Original Research Article

Classification and Segmentation of Brain Tumor Using EfficientNet-B7 and U-Net

ABSTRACT

Tumors are caused by uncontrolled growth of abnormal cells. Magnetic Resonance Imaging (MRI) is modality that is widely used to produce highly detailed brain images. In addition, a surgical biopsy of the suspected tissue (tumor) is required to obtain more information about the type of tumor. Biopsy takes 10 to 15 days for laboratory testing. Based on a study conducted by Brady in 2016, errors in radiology practice are common, with an estimated daily error rate of 3-5%. Therefore, using the application of artificial intelligence, is expected to simplify and improve the accuracy of doctor's diagnose.

Keywords: Convolutional neural network; U-Net; EfficientNet-B7; Machine learning; Brain tumor

I Introduction

Brain tumors are solid neoplasms within the skull. These tumors are caused by uncontrolled growth of abnormal cells. Classification of brain tumors is divided based on the location of the tumor, the type of tissue produced, and whether the tumor is malignant (malignant) or benign (benign) and several other considerations [1]. Magnetic Resonance Imaging (MRI) is modality that is widely used to produce highly detailed brain images. In addition, a surgical biopsy of the suspected tissue (tumor) is required to obtain more information about the type of tumor [2]. To diagnose brain tumors, doctors usually perform a biopsy and direct observation through the brain images that have been obtained. The laboratory testing process for biopsy requires more time, around 10 to 15 days, whereas manual diagnosis carries a risk of error [3]. Based on a study conducted by Brady in 2016 [4], errors and discrepancies in radiology practice are common, with an estimated daily error rate of 3-5%. So we need an alternative method that is fast and has a low error rate to help doctors make decisions.

Machine learning, the cornerstone of today's artificial intelligence (AI) revolution, is bringing a new method to clinical practice with medical images. For example, for diagnosing various conditions from medical images, machine learning has been shown to have performance on par with medical experts [5]. In practice, machine learning usually begins with a learning algorithm system by calculating the features or image augmentations used to make predictions or accurate diagnoses [6].

Transfer learning is an approach in machine learning that leverages the knowledge gained from training on one task to help solve a different task. In this paper, we use EfficientNet-B7 to classify and U-Net to segment the brain tumor. The results of the journal produce a new architecture named EfficientNet where EfficientNet-B7 achieves better or comparable accuracy results to other networks while having a smaller size and lower floating point operations per second (FLOPS). U-Net is an architecture in convolutional neural network (CNN) for medical segmentation.

Based on the description above, with the advantages of machine learning algorithms, especially the CNN method in classifying brain tumors and contour detection, this research was conducted on the performance of machine learning on MRI image results for patients with brain tumor using transfer learning EfficientNet-B7 and U-Net. In this paper, the supervised machine learning techniques are used to classify four types of abnormal brain MR Images such as glioma, meningioma, no tumor and pituitary. The performance metrics used in this paper to measure the effectiveness of machine learning in classifying brain tumors are accuracy, precision, sensitivity, and specificity. The Intersection over Union (IoU) was used to evaluate the performance of the model in detecting brain tumor segmentation, or how well it was able to detect the tumor.

II Convolutional Neural Network (CNN)

CNN is primarily used for image data due to its high network depth and effectiveness. MLP is not ideal for image classification because it does not retain spatial information from image data and treats each pixel as an independent feature, resulting in unfavorable outcomes [7]. CNN is famous these days for image recognition and image related prediction, which applied the concept of convolution from classical linear system screening. Therefore, the concept of convolution is naturally combined into a neural network (NN) to develop what is called a CNN [8].

