Recursive Journal of Informatics
Vol. 4, No. 1, March 2026

p-ISSN 2963-5551 https://journal.unnes.ac.id/journals/index.php/rji e-ISSN 2986-6588

Improving Brain Tumor Image Segmentation Accuracy Based on Residual Network (ResNet) Using Nearest Neighbor Upsampling

Muhammad Afifudin[1*], Endang Sugiharti[2]
[1,2] Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang, Indonesia

Abstract. Brain tumors are a critical disease due to the abnormal growth of cells in the brain, which can damage surrounding normal cells and increasing the risk of death in patients. With advancements in technology, artificial intelligence can be developed to segment brain tumor areas, aiding medical professionals in identifying tumor characteristics and determining appropriate treatment plans. Convolutional Neural Network (CNN) models can be utilized for segmentation tasks because their ability to classify each pixel of an image, assign specific labels, and map them into homogeneous groups. To enhance the capability of CNNs against the possibility of vanishing gradients, the Residual Network (ResNet) architecture can be applied to the segmentation model. The use of ResNet provides additional capability for the network to choose between the training results in the current epoch or skip to the next network when the training results approach the identity value. However, ResNet also reduces the scale of images and feature maps during downsampling operations, sacrificing spatial resolution. This study proposes the implementation of the Nearest Neighbor Upsampling method on ResNet to improve the model's accuracy in the task of MRI brain tumor segmentation.

Purpose: This research proposes a method to increase the accuracy of brain tumor MRI image segmentation using the ResNet model by implementing the Nearest Neighbor Upsampling method.

Methods/Study design/approach: The method used is Nearest Neighbor Upsampling on ResNet to enhance image dimensions and fill gaps in MRI brain tumor images during the learning process, preserving spatial information and context crucial for segmentation.

Result/Findings: The optimization of the brain segmentation model for classifying brain tumor regions using ResNet and Nearest Neighbor Upsampling achieves an increase in accuracy from 96.94% to 98.44% and a decrease in loss value from 0.0881 to 0.0874.

Novelty/Originality/Value: This research addresses the limitations of the Residual Network (ResNet) model, such as the drawback of reducing the scale of images and feature maps during downsampling, resulting in a loss of spatial resolution. To overcome this challenge, the study introduces the Nearest Neighbor Upsampling method applied to ResNet, demonstrating its effectiveness in representing spatial information and image context, thereby improving segmentation accuracy.

Keywords: Residual Network (ResNet), Nearest Neighbor Upsampling, Image Segmentation, Brain Tumor
Received August 2024 / **Revised** March 2026 / **Accepted** March 2026

INTRODUCTION

Brain tumors are a critical health concern due to the abnormal growth of cells within the brain, which can damage surrounding normal cells and increasing the risk of death in patients. The World Health Organization (WHO) estimated that nearly 10 million cancer deaths occurred in 2020 [1]. One of the organs of particular concern when tumor cells begin to develop is the brain, as tumor growth can significantly impact the overall functionality of the body. Identifying the characteristics of brain tumors is crucial for planning treatment and management strategies for patients.

Advancements in science and technology have spurred the development of artificial intelligence (AI) for the identification of brain tumors through magnetic resonance imaging (MRI) analysis. MRI data, consisting of complex and unstructured images, necessitates the use of adaptive and dynamic AI approaches. Consequently, deep learning has become the preferred method for image detection due to its capability to deeply understand data and process it to generate valuable information [2]. The task of

[1]*Corresponding author.
Email addresses: muhafifudin2306@students.unnes.ac.id (Afif)
DOI: 10.15294/rji.v4i1.12010

identifying brain tumors through MRI images can be classified under semantic segmentation in computer vision. This process involves classifying each pixel in the image and assigning specific labels, resulting in a map of homogeneous groups. Convolutional Neural Network (CNN) models are often utilized for such tasks due to their ability to analyze and recognize specific features within images [3]. Various CNN architectures can be employed for image segmentation tasks, including Residual Network (ResNet), U-Net, and VGG-19.

