Diyala Journal of Engineering Sciences

Journal homepage: https://djes.info/index.php/djes

ISSN: 1999-8716 (Print); 2616-6909 (Online)

Segmentation of Human Brain Gliomas Tumour Images using U-Net Architecture with Transfer Learning

Assalah Zaki Atiyah*, Khawla Hussein Ali

Department of Computer Science, College of Education for Pure Sciences, University of Basrah, 42001 Basrah, Iraq

ARTICLE INFO

Article history:
Received: November 2, 2021
Accepted: January 27, 2022

Keywords:
Gliomas images
MRI segmentation
Brain tumor
U-Net
Transfer learning
BraTS2020

ABSTRACT

The complexity of segmenting a brain tumour is critical in medical image processing. Treatment options and patient survival rates can only be improved if brain tumours can be prevented and treated. Segmentation of the brain is the most complex and time-consuming task to diagnose cancer utilizing a manual approach for numerous magnetic resonance images (MRI). The aim of MRI brain tumour image segmentation that to build an automated magnetic resonance imaging tumour segmentation system with separate the area of tumour and provided a clear boundary of the tumour region. U-Nets with different transfer learning models as backbones are presented in this paper, there are ResNet50, DenseNet169 and EfficientNet-B7. Brain lesion segmentation is performed using the multimodal brain tumor segmentation challenge 2020 dataset (BraTS2020). Based on MRI scans of the brain, the tumor segmentation technique is assessed using F_1-score, Dice loss, and intersection over union score (IoU). The U-Net encoder used with EfficientNet-B7 outperforms all other architectures in terms of performance metrics across the board. Overall, the results of this experiment are rather excellent. The Dice-loss score was 0.009435, and the score of IoU was 0.7435, F_1-score was 0.9848, accuracy was 0.9924, precision was 0.9829, recall was 0.9868, and specificity was 0.9943. The U-Net with EfficientNet-B7 architecture was shown to be crucial in the treatment of brain tumors, according to the findings of the experiments.

1. Introduction

Brain tumours are abnormal masses of brain tissue. A very hard skull protects the brain. Any development in such a tiny area might cause issues [1]. Four conventional magnetic resonance imaging (MRI) modalities are employed in brain image studies: native T1-weighted (T1), T2 fluid-attenuated inversion recovery (T2-FLAIR), T2-weighted (T2), and post-contrast T1-weighted (T1ce) [2]. Automatic methods for segmenting brain lesions are frequently used hand-crafted features such as edges, corners, histogram of gradient, local binary pattern. The classifier is given these features after they have been extracted. The training procedure of the classifier is not affected by the nature of those features. Figure 1 depicts the four different MRI modalities [2].

Semantic segmentation is commonly used in medical imaging to identify the precise location and the form of the body's structures and is essential to the proper assessment of medical disorders and their treatment.

Deep learning has recently achieved substantial advancements in a range of computer vision applications, particularly object detection, image classification, and semantic

* Corresponding author.
E-mail address: pgs2179@uobasrah.edu.iq
DOI: 10.24237/djes.2022.15102

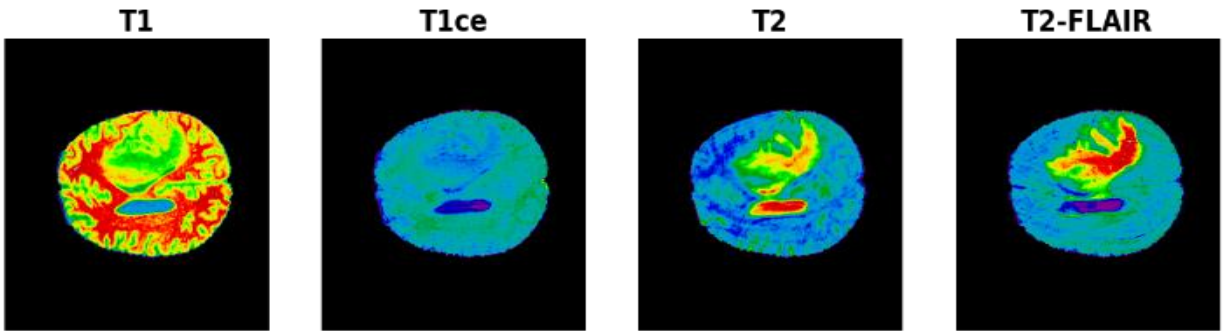


Figure 1. Four MRI Modalities. Sourced From [2]

segmentation, to name a few examples. With their remarkable performance and high accuracy in image segmentation, deep neural networks have recently gained popularity among researchers. Detection and segmentation of the human brain is a difficult problem to solve in reality due to its complexity. When it comes to tumour-bearing data, it is the 3D or 2D information that varies widely from patient to patient in terms of the structure, size, and location of the tumour. In addition, MRI information of brain tumours extracted from diagnostic scans or synthetic databases is inherently complex, requiring a large amount of device memory for tumour segmentation. U-Net with ResNet50, DenseNet169 and EfficientNet-B7 provided. In the proposed method, the brain MRI images has been augmented, performed data pre-processing methods to modify the actual data, evaluated different deep learning architectures, and offered a comparative study of these models.

This paper was conducted with various architectures for brain tumour segmentation. Enhanced output was achieved by the proposed architecture. It takes just a few seconds to segment the image using a trained model. Clinical professionals can take hours for manual segmentation of tumours, however. In the field of image diagnosis, this article is contributing to a model that can more accurately and efficiently diagnose the tumour.

This paper discusses techniques for segmenting brain tumours. This article is broken down into many parts. The introduction and background of brain tumours have been described in section one, literature review has been described in section two, the architectures explanation, dataset description and performance metrics have been described in section three, the proposed system, pre-processing the data, training and implementation details, the experimental results have been described in section four and conclusion with future work has been described in section five.