CNN performs the convolution process by moving a convolution kernel or a filter of a certain size to an image, the computer gets new representative information from the multiplication of the part of the image with the filter used. From the results of the convolution, smaller and overlapping image fragments will be produced. Each thumbnail of the convolution results is then used as input to produce a feature representation. This gives CNN the ability to recognize an object, regardless of where the object appears in an image. The images will then be stored in an array and downsampled. Downsampling is done to reduce the size of the array so that the computation is done lighter. Downsampling is divided into two, namely max pooling and average pooling. Max pooling takes the largest pixel value in each convolution kernel layer while average pooling takes the average pixel value of each convolution kernel.

CNN has three types of layers: Convolution, pooling, and fully connected layer. The convolution layer is made up of multiple convolution kernels which learn feature representations from the input and generate feature maps. feature maps are the first result of convolution of the input with the learned kernel, and then, by applying to the convolution result and non-linear activation function.

III Transfer Learning EfficientNet

Transfer learning is a machine learning method in which a model is developed for a task that is reused as a starting point for modeling in a second task. There are many types of transfer learning that can be implemented in an ML model. Selection of the type of transfer learning according to the use or needs of the desired task [9].

EfficientNet, first introduced by Tan and Le in 2019 is one of the most efficient models that achieve state-of-the-art accuracy on imagenet and general image classification transfer learning [10],[11]. EfficientNet is a collection of models built from base models developed through a search for neural network architectures. EfficientNet's performance is better than all its predecessors with the same computation and memory requiring a large margin [12].

The EfficientNet models are constructed using simple and highly efficient compound scaling methods. This technique enables the ConvNet baseline to be scaled up to any target with limited resources while maintaining the model's effectiveness in transfer learning datasets. EfficientNet consists of versions ranging from B0 to B7, each with different parameters ranging from 5.3M to 66M. In this research we will use efficientNet-B7 which the architecture shown in figure 1[13].

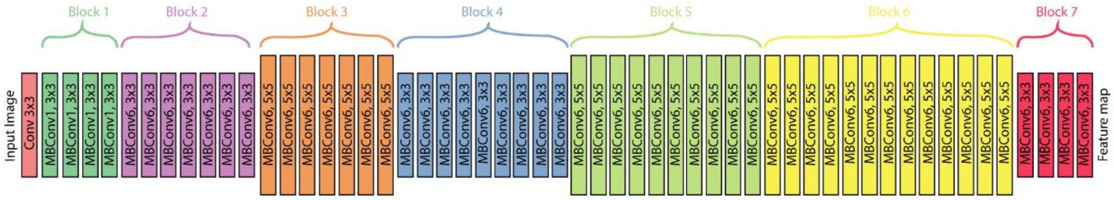


Figure 1. EfficientNet-B7 architecture

The complete design of the network can be categorized into seven distinct blocks that are represented using different colors. The fundamental component of the network is MBConv (mobile inverted bottleneck convolution), and the corresponding filter size for each MBConvX block is also displayed. The standard ReLU and ReLU6 activation functions are denoted by X=1 and X=6, respectively, for each MBConvX block.

IV UNet Architecture

U-Net is one of the usual types of CNN used for semantic image segmentation [14]. U-Net consists of the downsampler and upsampler. on layers downsampler, there are two 3x3 unpadded convolutions, which will then go through the ReLU activation function and a 2x2 layer maxpool with stride 2. On each stage of downsampling will occur doubling feature. On the upsampler layer, there is an upsampling process followed by a 2x2 convolution layer which will split the number of features in half, process feature pooling, two 3x3 convolution layers, each with a ReLU activation function. Layers finally in the form of 1x1 transposition convolution layer which used to transform the 64 component feature vectors into 3 classes [15]. Full architecture of U-Net shown in figure 2.