Residual Network (ResNet) architecture can be chosen for image segmentation due to its ability to mitigate the vanishing gradient problem, which occurs when the gradient diminishes exponentially during backpropagation that hindering effective training. ResNet provides the additional capability for the network to choose between the training results in the current epoch or skip to the next network when the training results approach the identity value.

This facilitates more efficient gradient flow during training and helps overcome the vanishing gradient problem. However, in some cases such as medical image detection, the use of ResNet may result in the loss of spatial details within feature maps. This is supported by a study conducted by Liu et al. (2022), which observed that the process of downsampling during the use of ResNet in segmentation tasks led to a reduction in image resolution and computational memory load, sacrificing spatial resolution at the pixel level [4]. This loss of fine or local pixel details can pose a significant challenge if not addressed through additional methods to preserve these details.

ResNet Architectural development for image segmentation tasks has been explored by Shehab et al. (2021) in their study titled “An Efficient Brain Tumor Image Segmentation Based on Deep Residual Networks (ResNets).” The researchers enhanced the ResNet architecture with several additional treatments to improve the performance of brain tumor MRI image segmentation [5]. The first treatment was applied during the preprocessing stage, utilizing the N4ITK method to address bias field distortion issues in MRI images. The processed images were then used by the convolutional and identity blocks in ResNet to capture features at various levels of abstraction. The second treatment involved downsampling, incorporating average pooling and dense layers to increase output accuracy according to the number of classes. The results of this study demonstrated impressive accuracy, reaching 92%, with a training time of 62 minutes.

The problem of losing spatial details and feature maps during the downsampling process in ResNet can be mitigated using the upsampling method. This method increases the image dimensions and fills empty gaps in the image columns/rows, allowing for better representation of spatial information and context [6]. Upsampling can be performed using various algorithms, one of which is the Nearest Neighbor Interpolation method [7]. This method is noted for its high processing speed and low computational complexity, minimizing potential errors during scaling operations.

Based on the aforementioned issues, this study focuses on examining the impact of the Nearest Neighbor Upsampling method on the challenges of residual network learning in the task of MRI brain tumor segmentation. This research aims to explore the benefits and effects of this modification in greater depth. This study is expected to contribute valuable insights to the development of technology, particularly in the early diagnosis of brain tumor diseases before subsequent treatment steps are undertaken.

PROPOSED METHOD

The segmentation model in this research relies on two main principles, there are an encoder layer that performs downsampling to extract relevant features from the image, and a decoder layer that performs upsampling and combines information from the encoder [8]. In this study, the encoder layer is constructed with multiple residual blocks, which are used to extract and learn features as the network depth increases. Not only have encoder layer, but the model segmentation also has decoder layer consists of multiple upsampling blocks, that restore image resolution and combine high-level context with local details. Figure 1 illustrates the encoder-decoder model used in this research.

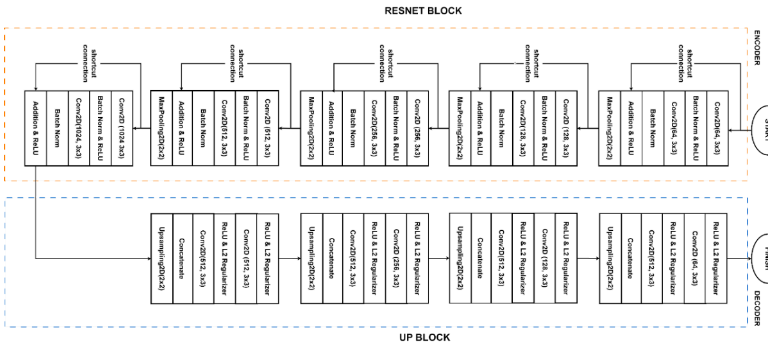

Figure 1. Image Segmentation Encoder and Decoder Model

Residual Network Downsampling

The Residual Network (ResNet) approach is built with a path that connects the input of the previous block directly to the output without any transformation that allows the gradient flow of the original information and fixes the bottleneck that usually occurs in training models with significant depth [9]. The presence of a ResNet allows the network to choose between learning the necessary changes to improve performance or learning to skip the transformation operation if the result is close to the identity value. Each ResNet block consists of two convolutional layers with batch normalization and ReLU, plus a shortcut connection to aid gradient flow. If dimensions differ, the shortcut adjusts them, and the final output is passed through ReLU.