2. Literature review

Over the last few decades, research into the automatic segmentation of brain tumours has increased, showing an increasing demand for this field of study, which is still ongoing. Several strategies for detecting and segmenting tumours on MRI Image data have indeed been proposed in the research. In 2015, O. Ronneberger et al. [3] provides a homogeneous fully CNN named U-Net. U-Net greatly increased effectiveness in medical image segmentation tasks, leaving the down sampling process to link with the feature graph of the down sampling process to capture contextual information and the correct up sampling process to establish the correct location. U-Net is frequently utilized in the sector of medical image analysis as a result of it being exceptionally efficient in the final training of a limited number of images.

In 2017, H. Dong et al. [4], Deep convolutional networks based on U-Nets was applied to segment brain tumours, this study was evaluated on multimodal brain tumor image segmentation (BraTS 2015) datasets, dice similarity coefficient (DSC) has been compute for high grade glioma (HGG) and low grade glioma (LGG) combined cases and obtained 0.86 of the complete tumor segmentation, the limitation of this segmentation method has been evaluated using a cross-validation scheme, which can be provide an unbiased predictor.

In 2020, R. A. Zeineldin et al. [5], residual neural network, dense convolutional network, and NASNet have been utilized in this study to build a fully programmed brain tumor recognition and segmentation, this deep learning architectures have been evaluated on-line based MRI datasets of brain tumor segmentation BraTS 2019, dice and Hausdorff distance scores of obtained segmentation results are about 0.81 to 0.84 and 9.8 to 19.7 correspondingly, the lack of this study was that false positives (FP) indicated was high values of both recall and specificity, which might not precisely reflect the actual performance. In 2020, X. Feng et al. [6] produced a 3D U-Net ensemble for brain tumour segmentation, multimodal brain tumor segmentation (BraTS 2018) challenge has been used in the study, the limitation of this structure , it hard to pick of the best model and/or hyper-parameter set because of that most models perform similarly. It is indeed one disadvantage of DCNN as the "black-box" nature of the network makes it challenging to analyze the effect of network structure and parameter except from the final performance. In 2021, T. Sadad et al. [7] developed U-Net with ResNet50 architecture for segmentation of tumours utilizing the Figshare dataset, getting an IoU score of 0.9504. In 2021, F. Isensee et al. [8], nnU-Net utilized , nnU-Net pipeline's segmentation performance has been demonstrated to be greatly enhanced by the addition of BraTS specific characteristics such as postprocessing, data augmentation, and region-based training. For the BraTS challenge 2020 segmentation problem, excellent results have been obtained using the nnU-Net configuration's baseline setup, however, one limitation of this approach is the lack spans a small number of the modifications and lacks sufficiently extensive experimental validation thereof. In 2020, P. K. Gadosey et al. [9] introduce the stripped-down U-Net (SD U-Net), experiments on the benchmark dataset of the Internatioanl symposium on biomedical imaging (ISBI) challenge for segmentation of neuronal structures in electron microscopic (EM) stacks and the medical segmentation decathlon (MSD) challenge brain tumor segmentation (BRATs) dataset show that the proposed model, a deep neural network that is highly fast, compact, and computationally effective for segmenting medical images on devices with low processing resources.

Although these approaches offer certain advantages in some areas, the problem of unequal tumour and background voxel distribution in brain tumour segmentation must be addressed immediately. As a result, this study presents a learning mechanism for improving the segmentation approach. However, the disadvantage of depthwise convolutions compared with standard convolutions is lower accuracy performance. In 2020, A. A. Pravitasari et al. [10] for the segmentation of MRI brain tumours proposed UNet-VGG16. This model or architecture is a fusion of the U-Net and VGG16 architectures, with transfer learning used to simplify the U-Net architecture, real dataset from general hospital has been used in this study, the learning dataset shows that this approach has a high accuracy of 96.1%. Calculating the correct classification ratio (CCR) and comparing the segmentation result with the ground truth are used to validate the segmentation result. With a mean CCR score of 95.69 %, this UNet-VGG16 was able to distinguish the brain tumour area, the size of datasets analyzed was limited.

3. Preliminaries

3.1 Method

3.1.1 U-Net architecture

In 2015 for the first time, O. Ronneberger et al.[3] suggested the U-Net architecture, a convoluted network designed only for biological-image analysis. The model is shaped like a "U". To put it simply, the encoder is a simple convolutional process, whereas its decoder is made up of 2D convolutional layers that have been transposed. The block diagram of U-Net architecture is depicted in Figure 3. The contraction route (also called the encoder component) is the primary route in the U-Net structure and is often used to capture the framework of the input image.

The encoder is made up of maximum pooling and convolution layers arranged one above the other. To permit precise localization using transposed convolutions, the second route

is an uneven increasing route (additionally called the decoder component). Fully convolutional networks can only interpret images of a certain size since they have no thick layers, making them unable to handle larger images. Before max-pooling, two consecutive convolution layers are implemented in this type of architecture. The dimensions of the data are halved when pooling is applied, as a result, a lot of data will be lost. So, before each convolution, pooling layers are stacked to build up richer data representation without losing all spatial information quickly. To the U-Net structure, we added skip connections among the first convolution layer and the max-pooling layer on each level. The key purpose of this was to increase the accuracy and uniform distribution of parameters in the layers.

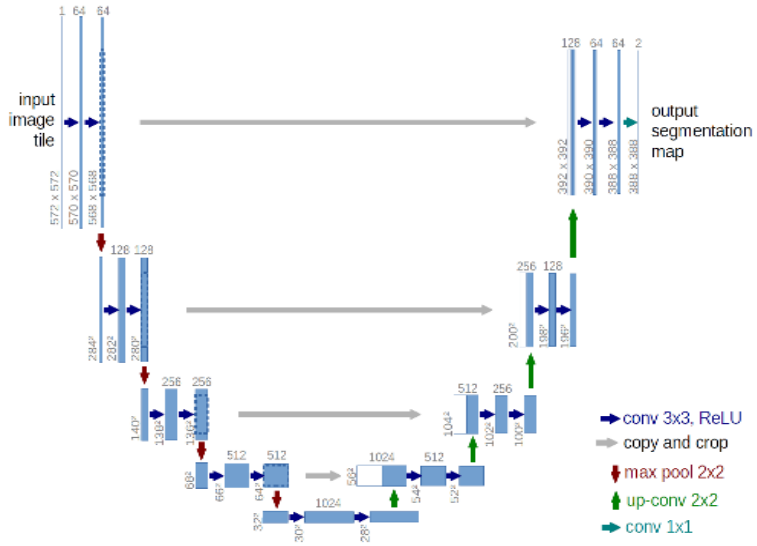


Figure 3. U-Net architecture sourced from [5]