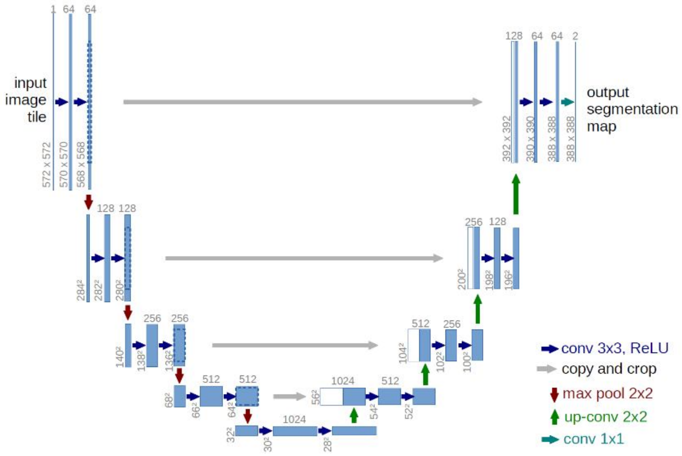


Fig 2 U-Net Architecture [14].

V Materials and Methods

Research on the classification and contour detection of brain tumors on MRI imaging results using machine learning algorithms was conducted at the Biophysics and Medical Physics Laboratory of the Physics Study Program at Udayana University with data taken from Kaggle containing 7022 brain tumor images with brain tumor classifications glioma, meningioma, no tumor brain and pituitary [16]. This data contains data taken from the Plos One article titled "Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition" by Jun Cheng, et al from the school of Biomedical Engineering Southern Medical University, Guangzhou, China [17],[18]. To evaluate the brain tumor classification model, data obtained from The Cancer Imaging Archive under the auspices of the University of Arkansas for Medical Science under contract from the National Cancer Institute (NCI) were used. Meanwhile, the dataset for detecting brain tumor segmentation was obtained from The Cancer Genomic Atlas (TCGA) [19].

This research was carried out utilizing the Python programming language and the Google Colaboratory code editor. The experiment was performed on a laptop equipped with an Intel(R) Core(TM) i7-4700HQ processor and 8 GB of RAM.. Model in this study uses several libraries as modules that contain a set of code that can be used repeatedly in different programs. After importing the model supporting library in this study, data preprocessing was carried out. Data preprocessing is the process of entering or combining data into the dataset model to be created. The dataset was obtained from Kaggle which contains 7022 MRI images with the classification of brain tumor images of glioma, meningioma, pituitary and no tumor. For segmentation detection of brain tumors using the UNet architecture, the dataset used was obtained from The Cancer Genome Atlas (TCGA) with images taken from 110 patients. At the data preprocessing stage, the data is divided into training and testing. After preprocessing the data, the process of making a machine learning model is carried out. At this stage, hyperparameters are determined according to the existing data and problems. After making the model is complete, the model training process is carried out to improve the model's ability to classify brain tumor image data according to the label that has been given. The next process is to evaluate the results of model training. If the model accuracy can still be improved, it can be improved by optimizing the hyperparameters used in machine learning modeling. The research procedure is shown in Figure 3.

