International Journal of Robotics and Control Systems
Vol. 5, No. 2, 2025, pp. 917-936
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

Impact of Hyperparameter Tuning on ResNet-UNet Models for Enhanced Brain Tumor Segmentation in MRI Scans

Yuri Pamungkas [a,1,*], Evi Triandini [b,2], Wawan Yunanto [c,3], Yamin Thwe [d,4]

[a] Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
[b] Department of Information System, Institut Teknologi dan Bisnis STIKOM Bali, Denpasar, 80234, Indonesia
[c] Department of Computer Science and Information Engineering, NTUST, Taipei, 106, Taiwan
[d] Department of Big Data Management & Analytics, RMUTT, Bangkok, 12120, Thailand
[1] yuri@its.ac.id; [2] evi@stikom-bali.ac.id; [3] d10615811@mail.ntust.edu.tw; [4] yamin_t@rmutt.ac.th
* Corresponding Author

ARTICLE INFO

Article history
Received February 05, 2025
Revised March 06, 2025
Accepted March 14, 2025

Keywords
Brain Tumor;
Data Segmentation;
Medical Imaging;
ResNet-UNet;
Deep Learning

ABSTRACT

Brain tumor segmentation in MRI scans is a crucial task in medical imaging, enabling early diagnosis and treatment planning. However, accurately segmenting tumors remains a challenge due to variations in tumor shape, size, and intensity. This study proposes a ResNet-UNet-based segmentation model using LGG dataset (from 110 patients), optimized through hyperparameter tuning to enhance segmentation performance and computational efficiency. The proposed model integrates different ResNet architectures (ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152) with UNet, evaluating their performance under various learning rates (0.01, 0.001, 0.0001), optimizer types (Adam, SGD, RMSProp), and activation functions (Sigmoid). The methodology involves training and evaluating each model using Loss Function, Mean Intersection over Union (mIoU), Dice Similarity Coefficient (DSC), and Iterations per Second as performance metrics. Experiments were conducted on MRI brain tumor datasets to assess the impact of hyperparameter tuning on model performance. Results show that lower learning rates (0.0001 and 0.001) improve segmentation accuracy, while Adam and RMSProp outperform SGD in minimizing segmentation errors. Deeper models (ResNet50, ResNet101, and ResNet152) achieve the highest mIoU (up to 0.902) and DSC (up to 0.928), but at the cost of slower inference speeds. ResNet50 and ResNet34 with RMSProp or Adam provide an optimal trade-off between accuracy and computational efficiency. In conclusion, hyperparameter tuning significantly impacts MRI segmentation performance, and selecting an appropriate learning rate, optimizer, and model depth is crucial for achieving high segmentation accuracy with minimal computational cost.

1. Introduction

Magnetic Resonance Imaging (MRI) data segmentation is an important step in the diagnosis and management of patients with brain tumors [1]. Brain tumors, which can be benign or malignant, are serious diseases with high morbidity and mortality rates. Early detection and a thorough understanding of the location, size, and type of tumor are essential for determining the appropriate treatment strategy,

such as surgery, radiation therapy, or chemotherapy [2]. In this context, MRI data segmentation helps separate tumor tissue from normal brain tissue, edema, or cerebrospinal fluid, thereby producing an accurate visual map for planning medical interventions [3]. Conventionally, the segmentation process is performed manually by radiologists, which is time-consuming and prone to interobserver variability. In addition, the heterogeneous nature of brain tumors, including variations in size, shape, and signal intensity, makes the segmentation process a major challenge [4]. With the increasing complexity of brain tumor cases, manual methods are no longer sufficient to meet urgent clinical needs. Therefore, an automated or semi-automated technology-based approach is needed to improve the speed and accuracy of segmentation. The use of machine learning and deep learning technologies in MRI data segmentation offers a potential solution. Models such as Convolutional Neural Networks (CNNs), including variants such as U-Net or 3D CNN, have been shown to provide accurate and efficient segmentation results [5]. With this technology, the model can learn complex patterns in MRI data to detect tumors, even at an early stage of development. This is very important because early detection of brain tumors is closely related to the success rate of therapy and patient prognosis [6]. In addition to being important for diagnosis, MRI data segmentation also plays a key role in medical research. Segmented data allows researchers to quantitatively analyze tumor characteristics, such as volume, growth, or response to therapy [7]. This information not only helps clinical decision-making but also contributes to the development of new therapies and personalization of care. With accurate segmentation, MRI data can be used to create predictive models that help predict disease progression and treatment effectiveness [8].

