the mobile application which is developed using the trained model and integrated it into the Flutter framework to allow users to classify brain tumor images on real-time basis just with their smartphone. The link for the mobile app with source code is available on Github. https://www.github.com/qubitrevolution/tumor_classification: Android Application

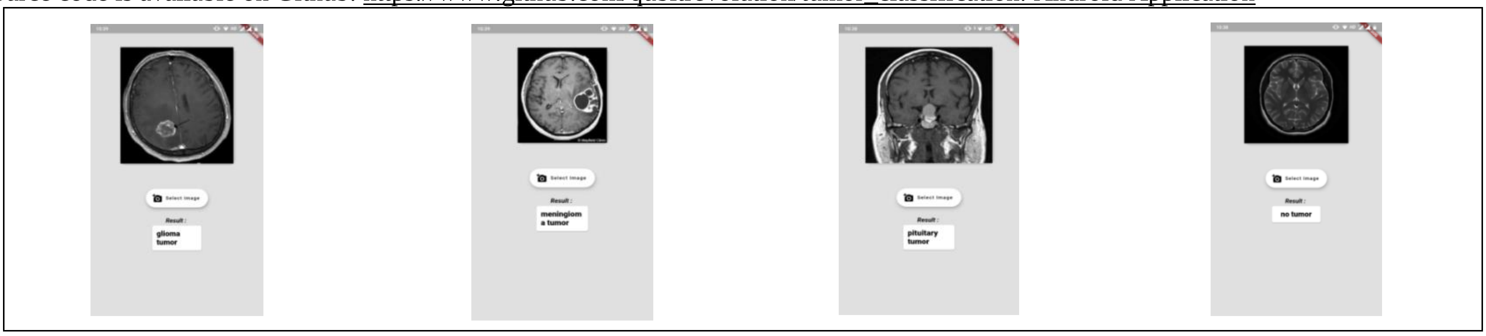


FIGURE 9. Android Application for Brain Tumor Classification available on Github

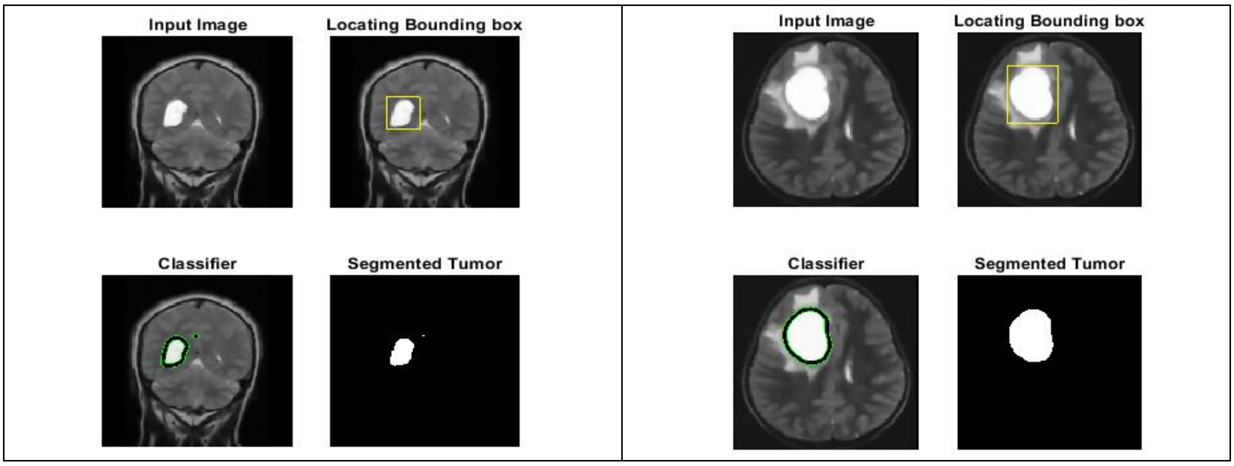


FIGURE 10. Classification of tumor after applying Transfer Learning to ResNet-50

CONCLUSION

In the field of medicine, brain image classification is critical in surgical preparation and treatment planning. We proposed a method for classification and tumor location detection using Deep Residual Networks (ResNet) in this paper.

Our model assists in anticipating a patient's Brian Tumor with greater exactness, particularity, precision, and analysis, all of which are important in the restorative world. In our model, validation accuracy of 95.3% is achieved across various classes of brain tumor datasets. We have compared the outcomes of multi-class classification of brain tumour using Transfer Learning utilising pre-trained ResNet50 model using CNN architecture in this paper.