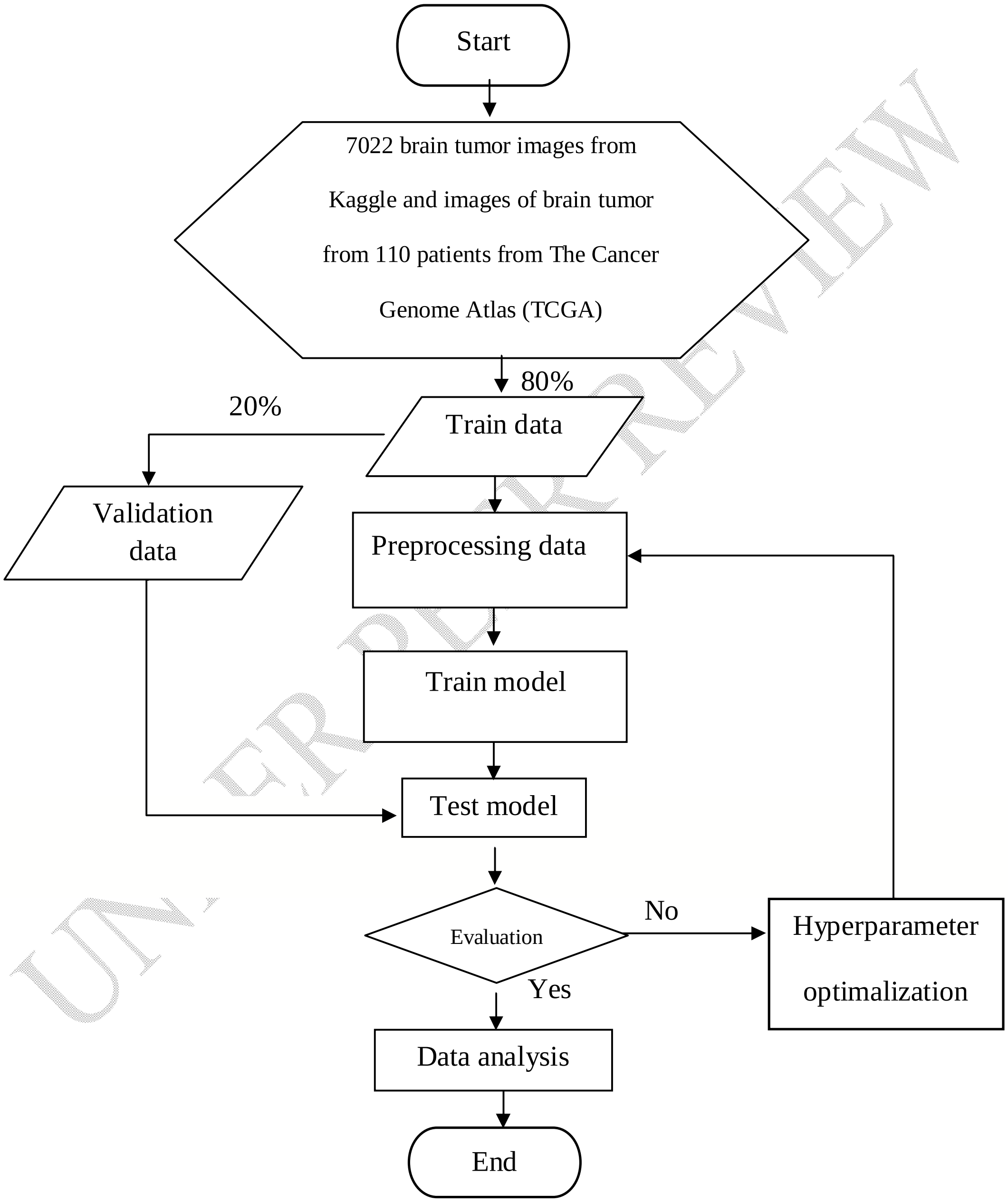


Figure 3 Research procedure

VI Result

Model Architecture

The CNN model shown consists of a base model that uses EfficientNetB7 transfer learning shown in figure 4. EfficientNetB7 has the same performance well with limited computing needs. This transfer learning has 66 million parameters for classifying and analyzing images. The model consists of 3 layers, with the first layer consisting of the Convolution2D layer, the Convolution2D layer is used for data in the form of images, this function is to recognize images based on the pixels contained in the image, using the Relu activation function. At this stage, 16 filter parameters are loaded, meaning that the number of filters issued in the convolution process is 16 with a size of 3x3 (kernel_size = (3, 3)). For the 2nd layer, 32 filter parameter are loaded. For 3rd layer, 64 layer are loaded and for 4th layer, 128 filter layer are loaded. Then input_shape for all layer are (150, 150, 3) this means that all image sizes that enter the input layer are 150x150 pixels in size and 3 color layers in matrix form. The activation used for all layer are the Rectifier Linear Unit (ReLu) because this function is a function commonly used for CNN. Furthermore, the MaxPooling2D layer is added which is used to reduce the image resolution while maintaining the information in the image, the way the MaxPooling2D layer works is to take the maximum value from the processed data. Where the pooling size on the first layer is 2x2. Thus, the analysis will be faster due to the small size of the feature maps.