Nearest Neighbor Upsampling

The decoder layer consists of multiple upsampling blocks, that restore image resolution and combine high-level context with local details [10]. Each up_block includes an upsampling operation using Nearest Neighbor Interpolation to increase feature representation resolution. The upsampled output is then combined with corresponding feature representations from the encoder layer and passes through two convolutional layers designed to extract features. These layers include ReLU activation and L2 regularization to learn richer feature representations and prevent overfitting. The output from the up_block is used in the decoding process to produce the final segmentation result. This block is vital in upsampling, merging high-level context with local details in image segmentation tasks.

METHODS

Research Approach and Design

This research will be focuses on the comparison of the accuracy Residual Network (ResNet) with and without Nearest Neighbor Upsampling for MRI brain tumor segmentation. The process includes data input and splitting, preprocessing, segmentation, and model evaluation, as shown in Figure 2.

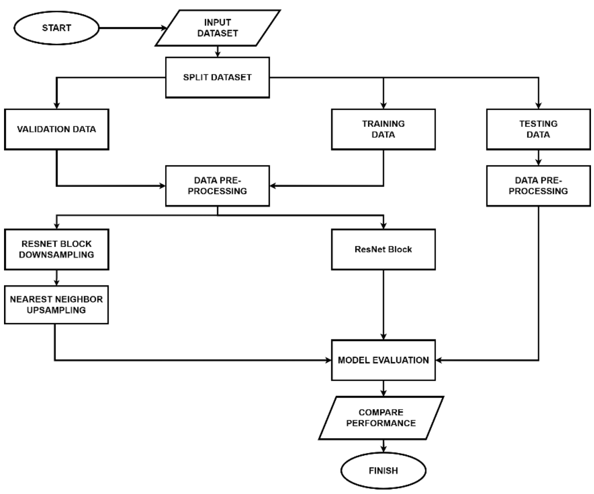

Figure 2. Flowchart of research processes

Dataset

This study uses the Brain Tumor Segmentation (BraTS) 2020 dataset, which includes 1,845 training samples consisting of four MRI modalities: T1, T1ce, T2, and FLAIR. The tumor annotations cover four subregions such as necrotic tumor core (TC), peritumoral edema (WT), enhancing tumor (ET), and non-tumor areas as background. Each image has dimensions of 240 × 240 pixels, with a total of 155 slices in each 3D scan [11]. The BraTS 2020 dataset was downloaded directly from the official website of the Center for Biomedical Image Computing & Analytics (CBICA).

Data Preparation

The research begins with the inclusion of the dataset that will be used during the network training process. Data splitting is performed first to separate the dataset into training data, validation data, and test data. The entire dataset will be split with a ratio of 65% training data, 15% validation data, and 20% test data. This separation helps ensure that the model not only learns from the training data but is also continuously evaluated during training (validation) to maintain the model's quality and generalization [12].

Data Preprocessing

Training a model on a large dataset requires significant computational power and time. Therefore, it is important to perform data preprocessing or a series of steps or techniques to help the network generalize and accelerate convergence in the neural network [13]. Some of the data preprocessing steps in this research that will be used image resizing, pixel normalization, and several image augmentation methods, such as rotation, shift, flip, zoom, and shear.

Training and Evaluation Model

The next stage after defining the model is training and evaluating the segmentation model. Training will be done by applying a certain number of epochs and batches until the iteration process is complete. In addition, callbacks objects are also implemented to monitor model training and perform certain actions on a condition in training. Pseudocode 1 describes the process of training an image segmentation model.