3.1.2 ResNet architecture

For ImageNet's greatest image identification task, K. He et al. [11] built ResNet, a neural network that had an error rate of 3.57%. This was done using a 152-layer deep CNN architecture. ResNet's design promoted the use of deeper networks than AlexNet and VGG-Net. Skip connections (also known as residual connections) are included in the ResNet architecture to reduce data redundancy during deep network training and optimization. The remaining connections might be used to train a 1001-layer model. The bulk of ResNet's architecture is made up of residual blocks. ResNet's residual blocks are connected, even though the layers of a shallow neural network are linked. By increasing the network's capacity, ResNet connections maintain their information and shorten the training period of the model.

3.1.3 DenseNet architecture

It was introduced by G. Huang et al. [12] as a feed-forward network that links one layer to the next, and it has since gained widespread acceptance. To ensure that the deep convolution network was trained more accurately and efficiently, a deeper convolution network with shorter connections between layers near to the input and close to the output was employed to ensure that the deep convolution network was taught more precisely and effectively. Non-linear transformation is prevented via direct connections from each layer to the next in DenseNet, unlike ResNet skip-syncs. When it comes to function mappings, all prior levels from x_0 all the way up to x_{l-1} are taken into account. According to the authors, "the wavelength of the feature-map created in layers 0, 1, 2,..., l-1 is described as $x_l = H_l([x_0, x_1, \ldots, x_{l-1}])$ where $[x_0, x_1,..., x_{l-1}]$" [12].

3.1.4 EfficientNet architecture

The EfficientNet model was developed by M. Tan and Q. V. Le [13] of Google Research's

Brain team, who published their findings in the journal "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". To begin, a new baseline architecture named "EfficientNetB0" was constructed and then scaled up to build the EfficientNet family through a compound scaling method. Eight EfficientNet variations are powered by this method, ranging from 5.3 million parameters to 66 million.

The researchers began by automating the construction of neural networks by constructing a primitive network using a method known as the search for neural network architecture. It maximizes both accuracy and efficiency by calculating the number of floating-point operations per second (FLOPS). This design makes use of convolution with a moveable inverted bottleneck (MBConv). The amount of these MBConv segments vary according to the EfficientNet family. As we proceed through EfficientNet B0-B7, the depth, width, resolution, and model size all rise, while the accuracy also increases. The EfficientNetB7 model outperforms earlier state-of-the-art CNNs on ImageNet [13].

3.2 Data set

The publicly accessible benchmark dataset was utilized in this paper. The BraTS 2020 dataset has been suggested for the identification and segmentation of brain lesions in an automated manner.

The Brain Tumor Segmentation Study (BraTS) has traditionally been designed to examine cutting-edge techniques for the segmentation of brain tumors in magnetic resonance imaging (MRI). The BraTS2020 dataset was produced entirely from preoperative data collected from a variety of institutions, and it focuses on segmenting brain tumors that are fundamentally diverse (in terms of shape, histology, and appearance), including gliomas.

This data consists of 293 HGG and 76 low grade glioma (LGG) instances are included in the BraTS 2020 training dataset. T_1, T_2, T_1ce, and T_2-FLAIR imaging modalities are co-registered with a voxel size of 240 x 240 and an isotropic resolution of one millimeter for all imaging modalities, Figure 2 shows four modalities of BraTS2020 dataset with ground truth.

