SINTECHCOM: Science, Technology and Communication Journal
Vol. 6, No. 3, June 2026, pp. 305-316, DOI: 10.59190/stc.v6i3.387

Implementation of U-Net as EfficientNet encoder for brain tumour type classification

Rahma Aulia*, Junadhi, Lusiana Efrizoni, Rini Yanti
Department of Informatics Engineering, Universitas Sains dan Teknologi Indonesia, Pekanbaru 28299, Indonesia

ABSTRACT

Brain tumor is one of the most dangerous diseases that requires fast and accurate diagnosis to support patient diagnose. The application of deep learning on magnetic resonance imaging (MRI) images has been widely used to assist automatic brain tumor classification. This study aims to implement a hybrid U-Net encoder-EfficientNet architecture for brain tumor classification using MRI images. In this study, the U-Net encoder was utilized to extract spatial features and generate an attention mask to highlight important regions before the classification process was performed by EfficientNet-B0. The dataset used was BRISC 2025, consisting of 6,000 MRI images divided into four classes: glioma, meningioma, pituitary, and no tumor. The experiments were conducted using three data splitting scenarios, namely 60:20:20, 70:15:15, and 80:10:10. The results showed that the proposed model achieved good classification performance across all testing scenarios. In the 60:20:20 scenario, the model achieved an accuracy of 82%, precision of 0.83, recall of 0.82, and F1-score of 0.81. In the 70:15:15 scenario, the model achieved an accuracy of 84%, precision of 0.85, recall of 0.84, and F1-score of 0.83. Meanwhile, the 80:10:10 scenario produced the best performance with an accuracy of 85%, precision of 0.86, recall of 0.85, and F1-score of 0.84. These results indicate that the use of the U-Net encoder was able to help the model focus on tumor regions, thereby improving the effectiveness of the classification process.

ARTICLE INFO

Article history:

Received Jun 12, 2026
Revised Jun 15, 2026
Accepted Jun 16, 2026

Keywords:

Brain Tumor
EfficientNet
Encoder
Image Classification
U-Net

* **Corresponding Author**
E-mail address: raaulia98@gmail.com

1. INTRODUCTION

Brain tumors are abnormal cell growths occurring within or around the brain that are often aggressive and potentially life-threatening [1]. In Indonesia, brain tumor cases continue to be reported across various age groups each year [2]. Therefore, early detection and accurate diagnosis play a crucial role in determining appropriate treatment strategies and patient prognosis [3]. In clinical practice, Magnetic Resonance Imaging (MRI) is widely utilized for brain tumor diagnosis due to its ability to produce high-resolution images of soft tissues [4].

However, MRI interpretation is still predominantly performed manually by radiologists, making the diagnostic process time-consuming and potentially subject to inter-observer variability [5]. Such limitations may affect the consistency and reliability of diagnostic outcomes, particularly in cases involving complex tumor characteristics. Consequently, the advancement of artificial intelligence, especially deep learning, has attracted considerable attention for assisting automated medical image analysis [6].

One of the most widely adopted deep learning approaches is the Convolutional Neural Network (CNN), which has demonstrated remarkable effectiveness in extracting complex features from medical images [7]. In brain tumor classification tasks, CNN models are expected not only to distinguish tumor types, such as glioma, meningioma, and pituitary tumors, but also to capture the spatial characteristics of brain structures [6]. Spatial representation learning is particularly important

because tumor characteristics are often defined by the shape, texture, and location of abnormal regions within MRI scans.

In the field of medical image analysis, the U-Net architecture has been recognized for its effectiveness in image segmentation and pixel-level spatial feature extraction [8]. The encoder component of U-Net is capable of capturing rich visual representations from MRI images [9], generating features that effectively characterize tumor-related patterns [4]. These capabilities make U-Net a promising candidate for spatial feature extraction in medical image classification tasks.

For the classification stage, EfficientNet is one of the state-of-the-art CNN architectures that achieves high performance while maintaining computational efficiency through its compound scaling approach [10]. Previous studies have demonstrated that EfficientNet can achieve high brain tumor classification accuracy following appropriate preprocessing and hyperparameter optimization procedures [11, 12]. Nevertheless, the classification performance remains highly dependent on the quality of feature representations provided to the model.

Motivated by these challenges, this study implements U-Net as the encoder component of an EfficientNet-based architecture for brain tumor classification using MRI images. The U-Net encoder is employed to extract spatial features and generate an attention mask that highlights relevant tumor regions prior to classification by EfficientNet. This approach is expected to integrate U-Net's capability in understanding the structural characteristics of medical images with EfficientNet's classification efficiency, thereby improving the effectiveness and accuracy of brain tumor classification.

