EfficientNet FPN-Based Brain Tumor Segmentation on the T2-FLAIR Subset of BraTS 2020 Dataset.

Van Kiet Vo and Phuoc Huy Tran

May 18, 2023

EfficientNet FPN-based Brain Tumor Segmentation on the T2-FLAIR Subset of BraTS 2020 Dataset.

Van Kiet Vo[1] and Phuoc Huy Tran[1]

[1] Can Tho University, Can Tho, Vietnam
vovankiet1108@gmail.com

Abstract. Brain tumors are complex and dangerous conditions that require accurate diagnosis for effective treatment. While MRI is a crucial diagnostic tool, the process of interpreting and evaluating MRI is time-consuming and requires expertise. Developing AI and machine learning methods to predict brain tumors can speed up diagnosis, reduce wait times, and improve accuracy. In this study, the authors validated and used the EfficientNet model combined with FPN to segment brain tumors in reality. We trained model on the BraTS 2020 dataset, achieving good performance on the test and evaluation sets. The proposed method demonstrated an average IoU accuracy of 0.9083 and 0.8878 and an average Dice accuracy of 0.9336 and 0.9303 on the test and evaluation sets, respectively.

Keywords: Brain tumor, EfficientNet model, FPN, BraTS 2020.

1 Introduction

Brain tumor is a dangerous and complex disease in the nervous system, and accurate diagnosis is crucial. Magnetic resonance imaging (MRI) is a useful tool for identifying brain tumors, but interpreting and evaluating MRI images is a time-consuming and specialized task. Developing methods to predict brain tumors from MRI images can help speed up diagnosis and reduce waiting time for patients. Using artificial intelligence and machine learning to predict brain tumors can increase accuracy and reduce errors, which can lead to accurate and timely treatment decisions for patients, reducing the risk of complications and increasing the chances of recovery. Therefore, predicting brain tumors from MRI images is essential in healthcare to improve the accuracy and speed of diagnosis, and reduce waiting time for patients.

2 Related work

There are numerous scientific research articles investigating the issue of brain tumor partitioning. The article "Deep learning based brain tumor segmentation: a survey" [1] refers to over 150 scientific research articles on this issue. These studies employ various deep learning methods for brain tumor segmentation, such as Convolutional Neural

Networks with U-Net, V-Net, 3D-FCN and 3D-UNet networks, or methods using Recurrent Neural Networks, Fully Convolutional Networks and Encoder-Decoder Networks. The article also evaluates the advantages and limitations of each method and concludes that CNN models are the most popular and have higher accuracy compared to other methods [1].

One of these papers is "Tumor Segmentation in Brain MRI: U-Nets versus Feature Pyramid Network" [2]. This article compares the performance of two neural network architectures, U-Net and Feature Pyramid Network (FPN), in segmenting brain tumors from MRI images. The authors conducted their experiments on the BRATS 2017 and 2018 datasets. The results indicated that FPN outperformed U-Net with evaluation metrics such as Dice and Jaccard scores. Additionally, the authors performed experiments to examine the impact of input size and the number of intermediate layers on the model's performance. They found that using a gradually decreasing learning rate during training could enhance the model's performance.

In "Segmentation of glioma tumors in brain using deep convolutional neural network" [3], the author discusses the use of a deep convolutional neural network (DCNN) for the segmentation of glioma tumors in the brain. The study shows that the DCNN method achieves high accuracy and outperforms traditional segmentation methods, suggesting potential improvements to clinical practices The accuracy of the proposed DCNN method was evaluated on a set of magnetic resonance (MR) imaging data and achieved an average accuracy of around 89.4%. However, the specific accuracy may vary depending on the dataset used and the evaluation method [3].

3 Method

The input image is a brain MR image that needs to be segmented to detect brain tumors (in PNG, JPEG, or DIFF format). The image is then processed to transform it into a 256x256x3 size to be inputted into the pre-trained segmentation model for segmentation. The EfficientNet-B7 network is used for encoding and the FPN is used for decoding in the segmentation model. The output is the segmented image, also known as a mask.