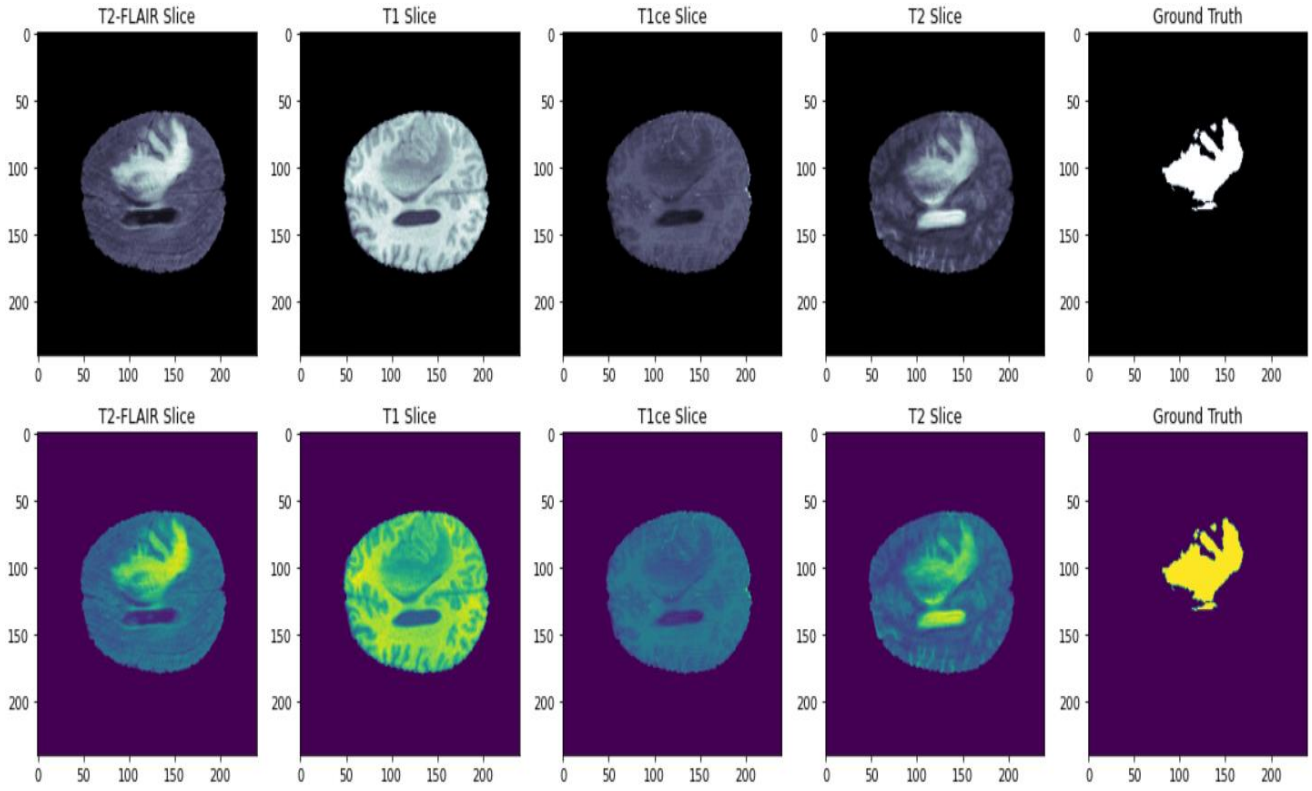


Figure 2. BraTS2020 dataset modalities with ground truth

The training data has annotations, but the validation and testing datasets (125 and 166 instances, respectively) do not. Estimated volumes of segmentation may be submitted to the organizer's website by participants to compare their methods. For the validation assessment, several contributions are permitted; however, for the final testing evaluation, only one submission is permitted per participant.

NifTI files with T_1, T_2, T_1ce, and T_2-FLAIR descriptions are supplied for all multimodal BraTS scans. These scans came from diverse clinical procedures and scanners from various organizations. Images were manually segmented by one to four assessors for all of the datasets. There are many types of tumor cores, including necrotic and non-enhancing tumor cores, enhancing tumors, peritumoral edema, and others.

3.3 Performance metrics

3.3.1 Accuracy

Accuracy is a measure of the number of correct predictions out of all predictions.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

Here, true positives (TP) and negatives (TN), as well as false positives (FP) and negatives (FN).

3.3.2 Precision

Precision is a measure of the accuracy of a positive prediction. In other words, if an outcome is predicted to be positive, how certain that it can be actually positive.

$$Precision = \frac{TP}{(TP + FP)} \quad (2)$$

3.3.3 Recall

The recall is the measure of how many true positives are predicted out of all actual positives in the dataset.

$$Recall(Sensitivity) = \frac{TP}{TP + FN} \quad (3)$$

3.3.4 Specificity

Specificity is the measure of how many true negatives are predicted out of all actual negatives in the dataset.

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

3.3.5 IoU Score

To compute the IoU score, divide the intersection point between the actual data (ground truth) and predicted segmentation by the point of union between the actual data (ground truth) mask and predicted segmentation mask." When assessing how much overlap there is between two masks or bounding boxes [14], it is a valuable statistic.

$$IoU = \frac{(ground\ truth\ \cap prediction)}{(ground\ truth \cup prediction)} \quad (5)$$

3.3.6 F_1-Score

The F_1 score is determined by using "the harmonic mean of recall and precision" as the starting point [15]. The F_1 score is also the same as the dice score.

$$F1 - Score = 2 \times \frac{(precision \times recall)}{(precision + recall)} \quad (6)$$

"An F_1-score may have the highest possible value of 1.0, which shows perfect precision and recall, and the highest possible value of 0 if either precision or recall is zero".

3.3.7 Dice loss

In this section, we'll talk about dice loss, which is computed as one minus the dice coefficient. The dice coefficient is a common metric for pixel segmentation that may also be used as a loss function if it is altered to work in this manner. " The dice coefficient is computed by multiplying the intersection area by the total number of pixels in both images" [16]. The dice coefficient is determined using the following formula:

$$Dice = \frac{2 \times TP}{2 \times TP + FN + FP} \quad (7)$$

$$Dice\ loss = 1 - Dice \quad (8)$$

4. Proposed system

To ensure the effectiveness of MRI image segmentation based on the human brain, different pretrained models as U-Net encoder addressed such as ResNet50, DenseNet169, and EfficientNet-B7. Figure 3 shows the proposed approach for segmentation. The brain MRI

images were administered and then used for training the segmentation model. The segmented image can be predicted using the trained model.