2. RESEARCH METHODS

This section describes the research methodology employed to develop a brain tumor classification model using a hybrid U-Net and EfficientNet-B0 approach. The research stages include data collection, data splitting, preprocessing, spatial feature extraction using the U-Net encoder, classification using EfficientNet-B0, and model performance evaluation. The overall research workflow is illustrated in Figure 1.

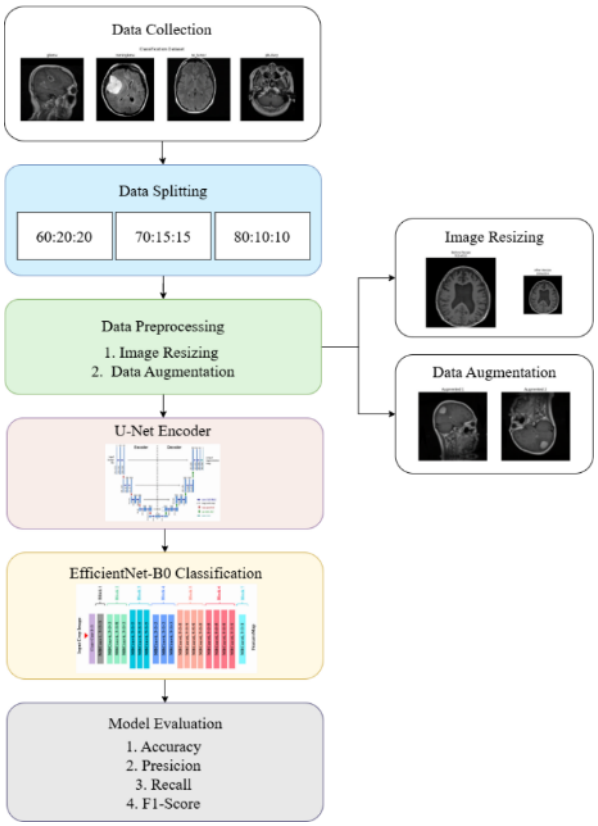


Figure 1. Research workflow.

2.1. Data Collection

This study utilizes a publicly available dataset obtained from the Kaggle platform. The dataset, entitled BRISC 2025, contains brain MRI images for both tumor segmentation and classification tasks [13]. The dataset is provided with annotations, enabling its use for training and evaluating brain tumor classification models. A total of 6,000 MRI images are included in the dataset, distributed across four classes: glioma, meningioma, pituitary tumor, and no tumor.

2.2. Data Splitting

The dataset was divided into training, validation, and testing sets to ensure objective model training and evaluation. Three data-splitting scenarios were employed in this study, namely 60:20:20, 70:15:15, and 80:10:10. These variations were designed to analyze the effect of training data quantity on the classification performance of the proposed model.

2.3. Data Preprocessing

The preprocessing stage was conducted to prepare MRI images according to the model requirements. The preprocessing procedures included image resizing and data augmentation.

2.3.1. Image Resizing

Image resizing was performed to standardize image dimensions before being processed by the deep learning model [14]. All MRI images were resized to 224 × 224 pixels to match the input requirements of EfficientNet-B0.

2.3.2. Data Augmentation

Data augmentation was applied to increase the diversity of the training dataset and reduce the risk of overfitting [15]. Augmentation was performed dynamically using several image transformations, including horizontal flipping, vertical flipping, rotation, and zooming. These transformations were applied exclusively to the training data, while the validation and testing datasets remained unchanged to ensure an objective evaluation process.

2.4. U-Net Encoder

U-Net is a CNN architecture specifically designed for medical image segmentation and is well known for its ability to extract spatial features through its encoder-decoder structure and skip connections [16]. The primary characteristic of U-Net lies in the implementation of skip connections that link feature maps between the encoder and decoder, enabling the recovery of high-resolution spatial information that may be lost during the downsampling process [17]. In this study, the encoder component of U-Net was utilized to extract spatial representations from MRI images before being forwarded to the classification model. The U-Net architecture is illustrated in Figure 2.

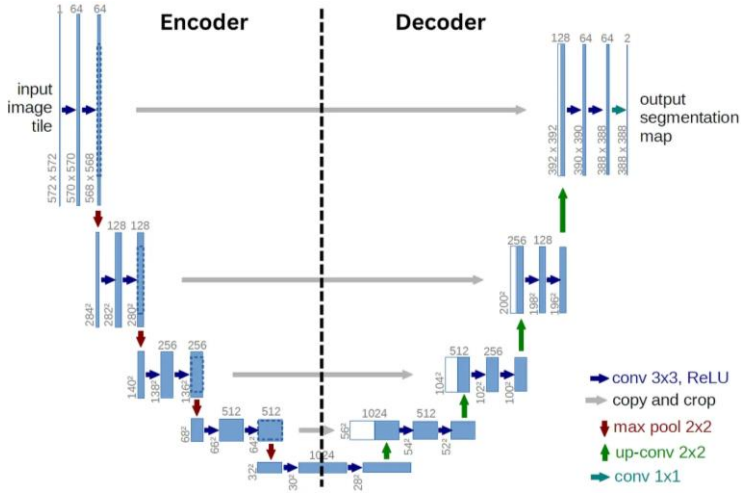


Figure 2. U-Net architecture [3].