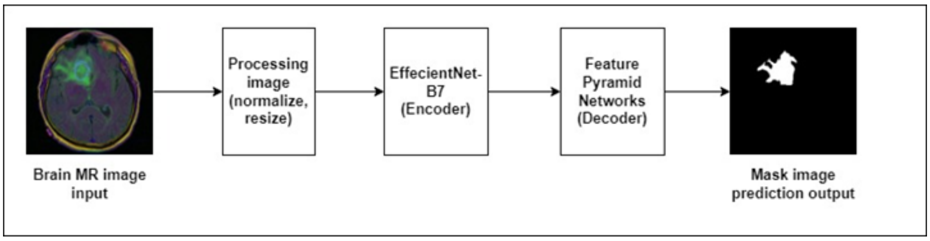
Brain MR image input
Processing image (normalize, resize)
EffecientNet-B7 (Encoder)
Feature Pyramid Networks (Decoder)
Mask image prediction output

Figure 1: Methodology diagram

4 Data Collection

We utilized the manually preprocessed MICCAI BraTS 2020 dataset [4] consisting of 110 patients. The dataset was used for training and testing of low-grade glioma brain tumor segmentation ` masks. The images in the dataset are displayed as shown in the Figure 2.

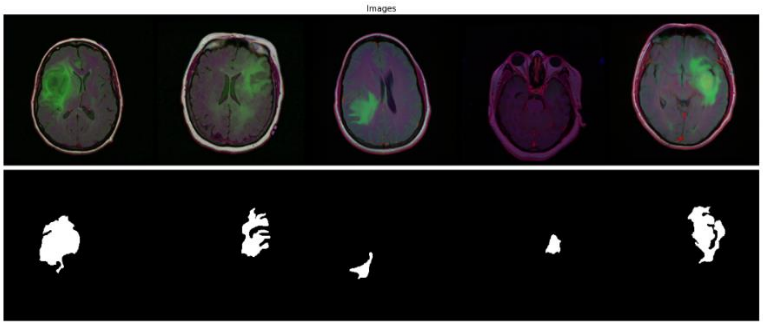

Figure 2: Image samples in the dataset

This dataset consists of 7858 images, including 3929 face images and 3929 MRI images. Additionally, the number of images with and without brain tumor corresponding to the Positive and Negative classes are 2556 and 1373, respectively. The below describes the data classification.

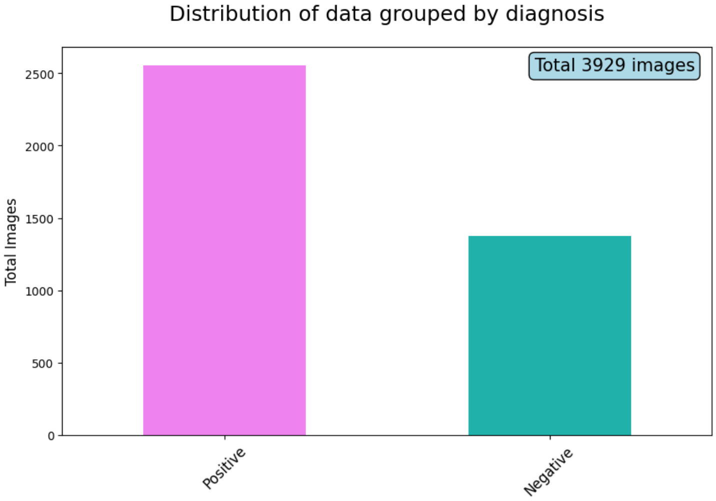

Figure 3: Description of the classification of the dataset.

The group prepared the dataset for model training by dividing it into three parts: train, validation, and test sets with the following ratios: 70% for train, 15% for validation, and 15% for test. Specifically, the number of images in each set is 2750 for train, 590 for validation, and 589 for test. This splitting into three sets is to avoid overfitting. The training set is used to train the model, where the model learns from the training data to predict new results. The test set is used to evaluate the performance of the model. After training the model on the training set, we evaluate it on the test set to see if the model performs well on new data. The validation set is used to evaluate and adjust the hyperparameters of the model. Hyperparameters are parameters not learned from data but are set by programmers or users. Examples of hyperparameters include the number of hidden layers in a neural network, learning rate, number of trees in a random forest, etc.