Afterwards, flatten the matrices using the command. By doing so, the matrices will be converted into a single vector with a certain size that can be utilized as input for the neural network in the future. The next step is to give dense or add a hidden layer. The first hidden layer is 1024 in size, which means that there are 1024 neurons in this layer. Then the next parameter is activation='relu'. Next is used hidden layer 512.

Dropout is used to reduce overfitting in the model. Then next is the definition of the output layer. Because the data type is categorical (glioma, meningioma, pituitary, and no tumor) there are 4 neurons according to the number of labels in the dataset.

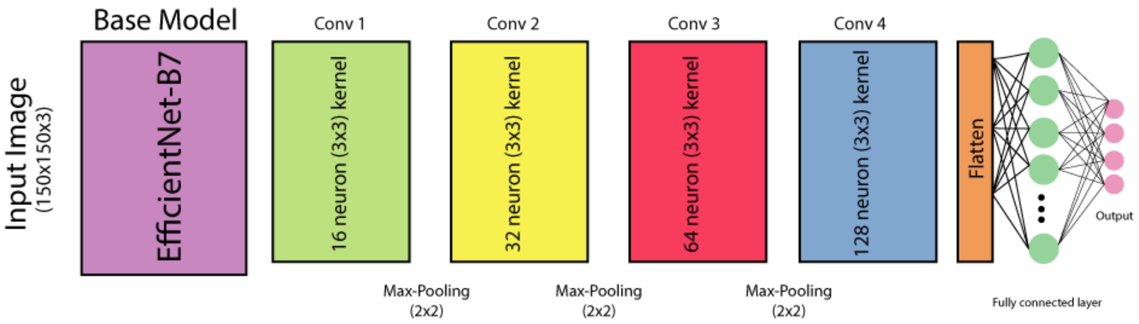


Figure 4. Model architecture

The UNet architecture in the model for detecting brain tumor contours uses 10 convolution layers. In layer 1 Conv2D with 64 filters and a kernel size of 3×3 is used to convolve the input image with the output being a feature map with a smaller size. The activation layer used is ReLU, which functions to provide non-linearity to feature maps which are the output of the previous layer. BatchNormalization is a batch normalization layer on the 3rd channel (RGB). This layer functions to speed up training and model stabilization by normalizing each batch of data that is input to the layer. Activation is an activation layer with the ReLU activation function. This layer is used to provide non-linearity to the normalization results generated by the BatchNormalization layer. MaxPooling2D is a pooling layer with a pool size of 2x2. This layer serves to reduce the spatial dimension of the feature map so that the information retrieved becomes more focused and not too much. For layer 2 Conv2D uses 128 filters and a kernel size of 3 × 3. With BatchNormalization, the same activation and Maxpooling2D functions as layer 1 to layer 9. Layer 3 uses Conv2D with 256 filters and the same kernel size as the other layers. Layer 4 uses 512 Conv2D filters with the same kernel size as the previous layer. On layer 5 use 1024 Conv2D filters with the same size as the previous layer. In layer 6, the concatenate layer is used which is a layer combining feature maps from two layers, namely the Conv2DTranspose layer and the Conv4 layer. This layer functions to combine the upsampled feature map with the feature map on a lower layer horizontally, thereby adding spatial information and enriching the features used for segmentation. In addition, the Conv2DTranspose layer is also used which is a convolution transpose layer with 512 filters and a kernel size of 2x2 and a stride of 2x2. This layer functions to carry out convolution transpose operations (upsampling) on the feature map produced by the previous layer so as to produce a larger feature map. Layer 6 also uses a Conv2D layer with a total of 512 filters with a kernel size of 3 × 3. Layer 7 just like layer 6 also uses Conv2DTranspose and a concate layer. At layer 7, it uses 256 filters with the same kernel size as the other layers. Layer 8 uses 128 filters with the same kernel size as the other layers, layer 8 also uses Conv2D Transpose and concanate layers with the same parameters as layers 6 and 7. Layer 9 uses 64 filters with the same kernel size as the other kernels. Layer 9 also uses the Conv2D and concanate layers. Layer 10 uses 1 filter with a kernel size of 1 × 1. Layer 10 uses sigmoid activation which is used to determine the probability of an input data being included in the target class or not.