The U-Net architecture consists of three main components: an encoder, a bottleneck, and a decoder. The encoder extracts features using convolutional blocks and max-pooling operations to capture contextual information from the images. Subsequently, the decoder performs upsampling and integrates encoder features through skip connections to preserve spatial information. The output of this process is an attention mask, which is multiplied with the original input image to generate an attention-guided image that serves as the input to the classification model.

2.5. EfficientNet-B0 Classification Model Training

EfficientNet-B0 was selected as the classification model due to its ability to provide a balance between classification accuracy and computational efficiency through the compound scaling approach [10]. Pooling layers progressively reduce the spatial dimensions of feature representations, thereby decreasing the number of trainable parameters and helping to mitigate overfitting [18]. EfficientNet-B0 employs Mobile Inverted Bottleneck Convolution (MBConv) blocks and squeeze-and-excitation modules to enhance feature extraction efficiency. The architecture of EfficientNet-B0 is illustrated in Figure 3.

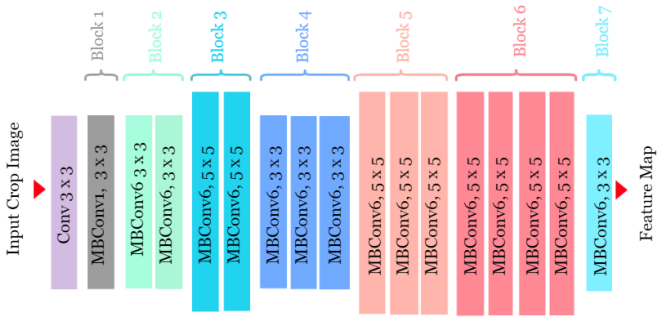


Figure 3. Illustration of the EfficientNet-B0 architecture [10].

In this study, the model receives attention-guided images generated by the U-Net encoder as input for classifying MRI scans into four brain tumor categories. EfficientNet-B0 was initialized using ImageNet pre-trained weights, while the feature extraction layers were retained in a frozen state. The extracted features were subsequently processed through Global Average Pooling, Batch Normalization, Dense layers, Dropout, and a Softmax layer to produce classification probabilities.

2.6. Model Evaluation

Model evaluation was performed using a confusion matrix to compare predicted labels with ground-truth labels [5]. Model performance was assessed using accuracy, precision, recall, and F1-score metrics [19]. Accuracy measures overall prediction correctness, precision evaluates the correctness of positive predictions for each class, recall assesses the model's ability to identify instances belonging to each class, and F1-score provides a balanced evaluation between precision and recall [5].

The confusion matrix consists of four fundamental components used in the calculation of classification evaluation metrics [20], as presented in Table 1.

Table 1.Confusion matrix.

Classification	Classified as positive	Classified as negative
Actual positive	TP	FN
Actual negative	FP	TN

Description: TP (True Positive) : Positive samples correctly classified into the positive class.
TN (True Negative) : Negative samples correctly classified into the negative class.
FP (False Positive) : Negative samples incorrectly classified as positive.
FN (False Negative) : Positive samples incorrectly classified as negative.

The performance metrics are computed using the following equations:

- Accuration

$$Accuration = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (1)$$

- Precision

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

- Recall

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

- F1-Score

$$F1 - Score = \frac{2\times(Precision\times Recall)}{(Precision+Recall)} \times 100\% \quad (4)$$

3. RESULTS AND DISCUSSIONS

This chapter presents the implementation results of the U-Net Encoder–EfficientNet model for brain tumor classification using MRI images. The discussion covers the data collection process, data splitting, preprocessing, U-Net encoder implementation, model training, and performance evaluation under various data-splitting scenarios.

3.1. Data Collection

The dataset used in this study was the BRISC 2025 dataset, which consists of 6,000 brain MRI images for tumor classification. The dataset is divided into four classes, there are glioma, meningioma, pituitary tumor, and no tumor. The dataset comprises 1,401 glioma images, 1,635 meningioma images, 1,207 no tumor images, and 1,757 pituitary tumor images. The distribution of samples across the four classes is relatively balanced, providing sufficient representation for each category and enabling effective model training and evaluation. The brain tumor classification dataset used in this study is presented in Figure 4.