Pseudocode 1. Image Segmentation Model Training Illustration

Image Segmentation Training
Input: data training, data validation, and segmentation model
Output: trained segmentation model
Begin:

1. Classify image pixels to distinguish between pixels with phenomenon and pixels without tumor
2. Assign image pixels with phenomenon to label variables and ignore pixels without tumor
3. Use segmentation model to classify label pixels into 4 segmentation labels (0: 'not tumor', 1: 'necrotic/core', 2: 'edema', 3: 'enhancing') with identification image pixel process.
   for each pixel in the image:
       if pixel_value >= threshold: segmentation_label = low_segmentation_label
     else: segmentation_label = high_segmentation_label
     endif
     endfor
4. Train the model using fit method with specified parameters such as epochs, batch size, callbacks, learning rate, and data validation.
5. Return the model as output from the trained segmentation model

End

Accuracy is calculated after model evaluation and serves as a reference for research success. It measures the percentage of correctly classified image pixels [14]. The segmentation accuracy ratio is obtained from the True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP) value operations. Pseudocode 2 shows the steps in processing image segmentation accuracy.

Pseudocode 2. Segmentation Accuracy Processing Steps

Image Segmentation Evaluation
Input: testing data and trained segmentation model
Output: accuracy (A)
Begin:

1. Call the trained segmentation model to be evaluated with specified parameters such as epochs, learning rate, callbacks, and testing data.

2. Define TP, TN, FN, and FP from the evaluation of the predicted class and the actual class.
3. Calculate the accuracy value with the formula:
   accuracy (A) = $\frac{TP+TN}{TP+FN+TN+FP}$
4. Return accuracy as output for model evaluation.

End

RESULT AND DISCUSSION

This research uses a quantitative approach to analyze and train the segmentation model in a sequential and systematic manner. The following is an explanation of the research results that have been carried out:

Dataset Preparation

The image data was obtained from a previously prepared dataset. The data used in this research contains 4 types of images, such as flair, t1, t1ce, and seg images. Flair images are structural brain anatomy images to determine the location of the tumor more clearly. T1-weighted (T1) images are images that show high contrast in brain tissue, which are useful for seeing changes in the brain. T1-weighted Contrast Enhanced (T1ce) images are images resulting from the administration of contrast to facilitate the identification of tumor activity areas. Seg (segmentation) images are useful for confirming and evaluating the results of brain tumor segmentation that has been carried out [15]. Figure 3 shows some results from image sampling on the dataset.

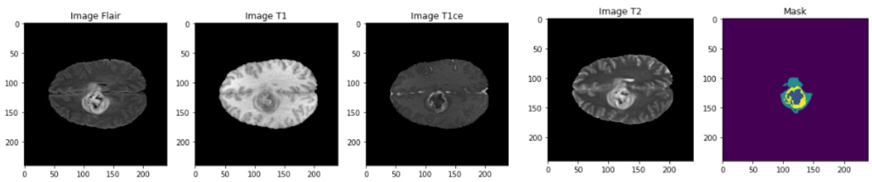

Figure 3. Sample Image from Dataset

In the data preparation stage, the number of slices to be taken from the image volume is also defined. Each brain tumor MRI image in the dataset has 155 slices. In this case, there is a limitation on the number of slices taken to maximize the potential of the image data and not put too much strain on the computational memory. The results of several experiments showed that the data slice starting from the 55^{th} volume can already facilitate well for the needs of the model and further analysis.

Data Splitting

The dataset used in this research has 1845 image data, which be split into training, validation, and testing dataset with 65%, 20%, and 15% ratio. Through this ratio, 1199 image data will be used as training dataset, 369 data will be used as validation dataset, and 277 data will be used as testing dataset. Table 1 shows results from data splitting on the dataset.

Table 1. Data Splitting Result

Type Data	Quantity
Training Dataset	1199
Validation Dataset	369
Testing Dataset	277
Total	1845

Data Preprocessing Results

The data preprocessing stage involves steps to enhance data quality and relevance before using it as input for the segmentation model. This research applies preprocessing methods such as image resizing, pixel value normalization, and image augmentation. Image resizing standardizes image dimensions and reduces model processing load, adjusting the original 240 × 240 pixel images to 128 × 128 pixels. This approach, inspired by Ali and Agrawal (2023), is essential for improving segmentation performance [16]. Additionally, pixel value normalization scales pixel values between 0 and 1, aiding in model convergence and ensuring consistency across images, as emphasized by Daimary et al. (2020) [17]. Furthermore, data augmentation is used in this research to increase the data variation in the dataset by creating variations from existing samples. Through this method, the segmentation model can become more adaptive in dealing with different variations in the data training process. The augmentation methods in this research include various transformations such as rotation, shift, flip, zoom, and shear. Figure 4 shows some results from data preprocessing results.

