

Brain Tumor Detection and Segmentation using Deep Learning

Mugdha Deokar

Electronics and Telecommunication
Savitribai Phule Pune University

Varun Godse

Electronics and Telecommunication
Savitribai Phule Pune University

ABSTRACT

Brain tumors are highly fatal and diagnosis can often be time-consuming. Timely attention and proper consultation is required to successfully detect the tumor. However, the detection of brain tumors is a very tedious task. Brain tumors are of numerous types which pose a challenging task of detection and classification as the tumors can be ill-defined with soft tissues. Magnetic Resonance Imaging (MRI), Computerized Tomography (CT) scans, and Ultrasound images are generally used to get the brain images. Brain tumors are common, and because of the associated risk of malignancy and hyperfunction, these tumors have to be examined thoroughly. To automate the process of detection and further segmentation, a robust system is required which can be used to produce accurate results. In this work, this observation is taken into consideration and a technique was proposed that will bridge the gap between diagnosing and detecting tumors. This process will further provide a valuable second opinion to medical professionals. The methodology involves brain tumor detection using transfer learning using Residual Networks (ResNet). Further, segmentation is done to identify the area of the tumor in the MRI scan. This work demonstrates Deep Learning's potential in processing and extracting information from MRI images to provide a non-invasive tool for automated tumor detection and segmentation for clinical applications.

General Terms

Machine Learning, Deep Learning, Computer Vision, Image Processing

Keywords

Deep Learning, Brain Tumor Segmentation, Brain Tumor Detection, MRI, Neural Networks.

1. INTRODUCTION

Tumors are an uncontrolled/atypical growth of abnormal cells in the body. There are multiple types of tumors, Brain tumor is one such type. The human skull is a very rigid space and any abnormal growth in this area causes serious problems for the human body. Depending on the origin, Brain tumors are classified as primary or secondary. A primary tumor originates in the brain which is also called a malignant tumor while a secondary tumor or benign tumor occurs when a cancerous cell from another organ such as lung or kidney travels to the brain. The brain tumor is also classified based on its location - local, regional and distant. The local tumors are limited to one hemisphere of the brain, while the regional tumors cross the brain's midline and spread into the bones, blood vessels, etc. Distant tumors invade the nasal cavity, nasal pharynx, and posterior tumors. Furthermore, tumors are also classified on how they appear radiologically. (Rajasekaran, 2018; Gounder, 2018) [1]

According to [2] Brain tumor accounts for 85 to 90 percent of total Central Nervous System tumors. It is estimated that 18600

(10,500 males and 8,100 females) adults die every year in the US due to brain tumors or related tumors. The 5-year survival rate for brain tumors after their discovery is around 31 percent, which means that 69 percent of people pass away in 5 years. As noted by (Bauer et al. 2013) [3]

Gliomas are the most frequent type of Primary Brain cancer. WHO classifies them into four grades. Higher the grade, the more harmful the tumor can be. Grading helps the patient, doctor, caregivers/family members understand the severity of the tumor. The Grade I and II are classified as low grade tumors which are also known as astrocytoma. Grade III and IV are high-grade Gliomas that lead to the patient's death. Gliomas Grade IV is also called Glioblastoma Multiforme (GBM) and the median survival time for Glioblastoma Multiforme is one year as observed by (Krex et al. 2007) [4], however there are exceptions where the patient survives more than 36 months. Hence, early detection of brain tumors can prove to be fruitful for a patient's life and in improving the treatment possibilities, and a higher gain of survival possibility can be accomplished as noted by (Hossain et al. 2019) [5].

Medical Imaging Modalities which include Magnetic Resonance Imaging (MRI), Computerized Tomography (CT) scans and Ultrasound images are used to get the data for analyzing brain images. However, MRI images are the most widely used technique of getting the images as it produces better quality images of softer tissues in comparison with CT scans. According to (Seetha, 2018; Raja, 2018) [6] Neural Networks (NN) and Support Vector Machine (SVM) were the usually used methods for research purposes, but the development of Deep Learning Networks in recent years have provided a much more efficient way and have made possible more complex networks. Manual segmentation is a time-consuming task and needs to be automated. In order to help medical professionals assess this data, the first step is classifying the image according to the presence of a brain tumor. After a tumor is identified, segmentation will be done to identify exactly which part of the brain is exposed as a brain tumor. As the first step of the proposed model, classification is done using Deep Learning-Convolutional Neural Networks (CNN) using Keras and Tensorflow. Deep Learning has gained a lot of ground in the area of medical imaging owing to the speed in training and the accuracy and results which are obtained by the particular diagnosis. ResNet was used for image classification. As the second part of this model, state-of-the-art Deep Residual UNET (ResUNet) which works significantly well for semantic segmentation is utilized.