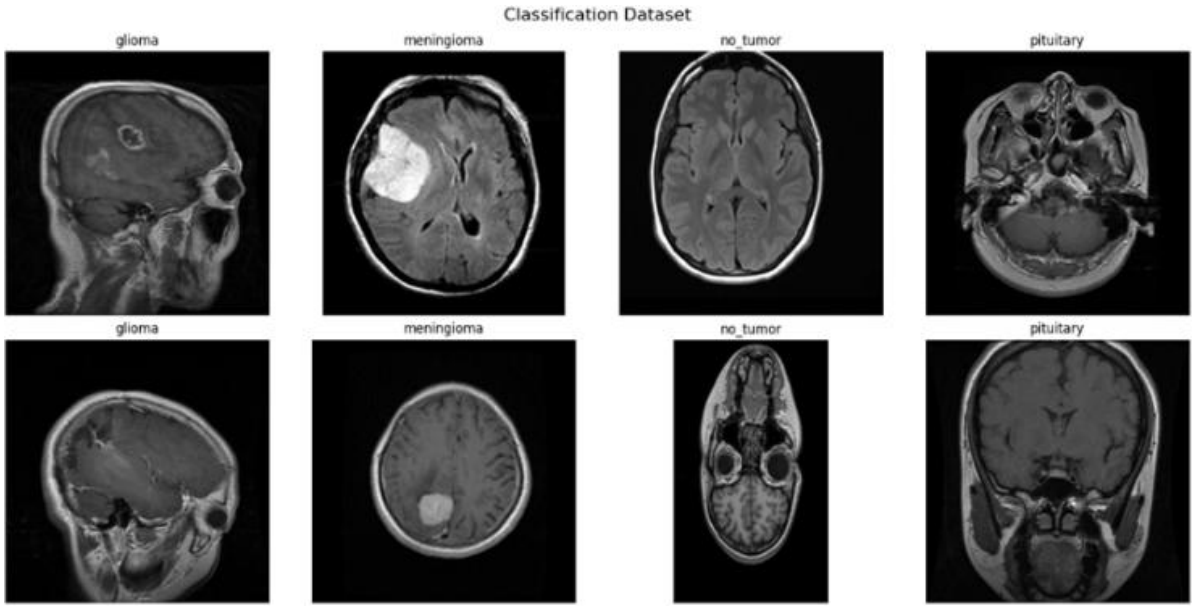


Figure 4. The dataset.

3.2. Data Splitting

Following the data collection stage, the dataset was divided into training, validation, and testing sets using three data-splitting scenarios: 60:20:20, 70:15:15, and 80:10:10. A stratified splitting strategy was employed to preserve the class distribution across all subsets. These configurations were designed to investigate the effect of training data proportion on the performance of the proposed classification model.

3.3. Data Preprocessing

The preprocessing stage consisted of image resizing and data augmentation.

3.3.1. Image Resizing

During preprocessing, all MRI images were resized from their original resolution of 512 × 512 pixels to 224 × 224 pixels. This step was performed to standardize the input dimensions and ensure compatibility with the EfficientNet-B0 architecture used in this study. In addition, image resizing reduces computational complexity and accelerates the training process. Examples of images before and after resizing are presented in Figure 5.

Figure 5. MRI images before and after resizing.

3.3.2. Data Augmentation

In this study, data augmentation was applied exclusively to the training dataset, while the validation and testing datasets remained unchanged. The augmentation process employed several image transformation techniques, including horizontal and vertical flipping, random rotation with a factor of 0.15, and random zoom with a factor of 0.15.

These transformations generated additional image variations with different orientations, positions, and scales compared to the original images. Horizontal and vertical flipping were used to reverse image orientation, random rotation was applied to rotate images randomly, and random zoom was used to enlarge or shrink specific image regions. Examples of the augmented images are shown in Figure 6.

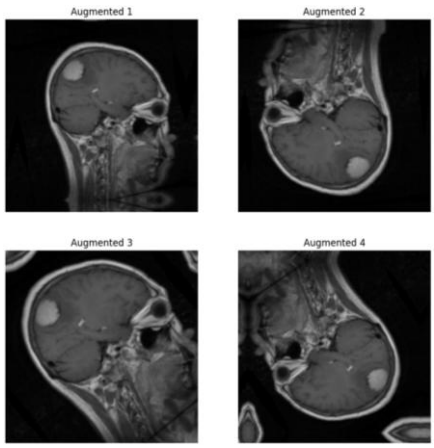


Figure 6. Data augmentation results.

3.4. U-Net Encoder

In this study, U-Net was employed as a spatial encoder to highlight relevant regions in MRI images before classification by EfficientNet-B0. The U-Net architecture consists of an encoder, bottleneck, decoder, and skip connections.