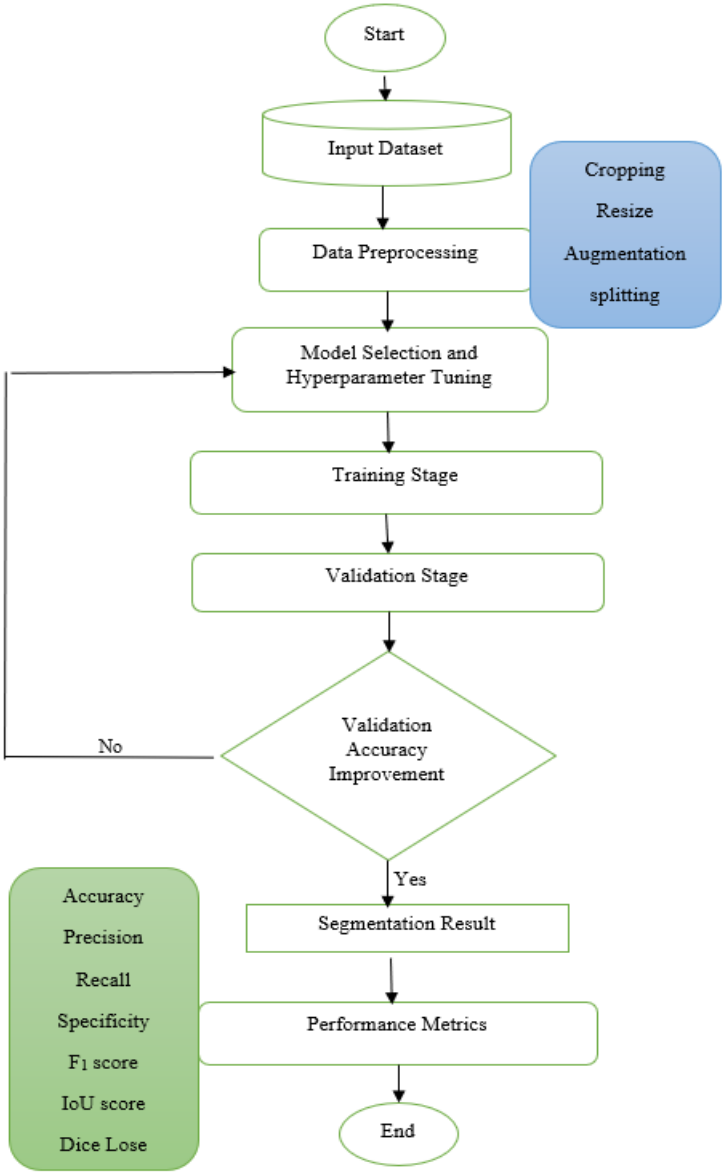


Figure 3. Flowchart of the proposed brain tumour segmentation model

4.1 Data preprocessing

Cropping images from the center is done by this method. Standardscaler normalization removes any distortion from the MRI intensity measurement, which is dependent on the imaging technology and scanner utilized.

The standard score for Sample x is computed in the following way:

$$Z = \frac{(x - u)}{s} \quad (9)$$

"Where s is the standard deviation of the training samples, or 1 if it is False, and u denotes the mean of the training samples, or 0 if it is False. According to the standard scalers operating concept, the data will be transformed into distribution with a standard deviation of one and a mean of zero".

This is done feature by feature in the case of multivariate data. Figure 4 shows the image visualization after pre-processing.

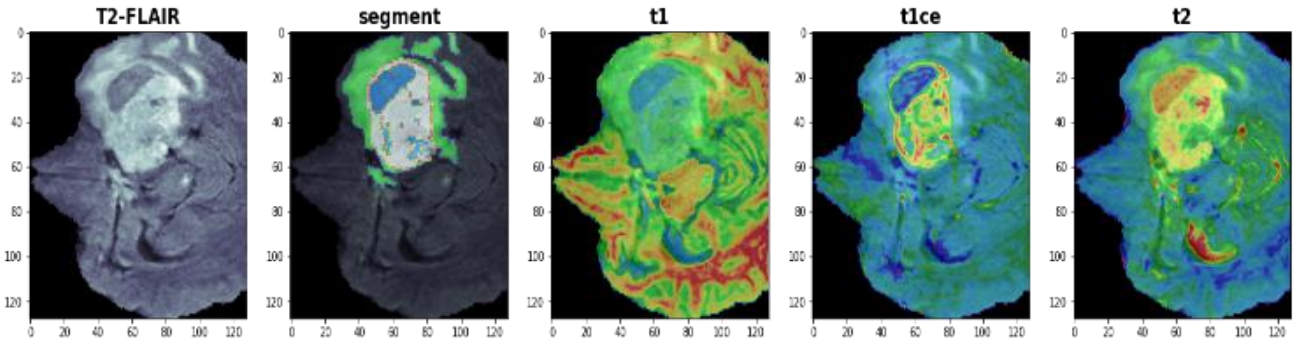


Figure 4. Image after preprocessing

4.2 Training and implementation details

The BraTS dataset contains MRI images and related segmentation output images for training and testing models. Clinically skilled neuroradiologists correct the output or ground truth labeled images. Folders for train and validation datasets are included in the BraTS dataset. All 369 subfolders in the train data folder are organized by modality, and each subfolder includes 5 nifty-format images of 5 distinct modalities. As a result, the train data folder has 1845 images. Each of the 125 subfolders in the validation data folder comprises four images of four distinct modalities, such as T_1, T_2-FLAIR, T_2, and T_1ce. As a result, there are a total of 500 images in the data folder for validation.

After evaluating the data, the images are pre-processed by standardizing the intensity value and cropping them. MRI records are split into three categories: train set, test set, and validation set. 60% of the images are used as the train set (2999 images), 20% are used as the test set (1001 images), and 20% are used as the validation set (999 images). The training images are data augmented, which aids generalization and enhances accuracy. Augmentation techniques are taken from the albumentation library such as grid distortion, random brightness contrast, elastic transform, optical distortion was employed to enhance the input image and offer more information for the model to learn. Before we can begin training our model, we need to build the learning process. This includes an optimizer, a loss function, and if desired, additional metrics such as F_1-score and IoU-score. The Adam optimizer was used, with a learning rate of 0.0001.

A range of approaches U-Net with a variety of transfer learning architectures such as ResNet50, DenseNet169, and EfficientNet-B7 in this paper.

In brain tumor representations, skip topologies provide perfect segmentation by using a high-level expression from dense sequencing layers. The three components of the design are bottleneck, contraction, and expansion. The expansion is made up of many contraction lengths. A convolution level addition input is followed by a maximum pooling limit for each block. In each block, each CNN layer has its sample level, with the lowest layer modulating between the expansion and contraction levels. After convolutions and activations, the ReLU activation function and batch normalization were utilized to prevent deep learning models from falling out of the experiment.

Features are extracted from the data loader to train the machine learning modes. The extracted features are then converted to a 2D array for better fitting the model.

4.3 Results & discussions

The proposed architectures' segmentation capabilities evaluated by using IoU, F_1 score, dice loss, accuracy, precision, recall, and specificity metrics. 200 epochs and 16 batch sizes are used to train the network. Four NVIDIA P40 GPUs are used to train our network in Pytorch.

The performance metrics of all architectures of the deep learning models are presented in

Table 1. All deep learning architectures are beaten in terms of performance by the EfficientNet-B7 encoder-based U-Net, according to Table 1. The U-Net with EfficientNet-B7 performed better with a dice loss of 0.009435, IoU score of 0.7435, F_1 score of 0. 9848, the accuracy of 0.9924, precision of 0.9829, recall of 0.9868, and specificity of 0.9943. In the U-Net with the EfficientNet-B7 model, the proportion of intersection between the targeted mask and our predicted output is higher.