Several recent studies have proposed various segmentation methods based on Convolutional Neural Networks (CNN) and optimization algorithms to enhance tumor detection accuracy. Zheng et al. (2022) proposed an enhanced U-Net model for brain tumor segmentation to overcome limitations in conventional segmentation approaches, such as poor edge detail extraction and inadequate feature reuse [9]. The study introduced a hybrid dilated convolution (HDC) module within a serial encoding-decoding structure, which enhances segmentation accuracy. Additionally, a new loss function was designed to improve segmentation of difficult-to-classify tumor regions. Experimental results showed that the modified U-Net outperformed existing models in Dice similarity coefficient, precision, and Hausdorff distance. This advancement is crucial for more accurate brain tumor detection and could improve automated diagnosis efficiency. Ullah et al. (2023) presented a hybrid Convolutional Neural Network (CNN)-based segmentation approach that integrates handcrafted features with CNNs to improve segmentation accuracy [10]. The proposed framework extracts intensity, texture, and shape-based features, which are then fused with CNN-extracted deep features. Evaluations on the BraTS dataset demonstrated superior performance in segmentation accuracy, Dice score, and specificity compared to standalone CNN-based methods. This study highlights the importance of integrating domain-specific handcrafted features with deep learning techniques to enhance segmentation robustness.

Aggarwal et al. (2023) proposed an improved ResNet-based segmentation model that enhances traditional deep learning models by addressing gradient diffusion issues [11]. The study optimized information flow through residual blocks and projection shortcuts, resulting in improved precision and computational efficiency. Compared to traditional CNNs and Fully Convolutional Networks (FCNs), the enhanced ResNet model achieved over 10% improvement in accuracy, recall, and F1-score on the BraTS 2020 dataset. The research demonstrated that residual learning could significantly improve segmentation efficiency while reducing computational costs. Another study conducted by Walsh et al. (2022) proposed an implementation of U-Net that enables real-time segmentation with superior performance compared to other conventional segmentation algorithms [12]. This model was evaluated using the BITE dataset, with results showing an average Intersection-over-Union (IoU) score of 89%, which is higher than other benchmark methods. The main advantage of this approach is the use of images from three perspective planes (coronal, sagittal, transversal) instead of 3D volumetric images, thereby improving computational efficiency in brain tumor analysis. In the study by Ranjbarzadeh et al. (2023), AI-based brain tumor segmentation was examined using supervised learning, unsupervised learning, and deep learning [13]. This study highlighted that deep learning-

based methods, particularly CNNs, have significantly improved segmentation efficiency compared to conventional methods such as thresholding and k-means clustering. The research also emphasized the importance of using multi-modal MRI (T1, T2, T1ce, and FLAIR) to enhance tumor detection accuracy. One of the main challenges in brain tumor segmentation is overfitting due to the limited availability of annotated data for training deep learning models.

Therefore, this study aims to develop a brain tumor segmentation model based on ResNet-UNet that has been optimized through hyperparameter tuning to improve segmentation performance and efficiency. ResNet-UNet is a hybrid architecture that combines ResNet (Residual Network) as an encoder for deeper feature extraction and UNet as a decoder for segmentation image reconstruction. This combination provides advantages in handling varying tumor structure complexity and improves the model's generalization ability to various types of MRI data. However, model performance is greatly influenced by the selection of hyperparameters, such as learning rate, batch size, number of layers, and activation function, which directly affect the convergence process and segmentation accuracy. Experiments were conducted using the LGG (Low-Grade Glioma) Segmentation Dataset, which is a benchmark for evaluating brain tumor segmentation models. Model performance evaluation was carried out using metrics such as Loss, Dice Similarity Coefficient (DSC), and Mean Intersection over Union (mIoU), to ensure more accurate and consistent segmentation compared to conventional methods. The hypothesis is that optimal hyperparameter tuning can improve model generalization and reduce overfitting, thereby improving brain tumor segmentation performance on LGG datasets. In addition, with optimal hyperparameter tuning, the model is expected to work efficiently without losing precision in detecting complex tumor boundaries. The implementation of this model is expected to assist medical personnel in making faster and more accurate clinical decisions, as well as opening up opportunities for the development of artificial intelligence-based diagnostic support systems in the fields of radiology and oncology.