The encoder extracts image features through convolutional blocks and max-pooling operations, while the decoder performs upsampling to reconstruct spatial representations. Skip connections are utilized to preserve spatial information throughout the feature extraction process. The generated attention masks and attention-guided images are illustrated in Figure 7.

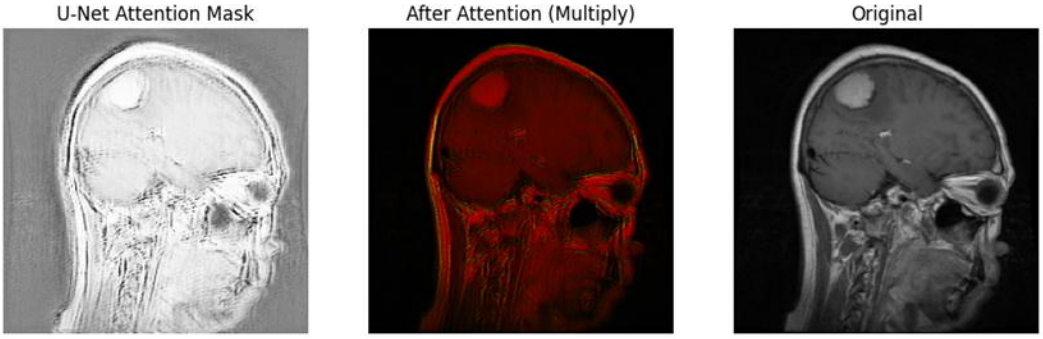


Figure 7. U-Net encoder output.

The decoder generates an attention mask, which is subsequently multiplied with the original image using a multiplication operation to produce an attention-guided image. The resulting attention maps indicate that regions containing tumor-related characteristics become more prominent than the background, enabling the classification model to focus more effectively on tumor-specific features.

3.5. EfficientNet-B0 Classification Model Training

The proposed U-Net Encoder–EfficientNet model was evaluated using three data-splitting scenarios: 60:20:20, 70:15:15, and 80:10:10. All models were trained for 10 epochs under identical training settings.

For the 60:20:20 splitting scenario, the model demonstrated a satisfactory learning process throughout training. Training accuracy gradually increased from 67.86% in the first epoch to 86.58% in the final epoch. Validation accuracy also improved and reached its highest value of 84.08% at epoch 9. Although minor fluctuations were observed in the validation loss, the overall loss trend decreased during training. The training results are presented in Figure 8.

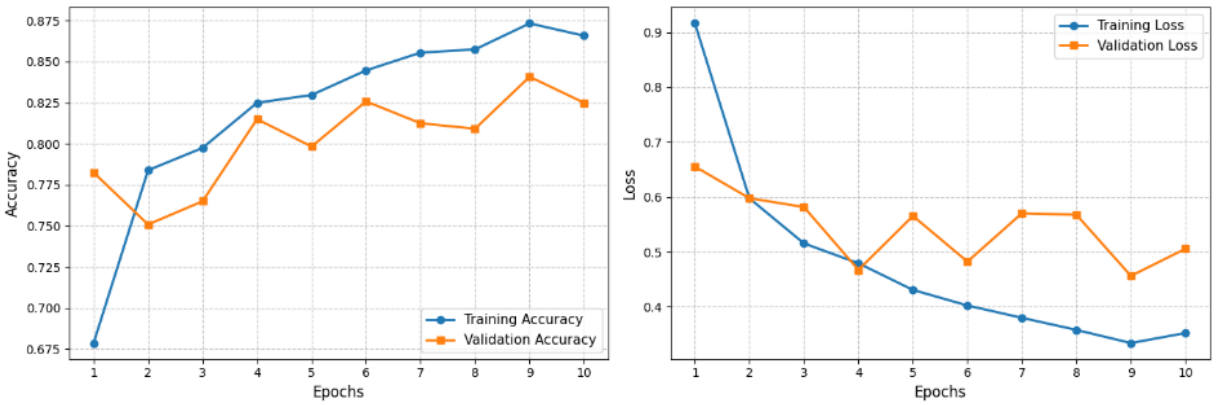
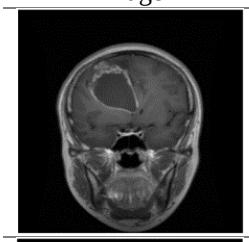
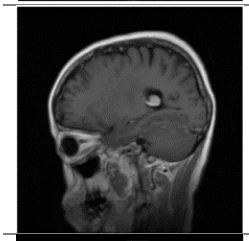
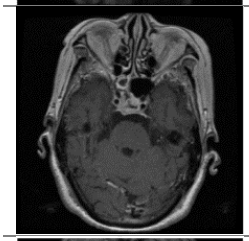
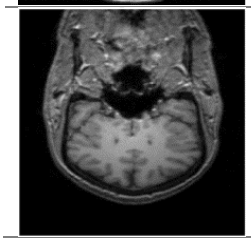


Figure 8. Training results for the 60:20:20 splitting scenario.