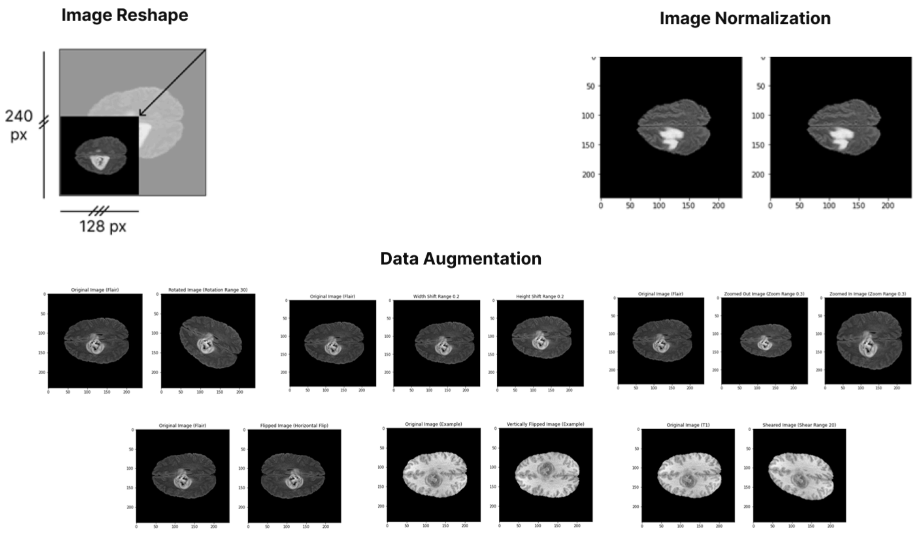

Figure 4. Results from data preprocessing results

Model Training and Evaluation Results

The two models are also compared in terms of loss and accuracy graphs during the data training process. Model 1 is built using ResNet-50 with three experimental variations: 15 epochs, 35 epochs, and 50 epochs. This model achieves the best accuracy at epoch 50 with a loss value of 0.0881 and an accuracy of 96.94%. Model 2 is built using ResNet optimized with Nearest Neighbor Interpolation with three experimental variations: 15 epochs, 35 epochs, and 50 epochs. This model achieves the best accuracy at epoch 50 with a loss value of 0.0875 and an accuracy of 98.44%.

Table 2. Data Splitting Result

	validation accuracy			validation loss		
	15	35	50	15	35	50
Model 1 (ResNet)	96,99%	97,6%	96,94%	0.0981	0.0906	0.0881
Model 2 (Proposed Method)	97,01%	97,8%	98,44%	0,1254	0,097	0,0874

Segmentation Evaluate

After training and testing the model, the next step is to perform image segmentation testing using the test data. This testing measures the model's ability to identify and separate brain tumor regions. An in-depth analysis of these results provides insights into the accuracy and precision of the segmentation achieved by Model 1 and Model 2. Test samples will be used to evaluate the brain tumor segmentation results for sequence images using both models. Figure 5 shows a sample segmentation result of a brain tumor region using Model 1.

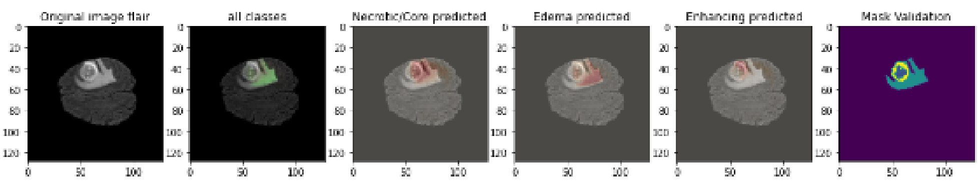

Figure 5. Segmentation Results with Model 1

In the case, Model 1, using ResNet without Nearest Neighbor Upsampling as the core architecture, demonstrates the ability to classify tumor regions. However, many fine details in the image remain unidentified. While segmentation testing using Model 2, which optimizes ResNet with Nearest Neighbor

Upsampling, successfully classifies brain tumor regions and enhances image detail, resulting in more optimal. Figure 6 shows a sample segmentation result of a brain tumor region using Model 2.