2. LITERATURE SURVEY

Multiple experts of Artificial Intelligence have studied vastly the use of deep neural networks for segmentation tasks. The importance of Machine learning is emphasized to facilitate better results. Deep learning architectures like UNET perform the tasks of semantic segmentation in biomedical images.

(S. Grampurohit et al. 2020) [7] Investigated two models; convolutional neural networks (CNN) and VGG-16 architecture. The results from both models were compared. They have used MRI scans for training the model. Decision tree classifier, radial basis functions were also studied during this. The network performance was evaluated on the basis of sensitivity, specificity, and precision. This method was based on combining convolutional neural networks and feature extraction techniques.

(Jha et al. 2019) [8] Proposes a ResUNet ++ architecture which is based on ResUNet, which can be a semantic segmentation neural network that has residual blocks, squeeze and excitation blocks, Atrous Spatial Pyramidal Pooling (ASPP), and spotlight blocks. This architecture has 1 stem block followed by three encoder blocks, ASPP, and three decoder blocks. Each encoder block has two 3×3 convolutional blocks and identity mapping. Further, every convolutional block consists of a ReLU activation layer and a convolutional layer. ResUNet++ was found to have improved the segmentation results when compared to other state-of-the-art methods. The proposed architecture works well with a smaller number of images.

(Ali et al. 2020) [9] Proposed an ensemble of as few as two networks; The author used a 3D CNN and a 3D U-Net network which are both trained on different data sets. In a straightforward technique, the ensemble of two networks gave better results and predictions were accurate when compared to uniform weighting. The goal is to develop an efficient neural network for the successful delineation of tumors into intratumoral classes with improved efficiency and accuracy in comparison to existing methods. It was observed that the proposed model showed improved results in comparison with state-of-the-art models.

(Hossain et al. 2019) [5] Approached this topic in a 7 step process which is skull stripping, filtering, and enhancement, segmentation, tumor detection, morphological operations, feature extraction, and classification. Two different sets of images were considered; Non-Tumor and tumor images. Skull stripping was done as the skull has less significance in brain tumor segmentation. Otsu thresholding method was employed to automatically calculate the threshold values and divide the image into foreground and background which helps in skull stripping. The fuzzy C Means algorithm is widely used in segmentation. It was observed that a 5 - Layer Convolutional Neural Network was introduced which gave better results to apprehend the tumor. In CNN after pooling, flattening is performed which is the process of converting a matrix of the input image into a single column of the vector. They have used ReLU as the activation function as it gives better convergence performance. Two different loss functions namely Adam optimizer and binary cross-entropy are used for compiling and finding the accuracy of the model. Classification of tumors is also achieved using six different classifiers and comparing their performance. SVM gave the most accurate results in this process. All the modalities like T1, T2, FLAIR are processed.

Pereria et al. (2016) [10] Based on their approach on first eliminating the bias field distortion in the pre-processing section using N4ITK (Improved N3 Bias Correction) method however it was not enough so instead the intensity normalization method put forth by Nyul et al. (2000) [11] is more facilitating. In CNN, for initialization; Xavier Initialization was used to achieve good convergence. By this approach, the activations and gradients are maintained at controlled levels. The activation function which is generally ReLU in CNN was substituted with leaky rectifier linear units LReLU as it was found to give better results than the regular rectifier linear units. The Regularization process is followed next to avoid overfitting which is an issue where the data model performs too well on a training dataset and fails on a test dataset. Due to the small size of the training dataset and also to

avoid overfitting, data augmentation was done with two different techniques to better evaluate the model and to compare other techniques of segmentation on the same dataset. This was done by generating a new patch by rotating the original patch in the multiple of 90 degrees.