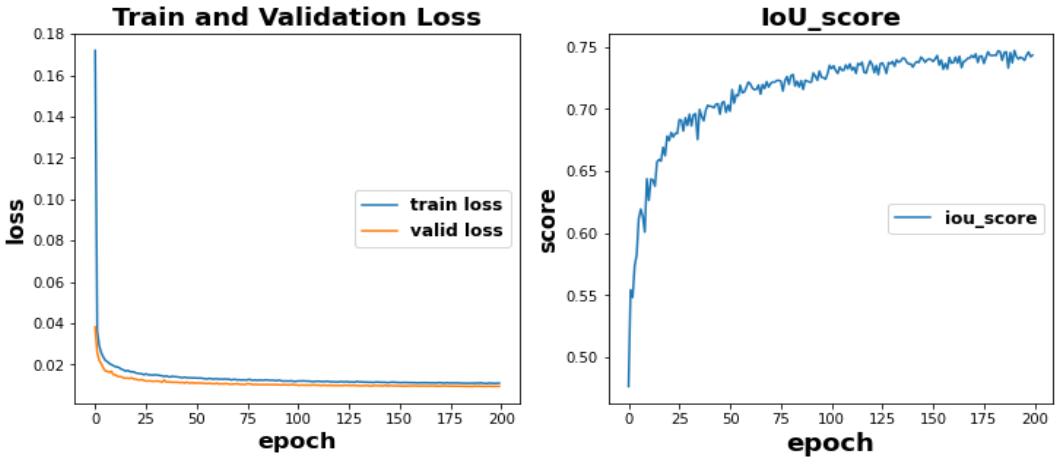
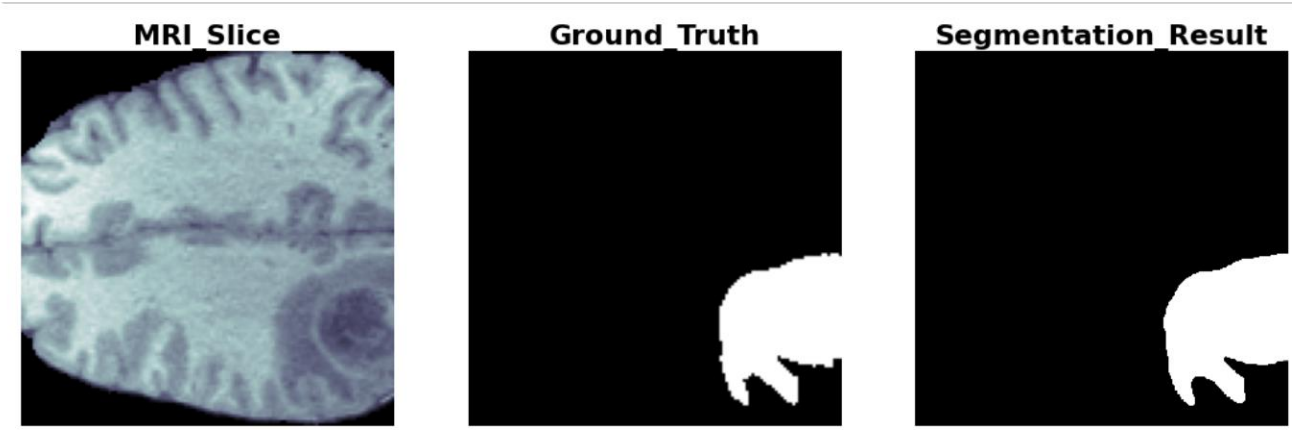
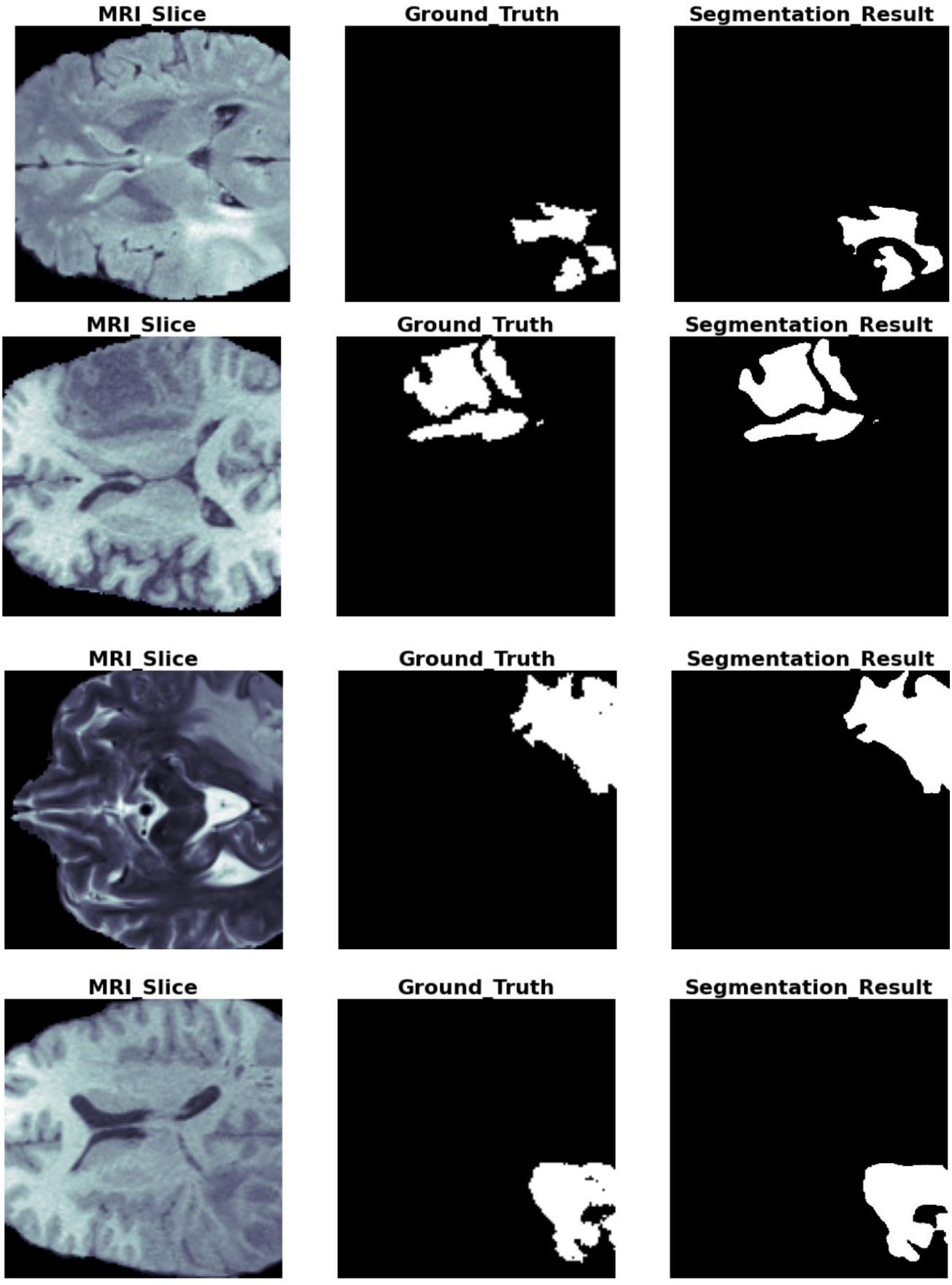