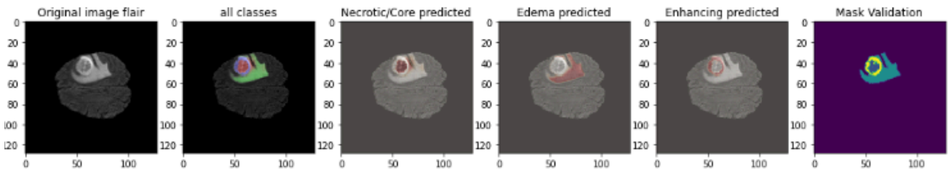

Figure 6. Segmentation Results with Model 2

Compare Other Model

In evaluating the results of this research, it was found that optimizing the ResNet-based MRI brain tumor segmentation model with the Nearest Neighbor Upsampling method achieved a high level of accuracy. Specifically, this model reached a maximum accuracy of 98.44%. In addition to comparing accuracy with related studies, the evaluation also included a comparison of segmentation results with other inspiring methods, such as UNET-VGG16 [18], UNET-DenseNet121 [19], and UNET-EfficientNetb2 [20]. Figure 7 shows a graph comparing the validation accuracy of the ResNet model optimized with Nearest Neighbor Upsampling against the three other methods over 50 training epochs.

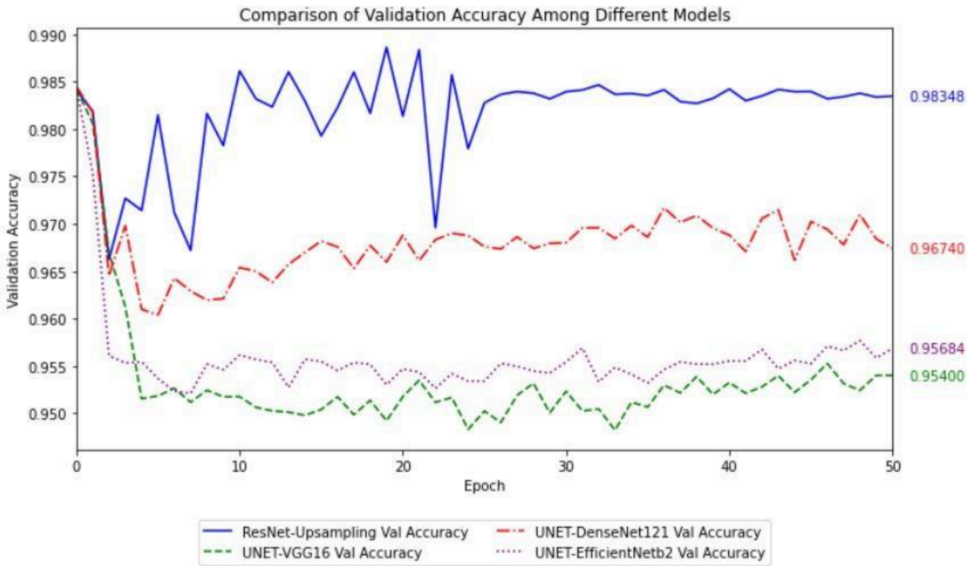

Figure 7. Comparative Graph Validation Accuracy Model

Figure 7 illustrates the validation accuracy comparison, where the highest validation accuracy achieved 98.35% by the ResNet model optimized with Nearest Neighbor Upsampling. While the lowest accuracy achieved 95.39% by UNET-VGG16 model. This indicates that the ResNet model optimized with Nearest Neighbor Upsampling better generalizes training patterns when validated against the validation data, leading to more accurate predictions on new data.