3. METHODOLOGY

3.1 Dataset

The Brain MRI segmentation is a publicly available dataset consisting of 3929 Brain MRI scans along with the location of the brain tumor. The MRI scans along with manual FLAIR (modality) abnormality segmentation masks are used. All images have FLAIR modality. There are two classes (class 0 and class 1) of the images i.e. whether it has a tumor or not. The MRI scans have corresponding ground truth images with the masks. The images were collected from The Cancer Imaging Archive (TCIA). The MRI images correspond to 110 patients included in The Cancer Genome Atlas (TCGA) lower-grade glioma collection with FLAIR modality i.e. fluid-attenuated inversion recovery sequence and genomic cluster data that was available [12]. The dataset was split into 70% training set, 15 % validation set, and 15% test set. The images are of the dimension $256 \times 256 \times 3$ which is the input to the ResNet model.

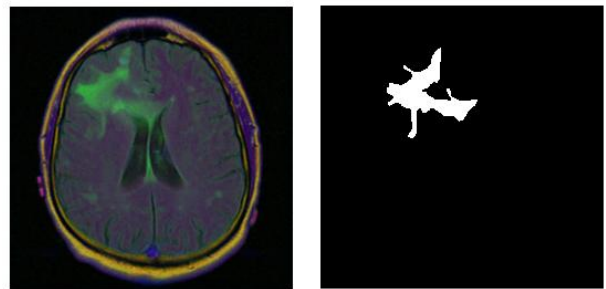


Figure 1: Brain MRI 1 with tumor (With Mask)

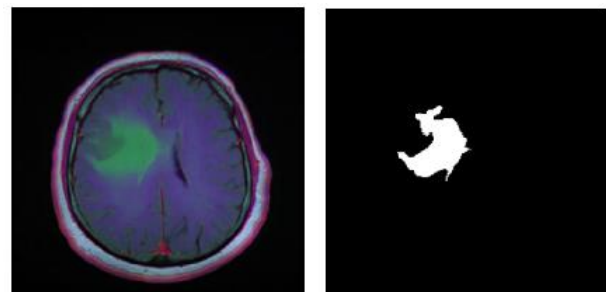


Figure 2: Brain MRI 2 with tumor (With Mask)

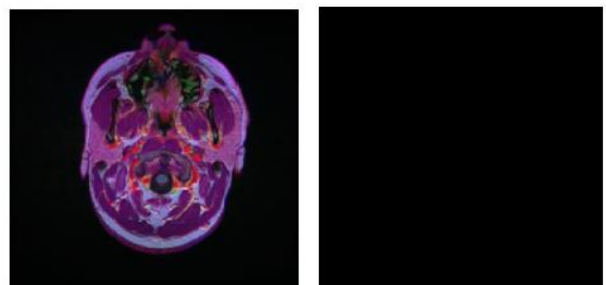


Figure 3: Brain MRI without tumor (No mask)

3.2 Block Diagram

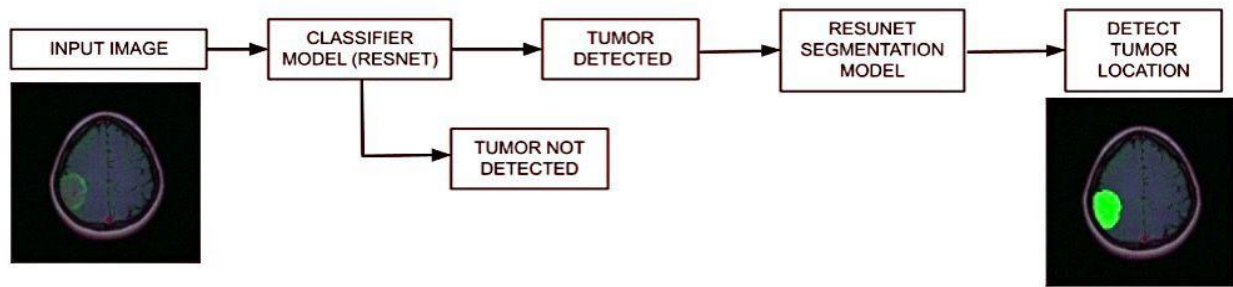


Figure 4: Block Diagram

4. IMPLEMENTATION

Convolutional neural networks in deep learning have gained a lot of importance and are extensively used in medical imaging in recent times and as a method for the non-invasive diagnosis of various diseases. It is a very complex model and requires high computational power. It is quite expensive computationally to use optimized GPUs and to use high-resolution images for the same. Also, traditionally a lot of feature extraction and image augmentation or enhancement techniques are needed to train the model efficiently. This is the reason the concept of Transfer Learning was introduced and is continuously evolving in Deep Learning. In Transfer Learning, an algorithm is pre-trained on a large dataset of images having many classes using heavy GPUs. This pre-trained model can then be used for relevant problem statements at hand by adding new fully connected convolutional layers and providing relevant data to overcome the problem of complex and time-consuming model training.