2. Material and Methods

2.1. Dataset

The LGG (Low-Grade Glioma) Segmentation Dataset is a collection of brain MRI images accompanied by manual FLAIR abnormality segmentation masks [14]. This dataset is designed to support research in brain tumor segmentation, particularly for detecting and classifying low-grade gliomas (LGG). FLAIR-based segmentation is crucial as this modality effectively highlights abnormal regions associated with tumors, including edema and structural changes in the brain that may not be clearly visible in other MRI modalities. The MRI images in this dataset were obtained from The Cancer Imaging Archive (TCIA), a widely used repository of medical imaging data for cancer research [15]. TCIA provides high-quality, curated datasets for various types of cancer, including brain tumors. The data in the LGG Segmentation Dataset has undergone preprocessing to ensure consistency in size, resolution, and intensity normalization, making it ready for use in various AI-based segmentation studies. This dataset consists of MRI scans from 110 patients, all of whom are part of The Cancer Genome Atlas (TCGA) lower-grade glioma collection. Each patient in this dataset has at least one Fluid-Attenuated Inversion Recovery (FLAIR) imaging sequence and available genomic cluster data. The inclusion of genomic information in this dataset enables research that correlates MRI imaging patterns with molecular tumor characteristics, improving our understanding of the relationship between tumor biology and imaging features [16].

2.2. Data Pre-Processing

Pre-processing is a crucial step in preparing the LGG Segmentation Dataset for deep learning-based brain tumor segmentation [17]. This process involves retrieving data, cleaning, transforming, and loading it efficiently to ensure consistency in model training. The first step in pre-processing is to gather all necessary files, including MRI scans, segmentation masks, and metadata containing patient-specific information such as clinical and genomic data. MRI scans are typically stored in NIfTI (.nii) format, while segmentation masks are provided as manually annotated ground truth labels. The dataset

includes a CSV file containing essential patient metadata, such as Patient ID, genomic cluster data (RNASeqCluster, MethylationCluster, miRNACluster, etc.), histopathological and clinical information (tumor tissue site, laterality, tumor location), demographic details (gender, age, race, and ethnicity), and survival data (death01). These features allow for a more comprehensive analysis by linking tumor characteristics in MRI scans with genetic and histological factors [18]. The CSV file is read into a pandas DataFrame, providing structured data processing for analysis. Handling missing values is a critical part of pre-processing to avoid inconsistencies in the dataset. For categorical variables, missing values are imputed using the most frequent category (mode), while for numerical values, missing data is filled using the mean or median, depending on the distribution. If a patient's metadata contains excessive missing values, the row may be removed to maintain dataset integrity. This step ensures that the dataset remains complete and unbiased for further analysis.

2.3. ResNet-UNet Architectures

ResNet-UNet is a deep learning architecture that combines the advantages of two well-known models, namely Residual Network (ResNet) and U-Net, to improve the performance of medical image segmentation, including in the detection and segmentation of brain tumors in MRI [19]. ResNet is known for its ability to prevent the vanishing gradient problem and enable deeper feature modeling, while U-Net is very effective in segmentation tasks with an encoder-decoder structure that allows for precise image detail reconstruction [20]. Fig. 1 is a diagram of the ResNet-UNet architecture.