CONCLUSION

In this research, we explored two segmentation models for the brain tumor region classification task: ResNet-50 and ResNet optimized with Nearest Neighbor Upsampling. The Brain Tumor Segmentation (BraTS) 2020 dataset was used, obtained from the Center for Biomedical Image Computing & Analytic (CBICA) website. Before the data was processed for model training, data preprocessing was performed to prepare the data for the model's needs, improve the model's performance and generalization ability, and address specific challenges in the image segmentation task. The data augmentation methods used in this research included various transformations such as rotation, translation, flipping, inversion, zooming, and cropping. After the data was optimized in the data preprocessing stage, the data training stage was carried out using several segmentation methods. The first model used the Convolutional Neural Network (CNN) architecture concept with ResNet-50 as the main architecture. The second model used the decoder-encoder architecture concept, with ResNet used as the decoder and Nearest Neighbor Upsampling used as the encoder. Both models were then trained and evaluated to determine their reliability in the brain tumor region classification task. The models were evaluated based on the loss and accuracy values obtained when testing the test data. The segmentation model using ResNet optimized with Nearest Neighbor Upsampling achieved the highest accuracy of 98.44% and the lowest loss value of 0.0874. The brain MRI tumor image segmentation model with ResNet optimized with Nearest Neighbor Upsampling is better at preventing the loss of spatial detail in feature maps by increasing the image dimension and filling in empty gaps in the image columns/rows during the model training process.

REFERENCES

[1] S. Hyuna et al., “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: A Cancer Journal for Clinicians*, 2021.
[2] S. Iqbhal H. “Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions,” *SN Computer Science*, 2021.
[3] A. Laith, et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of big Data*, 2021.
[4] L. Zhihua, et al., "Deep learning based brain tumor segmentation: a survey," *Complex & intelligent systems*, 2023.
[5] S. Lamia H., et al., "An efficient brain tumor image segmentation based on deep residual networks (ResNets).," *Journal of King Saud University-Engineering Sciences*, 2021.
[6] K. Anita, et al., "A deep Residual U-Net convolutional neural network for automated lung segmentation in computed tomography images," *Biocybernetics and Biomedical Engineering*, 2020.
[7] S. Jiawei, et al., "DRRNet: dense residual refine networks for automatic brain tumor segmentation," *Journal of medical systems*, 2019.
[8] Z. Sihang, et al., "High-resolution encoder–decoder networks for low-contrast medical image segmentation," *IEEE Transactions on Image Processing*, 2019.
[9] B. Jorge and A. Histace, “Computer-Aided Analysis of Gastrointestinal Videos,” *Springer*, 2021.
[10] W. Shengsheng, X. Hou, and X. Zhao. "Automatic building extraction from high-resolution aerial imagery via fully convolutional encoder-decoder network with non-local block," *IEEE Access*, 2020.
[11] B. Spyridon, et al., "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge.," a*rXiv preprint arXiv:1811.02629*, 2018.
[12] H. Sung Yang, et al., "Extensions of the external validation for checking learned model interpretability and generalizability," *Patterns*, 2020.
[13] H. Gousia, and S. Qureshi, "Optimization and acceleration of convolutional neural networks: A survey," *Journal of King Saud University-Computer and Information Sciences*, 2022.
[14] H. Intisar Rizwan and J. Neubert. "Deep learning approaches to biomedical image segmentation," *Informatics in Medicine Unlocked*. 2020.
[15] B. Subhashis, and S. Mitra, "Novel volumetric sub-region segmentation in brain tumors," *Frontiers in Computational Neuroscience*, 2020.
[16] A. Sana and A. Jitendra, "Automated segmentation of brain tumour images using deep learning-based model VGG19 and ResNet 101," *Multimedia Tools and Applications*, 2024.
[17] D. Dinthisrang, et al., "Brain tumor segmentation from MRI images using hybrid convolutional neural networks," *Procedia Computer Science*. 2020.
[18] P. Anindya Apriliyanti, et al., "On comparing optimizer of UNet-VGG16 architecture for brain tumor image segmentation," *Academic Press*, 2022.
[19] L. Fangyu, and L. Wang, "UNet-based model for crack detection integrating visual explanations," *Construction and Building Materials*, 2022.
[20] C. Iulen, et al., "Automated segmentation of the ciliary muscle in OCT images using fully convolutional networks," *Biomedical optics express*, 2022.