Other networks like AlexNet, ResNet, and VGG-16 architectures were considered for solving the classification problem. On comparing the three algorithms, in AlexNet, there are 7 layers and 60 million parameters. Its error rate is 15.3% which was higher than the other two algorithms. In VGG-16, the error rate is 7.3% and 16 layers. But, it has 138 million trainable parameters which makes the algorithm complex and difficult to handle. ResNet has a very low error rate comparatively of 3.57%. It has 152 layers but still has lower complexity than the VGG-16 model. Therefore, the ResNet architecture for this problem at hand was more suitable.

4.1 ResNet

ResNet is trained on the ImageNet dataset that has around 11 million images and over 1000 classes. The algorithm is used to extract features from input images. ResNet works on the concept of residual connections. Neural networks are hard to train because of the famous vanishing gradient problem that occurs during backpropagation. This can saturate the performance or even degrade the results in the worst case. 'Identity shortcut connections' are introduced in ResNet which skips one or more layers which facilitate the training of the model without the vanishing gradient problem. The skip connections allow the model to learn an identity mapping function which makes sure that the next higher level will perform at least as good as the lower or previous layer as it trains the 152 layers without the vanishing gradient problem and not worse than that. The ResNet-50 model involves a design of 5 stages. Each stage is made up of an identity block and a convolution block. The convolutional block and the identity block both consist of three convolutional layers. The

ResNet-50 model has over 23 million trainable parameters. Binary cross-entropy was used as a loss function and Adam optimizer.

4.2 ResUNet

RESUNET refers to Deep Residual UNET. It's an encoder-decoder architecture with a bridge/bottleneck between them acting as the connector developed for semantic segmentation. RESUNET is an advancement of the UNet and is a fully connected convolutional neural network that is developed to get high performance by using fewer parameters. It is an improvement of the existing UNET architecture. RESUNET takes the advantage of both the UNET architecture and Deep Residual Learning to design a better and efficient segmentation model. The use of residual blocks helps in developing a network that is deeper without worrying about the vanishing gradient problem or exploding gradients. It also helps in the efficient training of the network. The rich skip connections in the RESUNET facilitate better flow of information between different layers, which helps in better flow of gradients while training - i.e. during backpropagation. This architecture consists of an encoder or a contraction path, a decoder or expansion path, and a bridge or a bottleneck connecting both these paths, similar to a U-Net. In the case of RESUNET, these layers have pre-activated residual blocks. The term gets its name because of its U shape architecture.

4.2.1 Encoder

The encoder, also referred to as the contraction path, takes the input image and passes it through different encoder blocks of residual blocks and max-pooling layers, which helps the network to learn the representation in an abstract way. Each encoder block's output acts like a skip connection for the corresponding decoder block present in the decoder. To reduce the spatial dimensions, strides are used.

4.2.2 Decoder

The decoder, also referred to as the expansion path takes the feature map from the bridge, also referred to as the bottleneck, or the upsampled input from the previous layer of the decoder and concatenates the various output features from the corresponding encoder block or the contraction path and learns a better semantic representation, which is used to generate a segmentation mask. The decoder consists of three/four decoder blocks, and after each block, upscaling is performed, i.e. the spatial dimensions of the feature map are doubled and the number of feature channels is reduced.

The output of the last decoder is passed through a 1×1 convolution filter with sigmoid activation function. The sigmoid

activation function is used to give the segmentation mask representing the pixel-wise classification.

4.3 Approach

4.3.1 Loading ResNet model:

The dataset is loaded and divided into train, validation, and test sets. Get the ResNet50 model and freeze the model weights. It has around 23 million trainable parameters. The newly added dense layers with randomly initialized weights are trained.