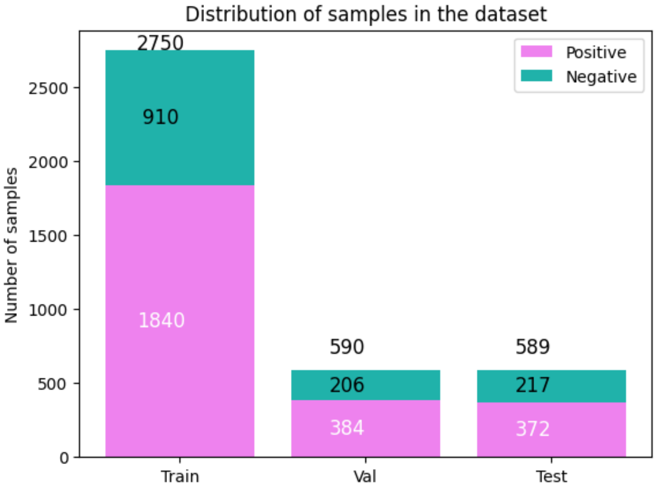

Figure 4: Description of the data split for training

They are represented as Figure 4 below with the train set having 2750 images, including 1840 Positive and 910 Negative images, the Val set having 2750 images, including 384 Positive and 206 Negative images. Similarly, the Test set has 589 images, including 372 Positive and 217 Negative images.

5 Implementation

We used EfficientNet-b7 as an encoder to extract features from the input image. FPN (Feature Pyramid Network) was used to gather information from different feature levels of the image to make the final prediction. It improves the ability to segment images by integrating information from different scales of the image, resulting in better and more accurate segmentation. Specifically, we used the segmentation-models-pytorch library to initialize a neural network model for semantic segmentation. A Fully Convolutional Network (FCN) model with a Feature Pyramid Network (FPN) architecture was built on an encoder model trained on the ImageNet dataset. The input image has a value of 256x256x3, which is the number of channels of the input image, a color image with 3 channels (R, G, B). The classes parameter was set to 1, so the model will predict a mask for each input image, with predicted values ranging from 0 to 1 at each pixel. The activation parameter was set to "sigmoid", which is the activation function applied to the output of the model to bring predicted values into the range of 0 to 1. We set the Early Stop Loss to 6 and monitored the validation loss function during the process. In this case, if the validation loss does not improve within 6 epochs, the training process will be stopped. We trained the model for 20 epochs using the Adam optimization algorithm, which is used to optimize the loss function during the machine learning model training process.

We used Colab Pro account with 15 GB GPU and 12 GB RAM for the entire training process.

6 Evaluate Metrics

In a Deep Learning training model, we can use various metrics to evaluate the quality of the model. A metric is a numerical value calculated based on the model's results on the test dataset. In our study, we use the Intersection over Union (IoU) and Dice similarity coefficient (Dice) as metrics for evaluation.

6.1 Dice

The Dice coefficient (also known as F1-score) is used in image segmentation to measure the similarity between the predicted object and the ground truth. The Dice coefficient is calculated by taking twice the intersection of the two regions and dividing it by the sum of the areas of the two regions. The Dice coefficient formula is: Dice = 2 * TP / (2 * TP + FP + FN) [5]. Both metrics have values ranging from 0 to 1, and the higher the value, the more effective the model is in segmenting the image. Figure 5 illustrates the formula for calculating IoU.

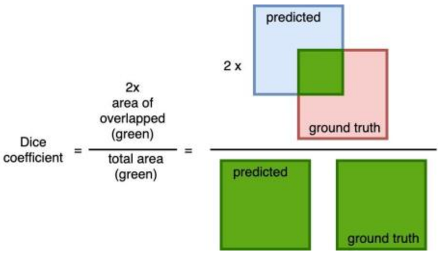

Figure 5: Illustration of Dice metric [6]

6.2 IoU

Used in image segmentation, measuring the similarity between the predicted object and the ground truth object. IoU (Intersection over Union) is commonly used metrics to evaluate the effectiveness of models in image segmentation. IoU is the ratio of the intersection area to the union area of two regions. The IoU formula is calculated as: IoU = TP / (TP + FP + FN) [5], where TP is the number of pixels correctly classified, and FP and FN are the number of pixels incorrectly classified. Figure 6 illustrates the formula for calculating IoU.