Confussion matrix

Figure 5 is the result of multiclass accuracy. The figure shows the number of comparison results from machine learning algorithm predictions with actual results (ground truth). From these images obtained 1291 true positive data for Glioma, 1148 true positive data for meningioma, 1550 true positive data for no tumor and 1446 true positive data for the pituitary.

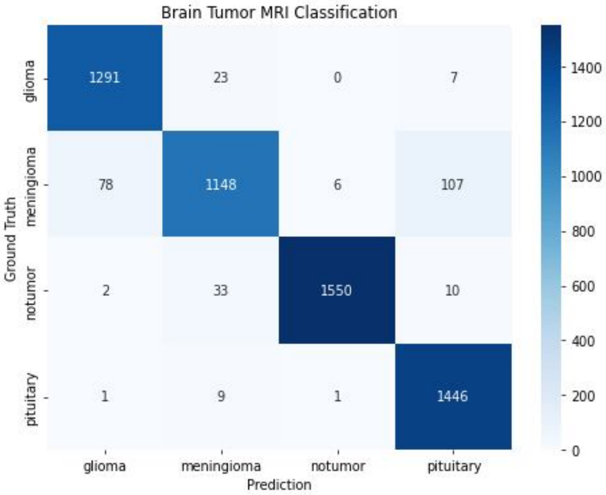


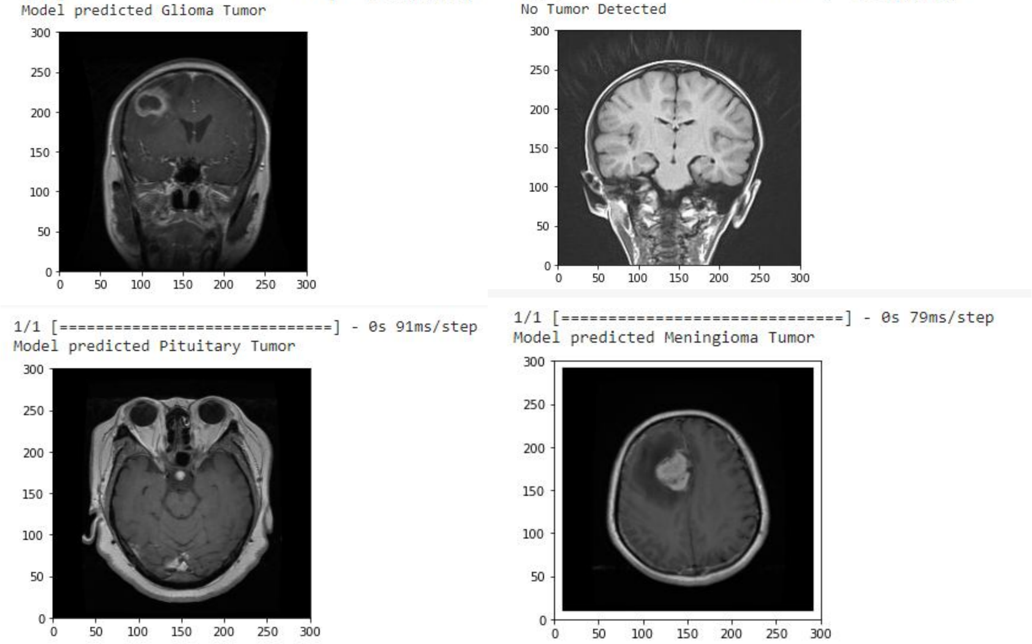
Figure 5. Confussion matrix result

Accuracy, Sensitivity, Precision and specificity

Table 1. Table of results for accuracy, sensitivity/recall, precision, specificity