To further evaluate model performance, several testing samples were used to assess the model's ability to predict brain tumor MRI images. The prediction results are presented in Table 2.

Table 2. Prediction results.

Image	Predicted label	Actual label
	Glioma	Glioma
	Meningioma	Meningioma
	Pituitary	Pituitary
	No Tumor	No Tumor

For the 70:15:15 splitting scenario, the model exhibited a more stable training process than the previous configuration. Training accuracy increased from 69.45% in the first epoch to 87.55% in the final epoch. Validation accuracy gradually improved and achieved its highest value of 83.56% at epoch 6. Furthermore, the loss values consistently decreased throughout training, as illustrated in Figure 9.

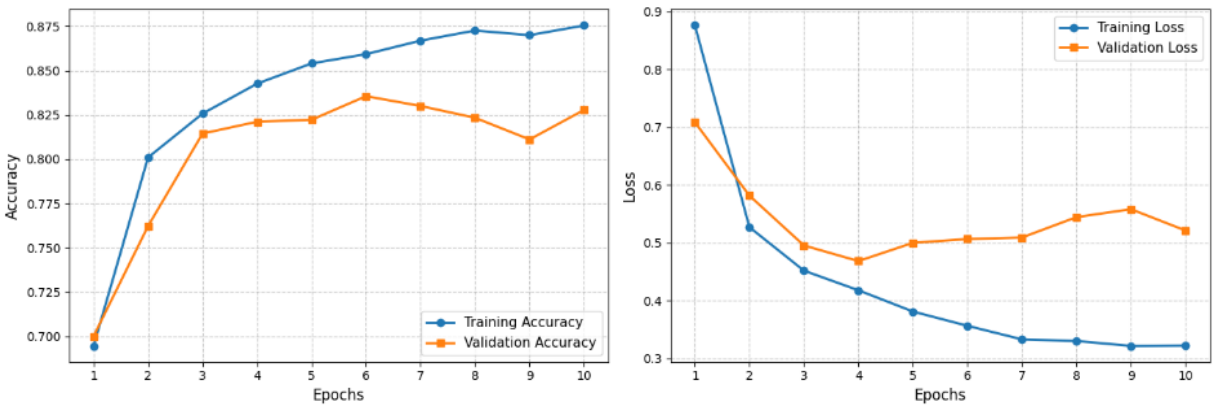
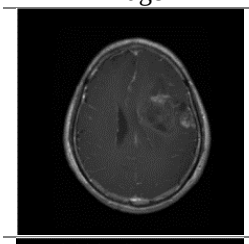
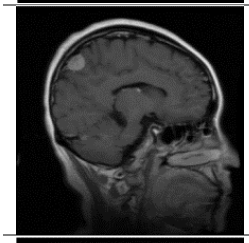
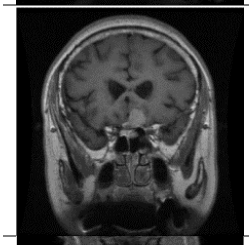
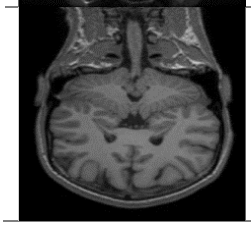


Figure 9. Training results for the 70:15:15 splitting scenario.

Testing was also performed using several MRI samples to evaluate prediction performance under the 70:15:15 scenario. The results are presented in Table 3.

Table 3. Prediction results.

Image	Predicted label	Actual label
	Glioma	Glioma
	Pituitary	Meningioma
	Pituitary	Pituitary
	No Tumor	No Tumor

The 80:10:10 splitting scenario produced the best training performance among all configurations. Training accuracy improved from 70.83% in the first epoch to 89.50% in the final epoch. Validation accuracy reached its highest value of 88.50% at epoch 9, while the lowest validation loss recorded was 0.2670. The corresponding training curves are shown in Figure 10.