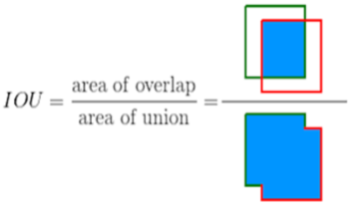

Figure 6: Illustration of IoU metric [7]

7 Results

The passage describes the training process of a segmentation model and the loss chart obtained from it. The model was stopped at epoch 13, with training loss of approximately 0.0026 and validation loss of approximately 0.0056. The chart shows a decreasing trend in both training and validation losses over time, indicating the model is learning to solve the segmentation problem effectively. However, the increasing gap between training and validation losses after the first epoch suggests the model may be overfitting, and the use of early stop loss helped stop the training process early without sacrificing too much time.

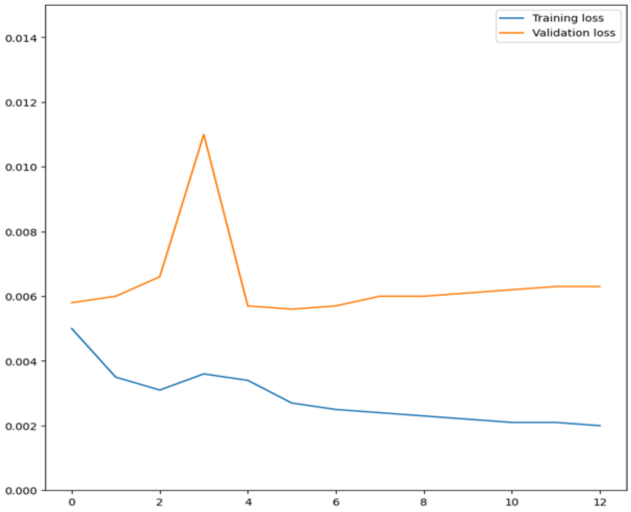

Figure 7: Loss graph

To visualize the performance of the model, the graph representing the IoU and Dice coefficient values is plotted in blue and orange, respectively. The graph is used to display the IoU and Dice coefficient values on the test set of the model during training, making it easier to evaluate the model's performance. The graph shows that the Validation Mean IoU and Validation Dice coefficient values both change during training and reach quite high values after training. This shows that the model has converged and achieved good results on the validation set. With a Validation Mean IoU value of

0.9083 and a Validation Dice coefficient value of 0.9336, the model is capable of good predictions. The graph is Figure 8.

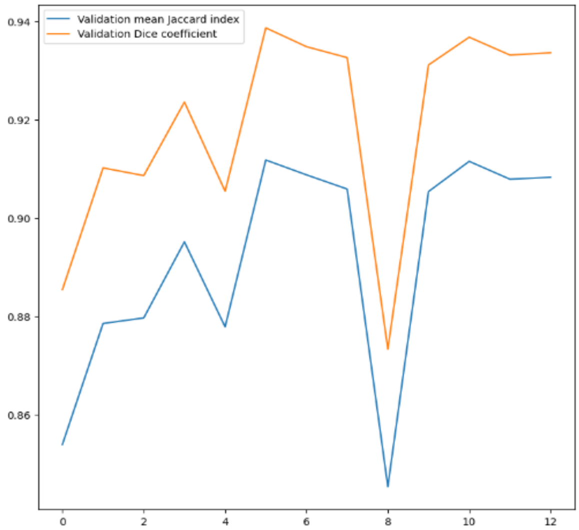

Figure 8: The Dice coefficient chart during training

The IoU and Dice metrics on the test set have a acceptable deviation compared to the validation set. This is shown in Figure 9. The results of the IoU and Dice metrics on the test set are 0.8878 and 0.9303, respectively, compared to 0.9083 and 0.9336 on the validation set.