	Precision	**Sensitivity/recall**	***Specificity***
Glioma	0.94	0.98	0.99
Meningioma	0.95	0.86	0.98
No Tumor	0.99	0.97	1.00
Pituitary	0.92	0.99	0.97
Accuracy			0.95

The results for glioma labeling have a precision value of 0.94, a sensitivity of 0.98 and a specificity of 0.99. For meningioma labels obtained a precision value of 0.95, a sensitivity value of 0.86 and a specificity of 0.98. If there is no tumor label, the precision value is 0.99, the sensitivity is 0.97 and the specificity is 1.00. For the pituitary label, a precision value of 0.92 was obtained, a sensitivity of 0.99 and a specificity of 0.97. With a total overall accuracy of 0.95.



```
1/1 [==============================] - 0s 105ms/step
Model predicted Glioma Tumor
```

```
1/1 [==============================] - 0s 53ms/step
No Tumor Detected
```

```
1/1 [==============================] - 0s 91ms/step
Model predicted Pituitary Tumor
```

```
1/1 [==============================] - 0s 79ms/step
Model predicted Meningioma Tumor
```

Figure 6. Result of Predicted Tumor

The output of brain tumor classification is shown in Figure 6. The classification of brain tumors defined in this model were glioma, meningioma, pituitary brain tumor and no tumor was detected. The results of the prediction of the MRI image are in the form of an image measuring 150×150 with 3 RGB color channels. The results of this classification prediction obtained an overall accuracy of 95%.

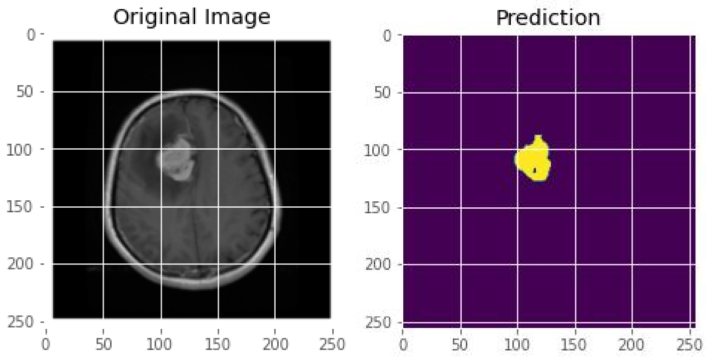


Figure 7. Result of Brain Segmentation

In Figure 7, the model predicted results of the tumor segmentation portion in the original image. The value of Intersection over Union (IoU) is 0.99. Intersection over union or also known as the Jaccard index is an evaluation metric used to measure how much the segmentation or detection object produced by a model matches the actual object in the image. This metric measures the overlap between two areas, namely the prediction area (the result of segmentation or detection from the model) and the reference area (the actual object that is expected to appear in the image). IoU is calculated by dividing the overlapping area by the combined area of the two areas.

V Conclusion

The objective of this study is to improve the accuracy of brain tumor diagnosis and segmentation, with the aim of aiding physicians in identifying specific brain tumors. By implementing transfer learning with the EfficientNetB7 model for brain tumor classification, we achieved an accuracy of 95%. For brain tumor segmentation using the UNet architecture, we achieved an Intersection over Union value of 99%. The use of EfficientNetB7 and UNet for brain tumor classification and segmentation, respectively, can effectively improve brain tumor diagnosis.

References

1 M.P. Narendran, V.K. Narendira Kumar, K Somasundara. "3D brain tumors and internal brain structures segmentation in Mr Images," *International Journal of Image, Graphics and Signal Processing*, vol. 4, no. 1, pp. 35–43, 2012.