Figure 5. Dice Loss and F_1 Score Graph of U-Net with EfficientNet-B7

The training parameters are the most important aspect to take into account when calculating the compute time of a CNN. Using the same dataset and setting up all training parameters, in the same manner, is thus critical. After the network has been trained, it may be used to image segmentation. Image segmentation is completed in a matter of seconds when using the previously trained model. Manual tumor segmentation by clinicians, on the other hand, might take many hours or even days. Accurate, rapid, and low-cost image segmentation approaches are recommended. There is a chance that this might save the lives of countless individuals by allowing physicians to diagnose a brain tumor quickly and precisely.

Figure 5 shows the dice loss and IoU score recorded for each epoch for U-Net with EfficientNet-B7 architecture Figure 6 depicts the prediction of U-Net with EfficientNet-B7.




MRI_Slice
Ground_Truth
Segmentation_Result
MRI_Slice
Ground_Truth
Segmentation_Result
MRI_Slice
Ground_Truth
Segmentation_Result
MRI_Slice
Ground_Truth
Segmentation_Result

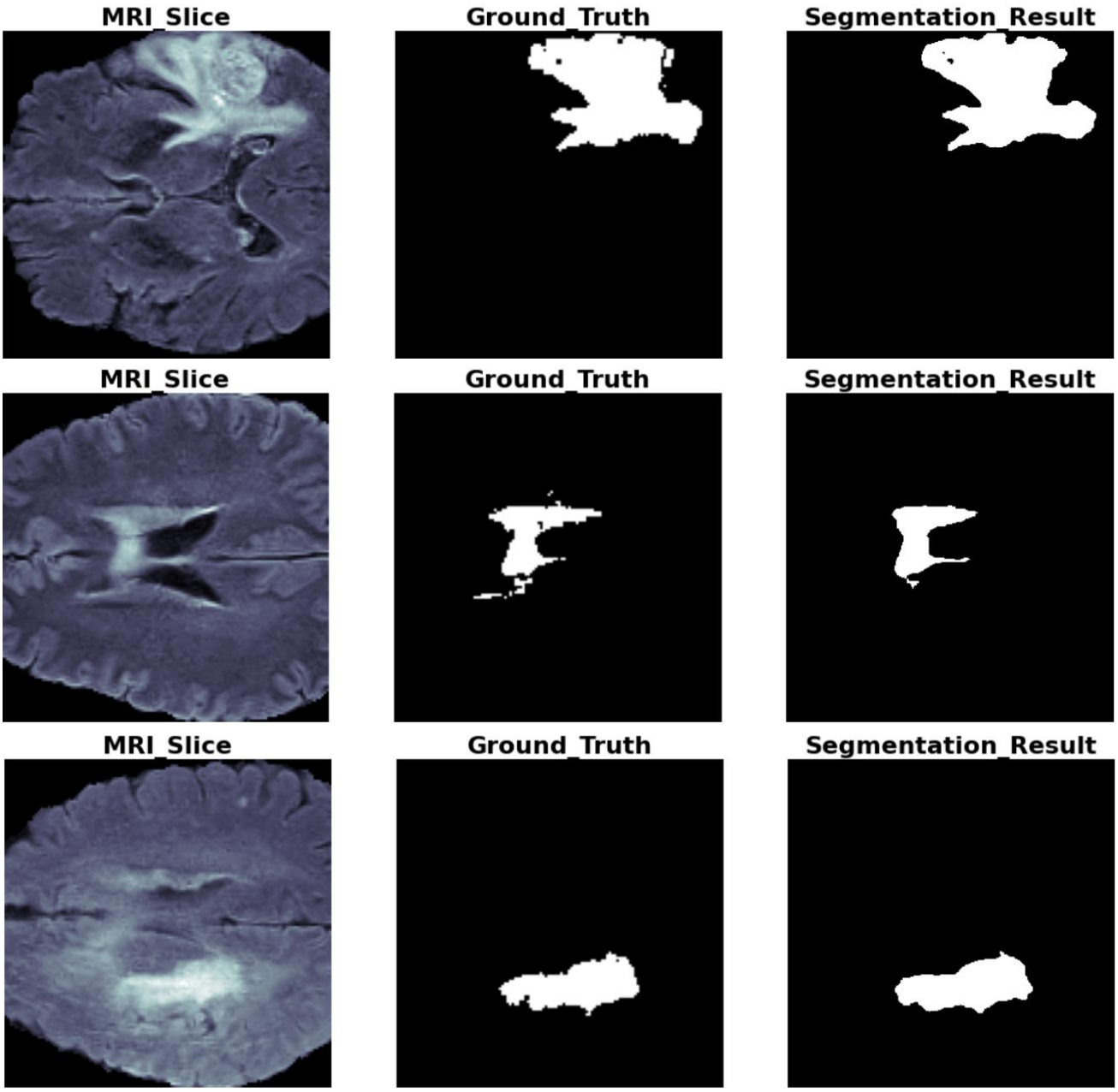


Figure 6. Segmentation result based on U-Net with EfficientNet-B7 encoder architecture

The segmentation is done to the eight-testing data under of U-Net with EfficientNetB7 as encoder. The visualization of segmentation results is described in Figure 6. This figure has shown that segmentation results of sample sequence could recognize the tumour area as region of interest in various tumour size and location, both on the right or left of the brain.