4.3.2 Classifier model for detecting whether tumor exists or not:

The output of the ResNet model is the input to the classifier model. The convolutional kernel is passed to a max-pooling layer. Average pooling is used to combine spatially nearby features in the feature maps. This is used to make the representation rigid to small insignificant changes and reduces the computational load too. Flattening is done to transform the whole matrix representing input images for processing. The output processing of the pooled feature map in a dense layer with 256 nodes. Added four more dense layers with ReLU activation function and dropout layers. It was done to avoid the effects of overfitting by dropping certain features so that the model does not simply learn the features. ReLU was used to transform the data non-linearly. It was observed that it performs much better than hyperbolic tangent

function or the sigmoid activation due to its non-linearity. The final dense layer used the softmax activation function which outputs 2 classes according to the problem - i.e. whether a tumor is present or not.

4.3.3 Segmentation model to localize the tumor:

The Residual block for segmentation was designed using Conv2D layers, batch normalization, and ReLU activation functions. Batch normalization is used to standardize the network's inputs. It helps to accelerate the training and provides regularization to the model which in turn improves the generalization. The contraction path or the encoder block consists of four encoder blocks where the inputs are passed on. 2x2 max pooling is followed. The feature maps after every block are doubled in size which helps in learning and generalizing the features efficiently. The bridge takes 256 nodes input and connects the contraction path and the expansion path. In the decoder block or the expansion path, up-sampling is done. Each block takes the previous input and concatenates features with the corresponding residual blocks in the encoder. 2x2 up-sampling is again done here. This helps in accurately reconstructing the images. The final Conv2D layer is added with a sigmoid activation function. A 1x1 convolution layer is applied to produce an output of the same size as the input.

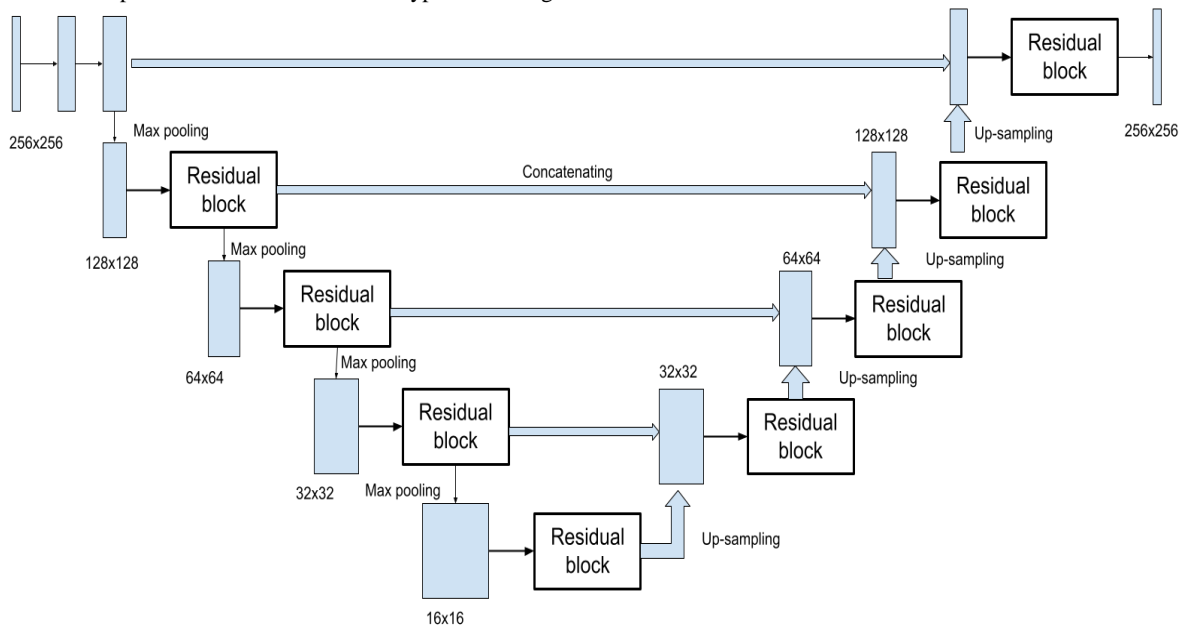


Figure 5: ResUNet Architecture

5. RESULTS AND DISCUSSION

After studying various algorithms used in convolutional neural networks, the algorithm that best suited the problem at hand was chosen. After performing a literature survey, ResNet and ResUNet architectures were chosen for the model. A dataset that contains tumorous and non-tumorous brain MRI scans was used. There was no need to perform feature extraction traditionally done in machine learning algorithms as convolutional neural networks were leveraged and the features from the network itself were used. The input images were successfully processed and the presence of a tumor was detected and further it was localized. An accuracy of 98.1% was achieved. If the confusion matrix is observed, there are 15 images that were incorrectly classified. Three images of class 0 i.e. no tumor present were identified as MRI with tumor. Twelve images that were of class 1 i.e. tumor was present were identified as MRI with no presence of a tumor.