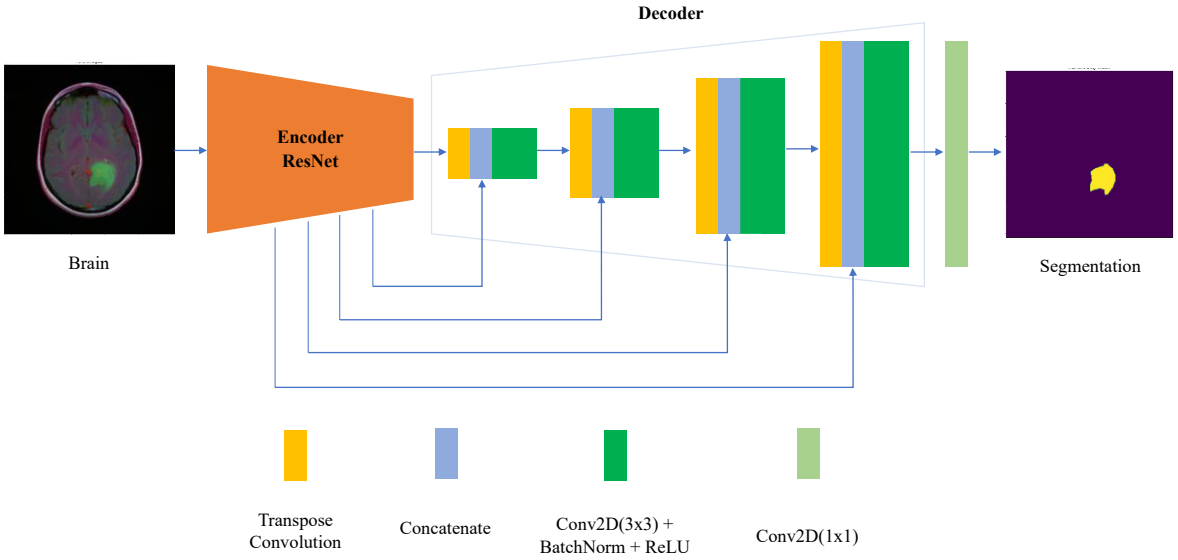

Fig. 1. ResNet-UNet architecture

The ResNet-UNet architecture utilizes ResNet as an encoder, which is responsible for extracting features from input images using residual blocks. Residual blocks in ResNet help retain important information by using shortcut connections, so that the model can learn more complex features without experiencing performance degradation due to too deep a network [21]. The use of ResNet as an encoder improves the feature extraction capability compared to the conventional encoder in U-Net. After the features are extracted by the ResNet encoder, the results are sent to the U-Net decoder section, which is responsible for reconstructing the high-resolution segmentation image [22]. The U-Net decoder consists of upsampling layers and skip connections that allow the recovery of spatial details from the extracted features. Skip connections in U-Net connect features from the initial encoding stage directly to the corresponding decoding stage, thus preserving important information that may be lost during the downsampling process [23].

2.3.1. UNet

UNet is a deep learning architecture based on Convolutional Neural Network (CNN) specifically designed for medical image segmentation tasks, including in the analysis of MRI scans of brain tumors

[24]. This model has become a standard method in pixel-based segmentation, where each pixel in the image is classified to determine the boundary between normal tissue and tumor. UNet uses an encoder-decoder approach with skip connections to improve segmentation accuracy and preserve spatial information in the image [25]. The main goal of UNet in brain tumor segmentation from MRI scans is to detect, classify, and isolate the tumor area from healthy brain tissue with a high degree of precision. Accurate segmentation is essential in early diagnosis, surgical planning, and evaluating the effectiveness of therapies such as radiation or chemotherapy [26]. By using UNet, doctors can obtain segmentation maps that show tumor boundaries in detail, allowing for more precise clinical decision making. In addition, this model is able to work with limited data through data augmentation and efficient utilization of spatial information [27]. Fig. 2 is a diagram of the UNet architecture.