2 A. M. Alqudah, "Brain tumor classification using deep learning technique - A comparison between cropped, uncropped, and segmented lesion images with different sizes," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3684–3691, 2019.

3 R. A. Rachmad, B. Wahyu, and R. Purbaningtyas, "Klasifikasi Tumor Otak Menggunakan Convolutional Neural Network dengan Arsitektur EfficientNet-B3," *Jurnal Sistem Informasi, Teknologi Informasi dan Komputer*, vol. 11, no. 3, pp. 55–59, 2021.

4 A. P. Brady, "Error and discrepancy in radiology: Inevitable or avoidable?," *Insights into Imaging*, vol. 8, no. 1, pp. 171–182, 2016.

5 G. Varoquaux and V. Cheplygina, "Machine Learning for Medical Imaging: Methodological Failures and recommendations for the future," *npj Digital Medicine*, vol. 5, no. 1, 2022.

6 B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine Learning for Medical Imaging," *RadioGraphics*, vol. 37, no. 2, pp. 505–515, 2017.

7 W. S. Eka Putra, A. Y. Wijaya and R. Soelaiman "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) Pada Caltech 101," *Jurnal Teknik ITS*, vol. 5, no. 1, 2016.

8 S. Cui, H. H. Tseng, J. Pakela, R. K. Ten Haken, and I. El Naqa, "Introduction to machine and Deep Learning for medical physicists," *Medical Physics*, vol. 47, no. 5, 2020.

9 I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. The MIT Press, 2016.

10 K. Team, "Keras documentation: Image Classification via fine-tuning with EfficientNet," *Keras*, 2020. [Online]. Available: https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning/.

11 M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning Research*, vol. 97, pp. 6105–6114, 2019.

12 T. A. Putra, S. I. Rufaida, and J.-S. Leu, “Enhanced skin condition prediction through machine learning using dynamic training and testing augmentation,” *IEEE Access*, vol. 8, pp. 40536–40546, 2020.

13 M. Ibrahim Khalil, S. Tehsin, M. Humayun, N. Z. Jhanjhi, and M. A. AlZain, “Multi-scale network for thoracic organs segmentation,” *Computers, Materials & Continua*, vol. 70, no. 2, pp. 3251–3265, 2022.

14 O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *Lecture Notes in Computer Science*, pp. 234–241, 2015.

15 M. S. Lui, E. K. Wijaya, and M. Hidayat, “Segmentasi citra Hewan dengan convolutional neural network Arsitektur U-NET,” 2014. [Online]. Available: http://yuita.lecture.ub.ac.id/files/2021/12/Tugas-4-Kelompok-8.pdf. [Accessed: 20-Feb-2023].

16 M. Nickparvar. *Brain Tumor MRI Dataset*. [Data set]. Kaggle, https://doi.org/10.34740/KAGGLE/DSV/2645886, K. M. Schmainda, M. Prah. Data from Brain-Tumor-Progression. *The Cancer Imaging* Archive, 2018, https://doi.org/10.7937/K9/TCIA.2018.15quzvnb.

17 J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, “Enhanced performance of Brain Tumor Classification via tumor region augmentation and partition,” *PLOS ONE*, vol. 10, no. 10, 2015.

18 K. M. Schmainda, M. Prah. Data from Brain-Tumor-Progression. *The Cancer Imaging* Archive, https://doi.org/10.7937/K9/TCIA.2018.15quzvnb, 2018.

19 N. Pedano, A.E. Flanders, L. Scarpace, T. Mikkelsen, J. M. Eschbacher, B. Hermes, V. Sisneros, J. Barnholtz-Sloan, and Q. Ostrom, “The Cancer Genome Atlas Low Grade Glioma Collection (TCGA-LGG) (Version 3)” [Data set]. *The Cancer Imaging Archive*, https://doi.org/10.7937/K9/TCIA.2016.L4LTD3TK, 2016.