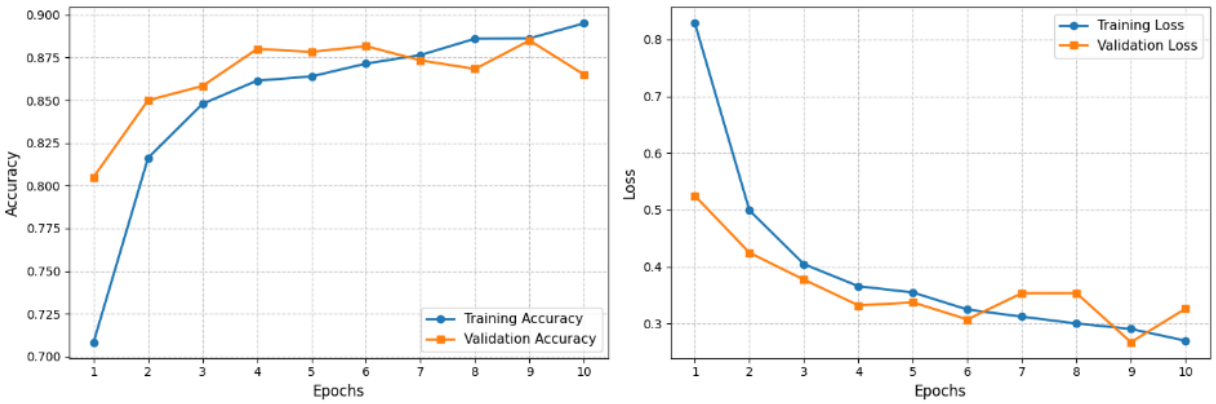
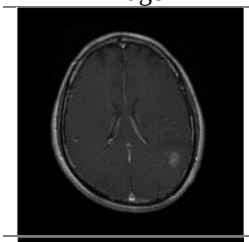
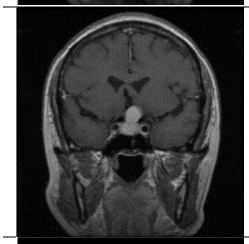


Figure 10. Training results for the 80:10:10 splitting scenario.

The training curves indicate that model learning improved as the proportion of training data increased. The decrease in loss and increase in accuracy suggest that the model learned MRI feature representations more effectively than in the other splitting scenarios. Further evaluation using testing

samples was conducted to assess the model's classification capability under the 80:10:10 scenario. The results are presented in Table 4.

Table 4. Prediction results.

Image	Predicted label	Actual label
	Glioma	Glioma
	Pituitary	Meningioma
	Pituitary	Pituitary
	No Tumor	No Tumor

The overall testing performance for each splitting scenario is summarized in Table 5.

Table 5. Testing results.

Data splitting	Accuracy (%)	Loss
60:20:20	82.25	0.5056
70:15:15	83.00	0.5172
80:10:10	85.00	0.3915

The classification performance improved as the amount of training data increased. The 80:10:10 scenario achieved the best performance, with an accuracy of 85.00% and a loss value of 0.3915. In comparison, the 60:20:20 and 70:15:15 scenarios achieved accuracies of 82.25% and 83.00%, respectively. These findings indicate that a larger training dataset enables the model to learn MRI image patterns more effectively, resulting in improved classification performance on unseen testing data.

4. CONCLUSION

This study successfully implemented a hybrid U-Net Encoder–EfficientNet architecture for brain tumor classification using MRI images. The U-Net encoder was effectively utilized to extract

spatial features and generate an attention mask that highlights relevant regions within the image prior to the classification process performed by EfficientNet-B0.

Based on the experimental results, the proposed model achieved satisfactory classification performance across all data-splitting scenarios. For the 60:20:20 split, the model achieved an accuracy of 82%, a precision of 0.83, a recall of 0.82, and an F1-score of 0.81. For the 70:15:15 split, the model performance improved, achieving an accuracy of 84%, a precision of 0.85, a recall of 0.84, and an F1-score of 0.83. Meanwhile, the 80:10:10 split produced the best performance, with an accuracy of 85%, a precision of 0.86, a recall of 0.85, and an F1-score of 0.84.

The results indicate that increasing the amount of training data positively affects the model's ability to learn discriminative patterns from brain tumor MRI images. Furthermore, the incorporation of the U-Net encoder enables the model to focus on regions containing tumor-related characteristics, thereby improving the effectiveness of the classification process. Therefore, the proposed hybrid U-Net Encoder–EfficientNet approach demonstrates promising potential for supporting MRI-based brain tumor classification.

REFERENCES

[1] Nuresa, K. (2024). Analisis Perbandingan Klasifikasi Tumor Otak Menggunakan Deep Learning. *IJITECH: Indonesian Journal of Information Technology*, **2**(1), 16–21.

[2] Harahap, F. A. A., Nafisa, A. N., Purba, E. N. D. B., & Putri, N. A. (2023). Implementasi algoritma convolutional neural network arsitektur model MobileNetV2 dalam klasifikasi penyakit tumor otak glioma, pituitary dan meningioma. *Jurnal Teknologi Informasi, Komputer, Dan Aplikasinya (JTIKA)*, **5**(1), 53–61.