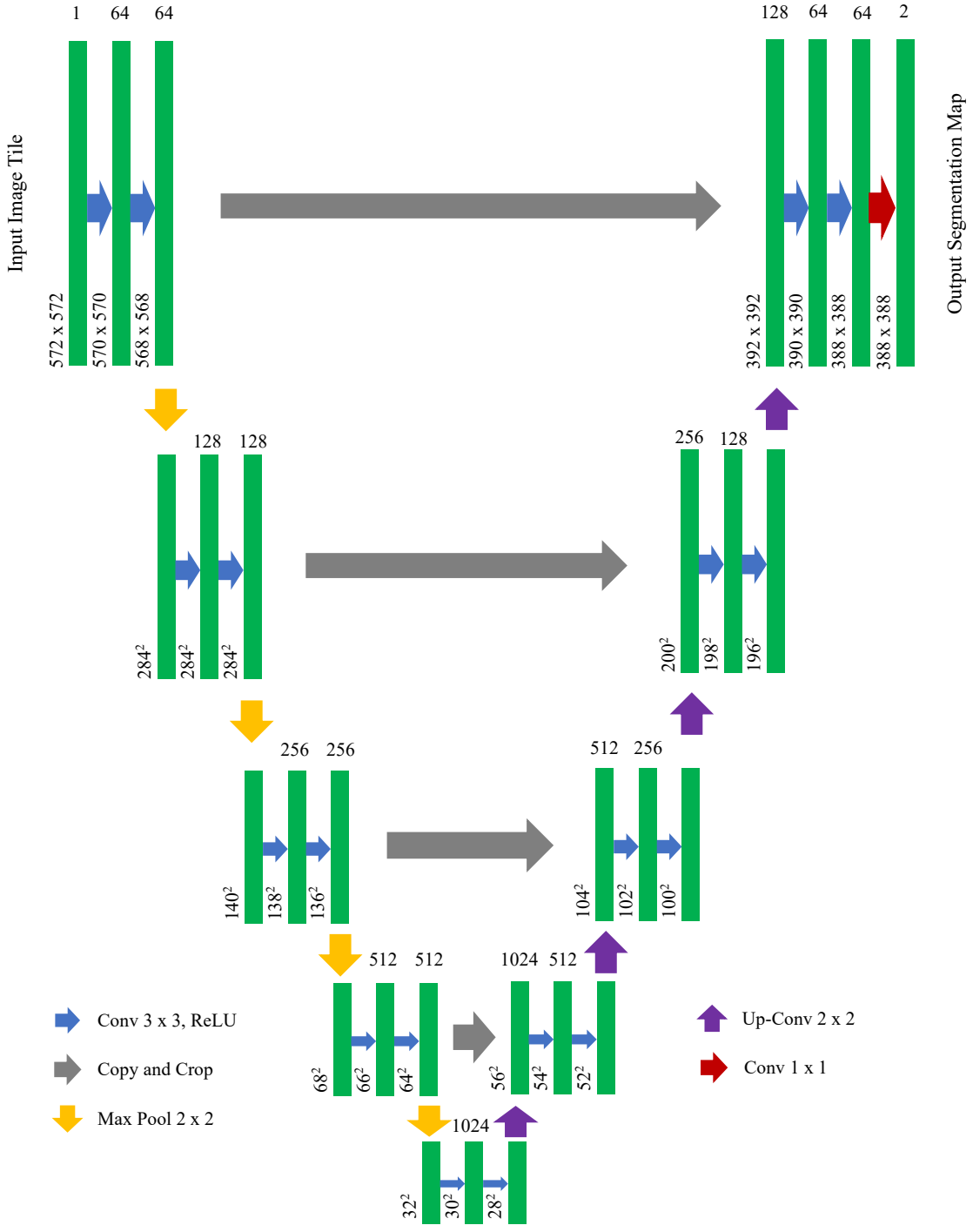

Fig. 2. UNet architecture

The diagram above illustrates the U-Net architecture, specifically designed for image segmentation tasks, including medical segmentation such as brain tumor detection from MRI scans [28]. This architecture begins with an input image tile of size 572 × 572 × 1 (grayscale) or larger for RGB format. The input is processed through an initial convolutional layer to produce an initial feature map with 64 channels. The U-Net structure consists of two main paths, such as the encoder (contracting path) and the decoder (expanding path), interconnected by skip connections [29]. The encoder is responsible for feature extraction using a series of stacked convolutional layers (3×3, ReLU) followed by max pooling (2×2) to progressively reduce the spatial dimensions. Each step in the encoder increases the number of feature channels, starting from 64 → 128 → 256 → 512 → 1024, while the spatial dimensions shrink from 572 → 284 → 140 → 68 → 32. The encoder aims to capture critical features at various levels of abstraction. At the center of the architecture lies the bottleneck, the deepest part with two stacked convolutional layers and 1024 channels, where abstract features are extracted to provide a richer representation [30]. These features are then passed to the decoder, which restores the spatial dimensions using up-convolution or transposed convolution (2×2). In the decoder, spatial dimensions progressively increase (e.g., from 32 → 68 → 140 → 284 → 572) while the number of channels decreases from 1024 → 512 → 256 → 128 → 64.