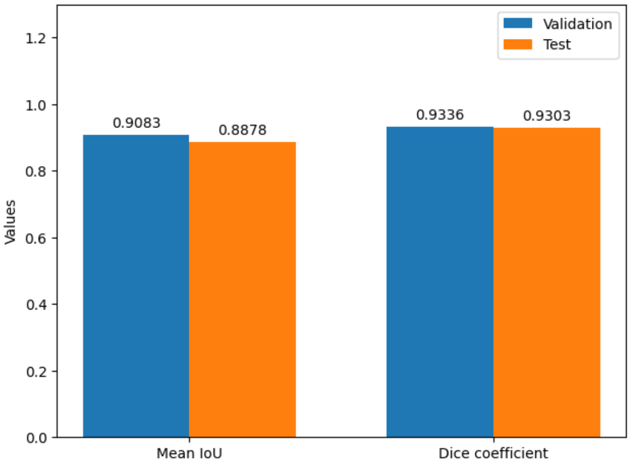

Figure 9: Comparison chart of Dice and IoU metrics between the validation and test sets.

Furthermore, they conducted experiments on real images and achieved good results with an IoU of 0.923 and a Dice coefficient of 0.96. The images below show MRI images in Figure 10, actual mask images in Figure 11 and predicted mask images drawn by the model in Figure 12.

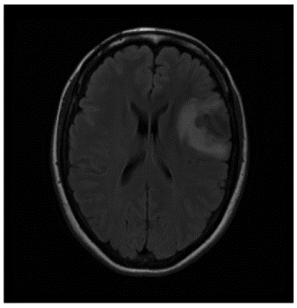

Figure 10: Brain MR image

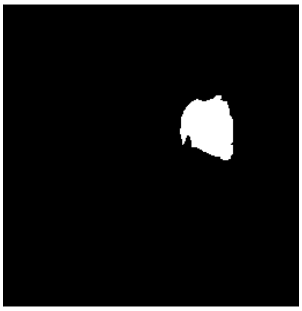

Figure 11: Origin mask image

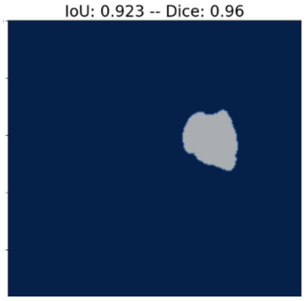

Figure 12: Prediction mask image

8 Conclusion

The EfficientNet and FPN model were trained and evaluated on the task of brain tumor segmentation. The experimental results showed that the proposed model achieved stable accuracy in brain tumor segmentation. The EfficientNet model with FPN achieved a Mean IoU of 0.8875 and a Dice coefficient of 0.9303 on the test set. Although the model's performance in brain tumor segmentation was effective, further improvement in accuracy is still needed by adding more data and adjusting model parameters. These results demonstrate the potential of using deep learning models for medical image analysis and its potential to improve the accuracy of brain tumor segmentation.

References

1. Zhihua Liu, Lei Tong, Long Chen, Zheheng Jiang, Feixiang Zhou, Qianni Zhang, Xiangrong Zhang, Yaochu Jin & Huiyu Zhou, "Deep learning based brain tumor segmentation: a survey," *Springer,* vol. 9, no. Complex & Intelligent Systems, p. 1001–1026, 2023.
2. Santosh, S. Ghosh and K. C., "Tumor Segmentation in Brain MRI: U-Nets versus Feature Pyramid Network," *EEE 34th International Symposium on Computer-Based Medical Systems (CBMS),* pp. 31-36, June 2021.
3. Saddam Hussain, Syed Muhammad Anwar, Muhammad Majid , "Segmentation of glioma tumors in brain using deep convolutional neural network," *ScientDirect,* vol. 282, no. Neurocomputing, pp. 248-261, 2018.
4. AWSAF, "Brain Tumor Segmentation(BraTS2020)," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/datasets/awsaf49/brats2020-training-data. [Accessed 4 5 2023].
5. Hanbury, Abdel Aziz Taha, Allan, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging,* vol. 15, 2015.
6. H. Jing, "Biomedical Image Segmentation: U-Net," Towards Data Science, 7 November 2019. [Online]. Available: https://towardsdatascience.com/biomedical-image-segmentation-u-net-a787741837fa. [Accessed 13 May 2023].
7. Rafael Padilla, Sergio L. Netto, Eduardo A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *2020 International Conference on Systems, Signals and Image Processing*, 2020.