If the nature of the dataset is observed, it can be seen that there are around 2600 brain images without any tumor(healthy) and around 1400 images where a tumor is present. From this observation, it can be concluded that due to a biased dataset, the model interpreted more images as non-tumorous than tumorous. The sensitivity and specificity are also calculated using the formula mentioned below. A specificity of 0.98 and a sensitivity of 0.99 is achieved. Precision, recall and F1-score are also calculated to evaluate the performance of the model. High precision was achieved as the false positive rate is low. F1-score is the weighted average of precision and recall. Table 3 shows the classification report of precision, recall and F1-score.

Table 1: Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	510	3
Actual Positive	12	260

The specificity and sensitivity are calculated as follows:

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} \quad (1)$$

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

Sensitivity is the ability of the model to identify positive patients correctly. Specificity is the ability of the model to identify negative patients correctly.

Table 2: Specificity and Sensitivity

Specificity	Sensitivity
0.98	0.99

Table 3: Classification Report

	Precision	Recall	F1-score
Negative	0.98	0.99	0.99
Positive	0.99	0.96	0.97
Micro avg	0.98	0.98	0.98
Macro avg	0.98	0.98	0.98
Weighted avg	0.98	0.98	0.98

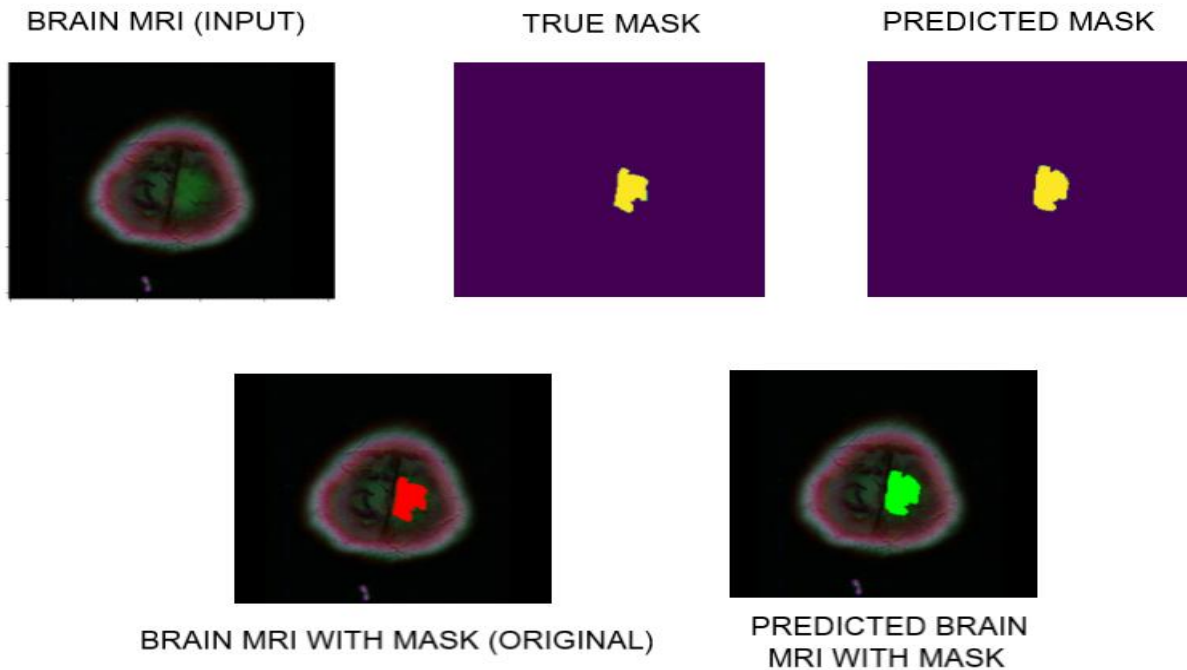


Figure 6: Segmentation result compared with Ground Truth

6. CONCLUSION

The diagnosis of brain tumor detection using MRI was studied. As brain tumor prevails extensively globally according to the surveys, it is vital to explore more in this field of medical imaging to find more accurate and more efficient solutions to reduce the gap between detection and treatment of the patient. Deep learning has proved to be extremely useful in the healthcare domain in recent years owing to its excellent capability to intelligently extract relevant features and process them. This solution will offer a valuable second opinion to medical professionals which can help to escalate further treatment in a timely manner. In this project, a model was successfully implemented that could identify the presence of a tumor and also identify the segmentation mask if the tumor was detected. The accuracy achieved was 98.1%. The accuracy can further be improved by using larger datasets and using hyperparameter tuning to fine-tune the model. Data augmentation techniques can be used to diversify the dataset further. It can be used to alleviate the problem of a biased dataset. More images could be gathered to train

the model. If the confusion matrix is observed, the number of images wrongly classified as non-tumorous was significantly higher than the number of images classified as wrongly tumorous. This is something that can be worked on in the future considering the importance of correctly identifying the tumor in a non-invasive manner quickly, efficiently, and accurately.