One of U-Net's strengths lies in the use of skip connections, which link the encoder layers to the corresponding decoder layers at the same resolution level [31]. These skip connections allow spatial information from the encoder to be transferred to the decoder, ensuring that critical details lost during downsampling are restored [32]. In the final stage, a 1×1 convolution layer is used to produce the output segmentation map with 2 channels, representing the pixel probabilities for each class (e.g., tumor or healthy tissue). The U-Net architecture is highly effective for segmentation as it combines global feature extraction in the encoder, spatial reconstruction in the decoder, and spatial information transfer through skip connections [33]. This structure ensures that the model can generate precise segmentation maps, even for complex images such as brain tumor MRI scans. This combination makes U-Net one of the most reliable models for deep learning-based segmentation tasks [34].

2.3.2. ResNet

ResNet (Residual Network) is a deep learning architecture based on Convolutional Neural Networks (CNN) designed to address the vanishing gradient problem that commonly occurs in very deep networks [35]. ResNet utilizes residual blocks with shortcut connections that allow information from earlier layers to be directly passed to subsequent layers. This approach enables the model to learn more complex features without performance degradation, making ResNet one of the most popular architectures for image segmentation tasks, including brain tumor segmentation from MRI scans [36]. The primary purpose of ResNet in brain tumor segmentation is to accurately identify and distinguish tumor regions from healthy tissue [37]. Accurate segmentation supported by ResNet plays a crucial role in early diagnosis, surgical planning, and therapy evaluation, such as radiation or chemotherapy. In segmentation tasks, ResNet is often used as the encoder in encoder-decoder architectures, such as ResNet-UNet, to capture deep features from MRI scans and detect relevant patterns in complex tumor structures [38].

ResNet's architecture consists of residual blocks, each containing two or three stacked convolutional layers (3×3) and shortcut connections that directly link the initial input to the block's output [39]. These shortcut connections employ an identity mapping mechanism that adds the input to the output after convolution, ensuring gradients flow efficiently through very deep networks. ResNet comes in various variants, such as ResNet-50, ResNet-101, and ResNet-152, differing in the number of layers and network depth, providing flexibility for tasks of varying complexity [40]. The working mechanism of ResNet begins with initial feature extraction using the first convolutional layer. Then, the image passes through a series of residual blocks where shortcut connections ensure that critical information from previous layers is preserved [41]. In the context of brain tumor MRI segmentation, ResNet is typically combined with a decoder to reconstruct the final segmentation map. The decoder enlarges the spatial dimensions of the image and generates a probability map for each

pixel, identifying whether the pixel belongs to the tumor region or healthy tissue [42]. Fig. 3, Fig. 4, Fig. 5, Fig. 6, Fig. 7 is a diagram of the ResNet architectures.