The training accuracy is 93.5 % and the validation accuracy is 90.0 % without the pre-trained PyTorch model. Although the training accuracy is greater than 90% when using a pre-trained PyTorch model, the total accuracy is poor. As a result, using a pre-trained PyTorch model reduced training and computing time by 50%. When developing a model from scratch rather than utilising a pre-trained model, the chances of over-fitting the dataset are greater. It also has a user-friendly interface for data enhancement. ResNet50 has the greatest overall accuracy as well as the greatest F1 score among the PyTorch models. We must ensure that we have sufficient of testing and validation data to avoid over-fitting. This is solved by Data Augmentation. We also created an app-based user interface that allows doctors to quickly assess the effects of the tumour and provide suitable treatment and advice to the patients.

FUTURE WORK

In the future, the proposed study could be expanded to include various types of modalities for detecting tumors, as well as the optimization method used to improve classification accuracy.

Employing classifier boosting approaches such as utilising larger amount of photos with more data augmentation, fine-tuning hyper parameters, training for a longer duration i.e. using more epochs, adding more applicable layers, and so on, you may improve testing accuracy and computation time.

We may utilise U-Net architecture instead of CNN for more complicated datasets, where the max pooling layers are simply replaced by up sampling ones. We eventually want to employ very large and deep convolutional nets on video sequences, where the temporal structure gives very useful details that are missing or less visible in static pictures.

I. REFERENCES

1. He, Kaiming & Zhang, Xiangyu & Ren, Shaoqing & Sun, Jian. (2016). Deep Residual Learning for Image Recognition. 770-778. 10.1109/CVPR.2016.90.
2. Ali Işın, Cem Direkoğlu, Melike Şah, Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods, Procedia Computer Science, 2016, ISSN 1877-0509
3. Akiba, Takuya & Suzuki, Shuji & Fukuda, Keisuke. (2017). Extremely Large Minibatch SGD: Training ResNet-50 on ImageNet in 15 Minutes.
4. Zhang, Jiachi & Shen, Xiaolei & Zhuo, Tianqi & Zhou, Hong. (2017). Brain Tumor Segmentation Based on Refined Fully Convolutional Neural Networks with A Hierarchical Dice Loss.
5. Ray, S. (2018). Disease Classification within Dermascopic Images Using features extracted by ResNet50 and classification through Deep Forest. *ArXiv, abs/1807.05711*.
6. Akram, M. & Usman, Anam. (2011). Computer aided system for brain tumor detection and segmentation. Proceedings - International Conference on Computer Networks and Information Technology. 10.1109/ICCNIT.2011.6020885.
7. Rathi, V.P. & Palani, S.. (2015). Brain Tumor Detection and Classification Using Deep Learning Classifier on MRI Images. Research Journal of Applied Sciences, Engineering and Technology
8. Roy, Sudipta & Bandyopadhyay, Samir. (2012). Detection and Quantification of Brain Tumor from MRI of Brain and it's Symmetric Analysis. International Journal of Information and Communication Technology Research.
9. G R Jothilakshmi, Arun Raaza & V Rajendran (2018). Pattern Recognition and size prediction of microcalcification based on physical characteristics by using digital mammogram images. Journal of digital imaging
10. G R Jothilakshmi, P Mohanapriya & V K Suvithra. (2018) Study of hardware implementation on size of the microcalcification detection using embedded systems. International Journal of Engineering & Technology
11. G R Jothilakshmi, Arun Raaza & Dr. Y Sreenivasa varma. (2018). A review of characteristic study of micro clacification using son mammogram images. International Journal of Engineering & Technology.
12. Madona B Sahaai & G R Jothilakshmi. (2021). A review and comparative study on an efficient brain tumor prediction by using two pathway CNN methodology with SVM for tumor classification in MRI images. Elementary Education Online.
13. Madona B Sahaai & G R Jothilakshmi & Raghavendra Prasath & Saurav Singh (2021). Brain tumor detection using DNN algorithm. Turkish Journal of Computer and Mathematics Education
14. Dong, Hao & Yang, Guang & Liu, Fangde & Mo, Yike. (2017). Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks.

15. Ghosal, Palash & Nandanwar, Lokesh & Kanchan, Swati & Bhadra, Ashok & Chakraborty, Debashis. (2019). Brain Tumor Classification Using ResNet-101 Based Squeeze and Excitation Deep Neural Network. 1-6. 10.1109/ICACCP.2019.8882973.
16. Gordillo-Castillo, Nelly & Montseny, Eduard & Sobrevilla, Pilar. (2013). State of the art survey on MRI brain tumor segmentation. Magnetic resonance imaging. 31. 10.1016/j.mri.2013.05.002.
17. S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629 (2018)
18. Cheng, Jun (2017): brain tumor dataset. figshare. DOI: https://doi.org/10.6084/m9.figshare.1512427.v5