7. REFERENCES

- [1] Kavitha AngamuthuRajasekaran and ChellamuthuChinna-Gounder (March 14th 2018). Advanced Brain Tumour Segmentation from MRI Images, High-Resolution Neuroimaging - Basic Physical Principles and Clinical Applications, Ahmet MesrurHalefoğlu, IntechOpen, DOI: 10.5772/intechopen.71416. Available from: <https://www.intechopen.com/chapters/58837>
- [2] Brain Tumor: Statistics, Cancer Net (<https://www.cancer.net/cancer-types/brain-tumor/statistics>), (Accessed 14 Aug 2021)

- [3] S. Bauer, R. Wiest, L. P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumour studies," 2013, [Online]. doi: 10.1109/INDISCON50162.2020.00037.
- [4] Dietmar Krex, Barbara Klink, Christian Hartmann, Andreas von Deimling, Torsten Pietsch, Matthias Simon, Michael Sabel, Joachim P. Steinbach, Oliver Heese, Guido Reifenberger, Michael Weller, Gabriele Schackert, for the German Glioma Network, Long-term survival with glioblastoma multiforme, *Brain*, Volume 130, Issue 10, October 2007, Pages 2596–2606, <https://doi.org/10.1093/brain/awm204>
- [5] T. Hossain, F. S. Shishir, M. Ashraf, M. A. Al Nasim and F. Muhammad Shah, "Brain Tumor Detection Using Convolutional Neural Network," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019
- [6] Seetha, J & Selvakumar Raja, S. (2018). "Brain Tumor Classification Using Convolutional Neural Networks. *Biomedical and Pharmacology Journal*". 11. 1457-1461. [10.13005/bpj/1511](https://doi.org/10.13005/bpj/1511).
- [7] S. Grampurohit, V. Shalavadi, V. R. Dhotargavi, M. Kudari and S. Jolad, "Brain Tumor Detection Using Deep Learning Models," 2020 IEEE India Council International Subsections Conference (INDISCON), 2020, pp. 129-134, doi: 10.1109/INDISCON50162.2020.00037.
- [8] Jha, Debesh & Smedsrud, Pia & Riegler, Michael & Johansen, Dag & de Lange, Thomas & Halvorsen, Pål & Johansen, Håvard & Simulmet., (2019). ResUNet++: An Advanced Architecture for Medical Image Segmentation.
- [9] M. Ali, S. O. Gilani, A. Waris, K. Zafar and M. Jamil, "Brain Tumour Image Segmentation Using Deep Networks," in *IEEE Access*, vol. 8, pp. 153589-153598, 2020, doi: 10.1109/ACCESS.2020.3018160.
- [10] Pereira S, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging*. 2016 May;35(5):1240-1251. doi: 10.1109/TMI.2016.2538465. Epub 2016 Mar 4. PMID: 26960222.
- [11] L. G. Nyul, J. K. Udupa, and X. Zhang, "New variants of a method of mri scale standardization," *IEEE Transactions on Medical Imaging*, vol. 19, no. 2, pp. 143–150, 2000.
- [12] [Dataset] Brain Image Segmentation; Mateusz Buda, Ashirbani Saha, Maciej A. Mazurowski "Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm." *Computers in Biology and Medicine*, 2019; <https://www.kaggle.com/mateuszbudalgg-mri-segmentation>