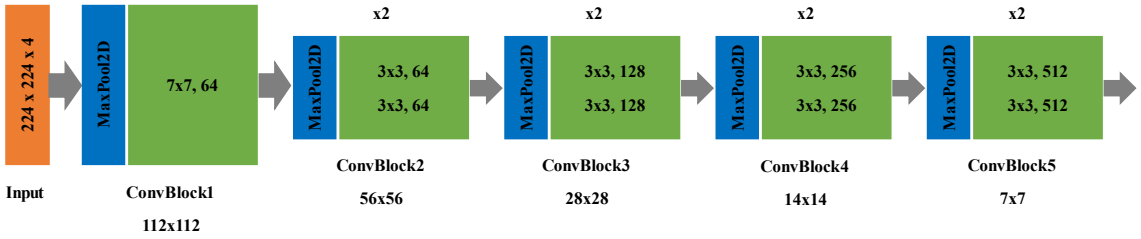

Fig. 3. ResNet-18

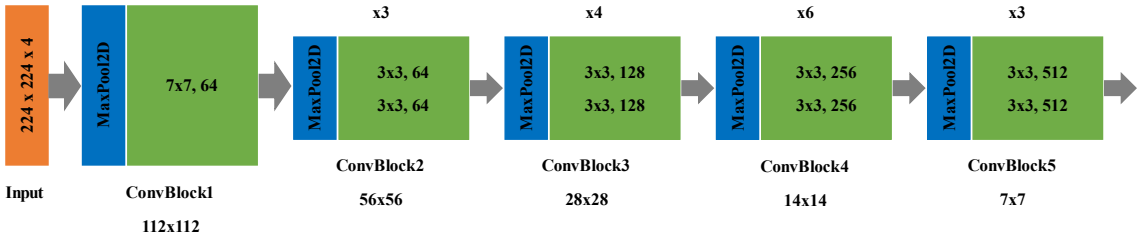

Fig. 4. ResNet-34

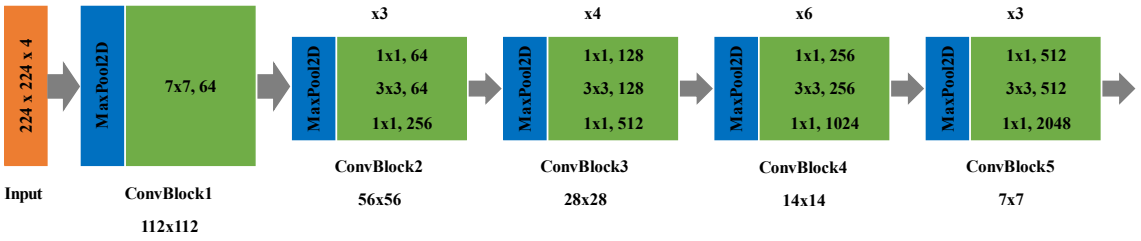

Fig. 5. ResNet-50

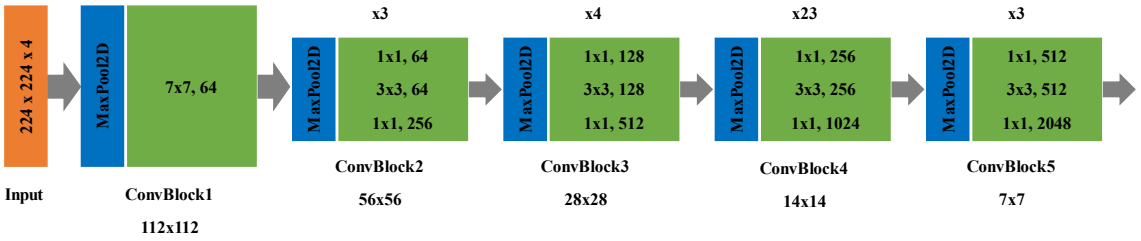

Fig. 6. ResNet-101

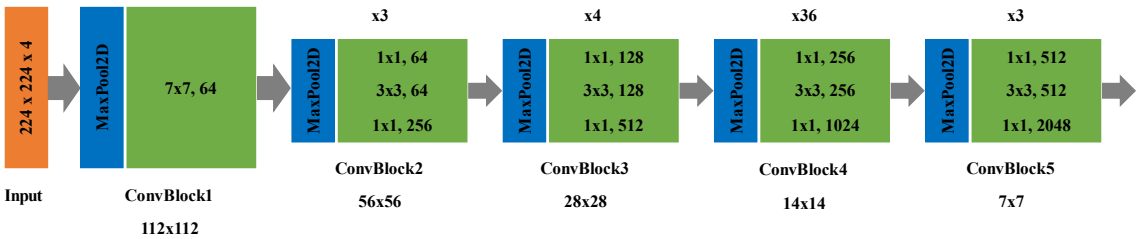

Fig. 7. ResNet-152

The ResNet model has several variants, namely ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152, which differ in the number of layers and network complexity [43]. Although all variants use residual blocks with shortcut connections, there are differences in the number of parameters, network depth, and the type of residual blocks used, which affect the model's performance and

efficiency [44]. ResNet18 and ResNet34 are lighter variants of ResNet, using basic residual blocks with two convolutional layers per block. ResNet18 consists of 18 layers and has fewer parameters, making it suitable for devices with limited computational power or tasks that do not require a very deep network. ResNet34 increases the depth to 34 layers, providing better accuracy while maintaining computational efficiency [45]. These two models are ideal for medical segmentation tasks with small to medium-sized datasets, as they do not require as much computational power as deeper models.