Table 1: Results of deep learning models

Model	Dice Loss	F_1 Score	IoU	Accuracy	Precision	Recall	Specificity
U-Net_ResNet50	0.009789	0.9801	0.7374	0.9903	0.9751	0.9864	0.9916
U-Net_DenseNet169	0.009848	0.9846	0.7337	0.9923	0.9821	0.9871	0.9940
U-Net_EfficientNetB7	0.009435	0.9848	0.7435	0.9924	0.9829	0.9868	0.9943

F. Isensee et al. [8] proposed segmentation of brain tumours using the BRATS2020 dataset, which is the same dataset that is used in this study. The model achieved the dice score of 88.95%, but our proposed model achieved 98.48% which is better than the F. Isensee et al. [8] model. An evaluation of U-Net using EfficientNet B7 and a comparison to other current approaches can be found in Table 2. The proposed model performed well when compared to the findings of earlier studies, according to the comparison.

Table 2: Comparison of the proposed model with previous works

Author	Network	Accuracy	F_1/Dice Score
F. Isensee [8]	U-Net ResNet50	-	0.8895
P. K. Gadosey [9]	SD-UNet	98.66	82.75
A. A. Pravitasari [10]	UNet-VGG16	96.1	-
C. Lyu et al. [17]	U-Net -VAE	-	0.8729
Proposed method	U-Net EfficientNetB7	0.9924	0.9848

5. Conclusion and future work

It may be challenging to segment brain tumours due to the intricacy of MRI brain imaging; nonetheless, its objectives to foresee malignancies via the use of artificial intelligence models make this effort worthwhile. For the automatic segmentation of brain tumours, the suggested system makes use of U-Net with several transfer learning models as encoder architecture. They simplify and expedite the imaging and segmentation of brain tumours. The U-Net architecture, which makes use of the EfficientNet-B7 encoder, outperforms all other architectures in terms of performance. Dice loss was 0.009435, IoU score was 0.7435, F_1 score was 0.9848, the accuracy was 0.9924, precision was 0.9829, the recall was 0.9868, and specificity was 0.9943 for the U-Net with EfficientNet-B7. According to the results of the tests, the U-Net with EfficientNet-B7 design plays a critical role in the treatment of brain tumours, particularly in the early stages. The suggested U-Net with the EfficientNet-B7 model offers a framework for predicting the segmentation of brain lesions and aids in the precise segmentation of the location of the lesions, both of which are important. Our study reveals that the suggested technique outperforms the current approaches for the segmentation of brain tumours by a significant margin.

MRI data cannot be completely exploited by the 2D U-Net model; hence the architecture lacks semantics and local characteristics between slices due to the model's restrictions.

The study could be expanded in the future by developing more powerful patch extraction techniques to help improve segmentation accuracy. So far, we have only introduced the use of 2D patches for extraction and training; however, this work can be extended to include the creation of models to be trained on 3D patches to find effective ways to extract 3D patches.

References

[1] S. Han,"Brain Tumor: Types, Risk Factors, and Symptoms, Diagnosis, Treatment," [Online], June 6, 2017. Available: https://www.healthline.com/health/brain-tumor.

[2] A. Z. Atiyah and K. H. Ali, " Brain MRI Images Segmentation Based on U-Net Architecture," IJEEE Journal, pp. 21-27, 2021.

[3] O. Ronneberger, Philipp Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," CoRR, vol. abs/1505.0, pp. 16591–16603, 2015.

[4] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-net based fully convolutional networks," Commun. Comput. Inf. Sci., vol. 723, pp. 506–517, 2017.

[5] R. A. Zeineldin, M. E. Karar, J. Coburger, C. R. Wirtz, and O. Burgert, "DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images," Int. J. Comput. Assist. Radiol. Surg., vol. 15, no. 6, pp. 909–920, 2020.

[6] X. Feng, N. J. Tustison, S. H. Patel, and C. H. Meyer, "Brain Tumor Segmentation Using an Ensemble of 3D U-Nets and Overall Survival Prediction Using Radiomic Features," Front. Comput. Neurosci., vol. 14, no. April, pp. 1–12, 2020.

[7] T. Sadad, "Brain tumor detection and multi-classification using advanced deep learning techniques," Microsc. Res. Tech., no. October 2020, pp. 1296–1308, 2021.

[8] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnU-Net for Brain Tumor Segmentation," pp. 118–132, 2021.

[9] P. K. Gadosey , "SD-UNET: Stripping down U-net for segmentation of biomedical images on platforms with low computational budgets," Diagnostics, vol. 10, no. 2, pp. 1–18, 2020.

[10] A. A. Pravitasari, N. Iriawan, M. Almuhayar, T. Azmi, Irhamah, K. Fithriasari, S. W. Purnami, W. Ferriastuti, "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation," Telkomnika, vol. 18, no. 3, pp. 1310~1318,2020.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, 2016.

[12] G. Huang, Z. Liu, G. Pleiss, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," IEEE Trans. Pattern Anal. Mach. Intell., 2020.

[13] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 36th Int. Conf. Mach. Learn. ICML, vol. 2019-June, pp. 10691–10700, 2019.

[14] "Intersection over Union (IoU) for object detection – PyImageSearch," [Online], Jul. 16, 2021,Available: https://www.pyimagesearch.com.

[15] "F-Score Definition | DeepAI," [Online], Jul. 16, 2021, Available: https://deepai.org/machine-learning-glossary-and-terms/f-score.

[16] "An overview of semantic image segmentation," [Online], Jul. 16, 2021, Available: https://www.jeremyjordan.me/semantic-segmentation.

[17] C. Lyu and H. Shu, "A Two-Stage Cascade Model with Variational Autoencoders and Attention Gates for MRI Brain Tumor Segmentation,". arXiv preprint arXiv:2011.02881,2020.