[3] Ardan, I. S. & Indraswari, R. (2024). Sistem Berbasis Deep Learning untuk Segmentasi dan Klasifikasi Tingkat Keganasan Tumor Otak Menggunakan Citra MRI 3D. *ILKOMNIKA*, **6**(2), 117–126.

[4] Suta, I. B. L. M., Sudarma, M., & Kumara, I. S. (2020). Segmentasi Tumor Otak Berdasarkan Citra Magnetic Resonance Imaging Dengan Menggunakan Metode U-NET. *Majalah Ilmiah Teknologi Elektro*, **19**(2), 151.

[5] Indriyani, F., & Rahadjeng, I. R. (2023). Klasifikasi Tumor Otak Berbasis Magnetic Resonance Imaging Menggunakan Algoritma Convolutional Neural Network. *Digital Transformation Technology*, **3**(2), 918–924.

[6] Hakim, M. N. M., Nugroho, A. B., & Minarno, A. E. (2023). Prediksi Tumor Otak Menggunakan Metode Convolutional Neural Network. *Inform. Mulawarman J. Ilm. Ilmu Komput*, **17**(1), 48.

[7] Febrianti, A. S., Sardjono, T. A., & Babgei, A. F. (2020). Klasifikasi tumor otak pada citra magnetic resonance image dengan menggunakan metode support vector machine. *Jurnal Teknik ITS*, **9**(1), A118–A123.

[8] Sonia, S. & Yohannes, Y. (2025). Penerapan Model U-Net untuk Segmentasi Gigi pada Citra Radiografi Panoramik Orang Dewasa. *Algoritme Jurnal Mahasiswa Teknik Informatika*, **5**(2), 188–198.

[9] Futrega, M., Milesi, A., Marcinkiewicz, M., & Ribalta, P. (2021). Optimized U-Net for brain tumor segmentation. *International MICCAI Brainlesion Workshop*, 15–29.

[10] Dianto, A. R., Anggraeny, F. T., & Maulana, H. (2025). Analisis efektifitas algoritma MobileNetV3-Large dan EfficientNet-B0 untuk klasifikasi citra penyakit daun jeruk. *Jurnal Informatika dan Teknik Elektro Terapan*, **13**(3).

[11] Tyas, D. L., Rumambi, F. R., Patanduk, A., & Mailangkay, R. C. J. (2025). Klasifikasi Jenis Tumor Otak Melalui Citra MRI dengan Menggunakan Convolutional Neural Network. *Informatik: Jurnal Ilmu Komputer*, **21**(1), 26–34.

[12] Ramadhani, R. A., Pangestu, B. W., & Purbaningtyas, R. (2022). Klasifikasi tumor otak menggunakan convolutional neural network dengan arsitektur EfficientNet-B3. *Jurnal Sistem Informasi, Teknologi Informatika dan Komputer*, 55–59.

[13] Fateh, A., Rezvani, Y., Moayedi, S., Rezvani, S., Fateh, F., Fateh, M., & Abolghasemi, V. (2026). BRISC: Annotated dataset for brain tumor segmentation and classification. *Scientific Data*, 1–18.

[14] Talebi, H. & Milanfar, P. (2021). Learning to resize images for computer vision tasks. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 497–506.
[15] Ayub, M. (2020). Augmentasi data pengenalan citra mobil menggunakan pendekatan random crop, rotate, dan mixup. Jurnal Teknik Informatika dan Sistem Informasi. *J. Tek. Inform. dan Sist. Inf.*, **6**, 311–323.
[16] Siddique, N., Sidike, P., Elkin, C., & Devabhaktuni, V. (2020). U-Net and its variants for medical image segmentation: theory and applications. *IEEE Access*, **9**.
[17] Salsabila, P. & Islamadina, R. (2026). Optimasi Segmentasi Kepala Janin Berbasis U-Net Melalui Preprocessing Citra USG. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, **10**(1), 12–25.
[18] Perdani, W. R., Magdalena, R., & Pratiwi, N. K. C. (2022). Deep Learning untuk Klasifikasi Glaukoma dengan menggunakan Arsitektur EfficientNet. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, **10**(2), 322.
[19] Aria, M. S. A., Slamet, C., & Firdaus, M. D. (2025). Klasifikasi Fake dan Real Menggunakan Vision Transformer dan EfficientNet-B0 pada Gambar Asli dan Generatif AI. *SMATIKA JURNAL*, **15**(01), 179–192.
[20] Fitriyana, K. H., Tyas, F. A., & Jamil, A. (2026). Peningkatan Performa Classification and Regression Tree Menggunakan Bagging pada Diagnosis Penyakit Jantung. *Jurnal Teknik Informatika dan Sistem Informasi*, **12**(1), 45–59.