On the other hand, ResNet50, ResNet101, and ResNet152 use bottleneck residual blocks, which are more complex compared to the basic blocks in ResNet18 and ResNet34 [46]. Each bottleneck block consists of three convolutional layers (1×1, 3×3, and 1×1), where the 1×1 layers are used to reduce and expand feature dimensions, keeping the number of parameters efficient even in deeper networks. ResNet50 has 50 layers and more parameters than ResNet34, making it more powerful in capturing complex features [47]. ResNet101 increases the number of layers to 101, providing deeper feature representation and higher accuracy in segmentation and classification tasks [48]. ResNet152 is the deepest variant with 152 layers, offering higher accuracy but requiring significant computational power, making it commonly used in brain tumor segmentation with large datasets, as it can capture fine structural details of tumors more effectively [49]. In terms of performance and application, ResNet18 and ResNet34 are faster in inference and well-suited for simple segmentation tasks or real-time applications. ResNet50 is often the optimal choice as it provides a balance between accuracy and computational efficiency, making it widely used in deep learning-based segmentation models such as ResNet-UNet [50]. ResNet101 and ResNet152 offer deeper feature representations and are ideal for tasks requiring high-level feature extraction, but they demand high-powered GPUs for training and inference. In general, the choice between ResNet variants depends on the required accuracy, computational capacity, and dataset size [51]. For MRI-based brain tumor segmentation, ResNet50 or ResNet101 are frequently used as they offer the best balance between complexity, accuracy, and computational efficiency.

2.4. Hyperparameter in CNNs

In Convolutional Neural Networks (CNNs), hyperparameters play a crucial role in determining model performance, training stability, and generalization ability to new data [52]. Unlike model parameters such as weights and biases that are updated during training, hyperparameters are set before training begins and need to be optimized for the best results [53]. Some of the key hyperparameters in CNNs that significantly affect training effectiveness include learning rate, batch size, number of epochs, optimizer, and activation function [54]. The learning rate (α) is a hyperparameter that defines how much the model updates its weights during each iteration. A learning rate that is too high can cause unstable convergence, where the model overshoots the optimal point during optimization. Conversely, a learning rate that is too low can make training excessively slow or cause the model to get stuck in a local minimum [55]. Typically, the learning rate is set within the range of 0.0001 to 0.01, depending on the model architecture and dataset. Techniques such as learning rate decay or adaptive learning rate are often used to dynamically adjust the learning rate during training to improve optimization [56].

Another crucial hyperparameter is batch size, which determines the number of data samples processed before updating the model's weights in a single iteration. The choice of batch size significantly affects training stability and efficiency. A small batch size (e.g., 16 or 32) results in more frequent weight updates, which can help the model find optimal solutions faster but requires more iterations [57]. A large batch size (e.g., 128 or 256) allows the model to use more data before updating weights, improving training stability but requiring more GPU memory. The optimal batch size is usually determined through experimentation, as a batch size that is too large may reduce model generalization, while a batch size that is too small may make training noisy and unstable [58]. The number of epochs refers to the number of times the entire dataset is used to train the model. If the number of epochs is too low, the model may not learn effectively, leading to underfitting [59]. Conversely, if the number of epochs is too high, the model may memorize the training data and overfit, resulting in poor performance on new data. To prevent overfitting, techniques such as early stopping

are often used, which stop training when no further improvement is observed in model performance over several epochs. Typically, the number of epochs is chosen within the range of 10 to 100, depending on the dataset and model complexity [60].

An important aspect of CNN training is the optimizer, which determines how the model updates its weights based on gradient calculations [61]. Several commonly used optimizers include Stochastic Gradient Descent (which updates weights using gradients and is often combined with momentum to enhance stability), Adam (which dynamically adjusts the learning rate for each parameter and accelerates convergence), RMSprop (which stabilizes gradient updates and prevents exploding gradients), and Adagrad (which adapts the learning rate automatically but may reduce learning efficiency over time) [62]. The choice of optimizer depends on the dataset and model architecture, with Adam being widely used in medical image segmentation and image classification due to its robust performance in deep networks. Finally, the activation function plays a crucial role in introducing non-linearity to the network, allowing the model to learn complex patterns [63]. Common activation functions in CNNs include ReLU (which is widely used due to its simplicity and effectiveness in mitigating the vanishing gradient problem), Leaky ReLU (which allows small negative values to ensure gradient flow), Sigmoid (which is used in binary classification output layers but is less effective in hidden layers due to vanishing gradient issues), and Softmax (which is applied in multi-class classification tasks as it converts outputs into probability distributions) [64]. The choice of activation function depends on the specific task, with ReLU and Leaky ReLU commonly used in hidden layers, while Sigmoid and Softmax are applied in output layers depending on the classification type. The following are the hyperparameter tuning scenarios on the ResNet model tested in the study.

Based on Table 1, various ResNet architectures (18, 34, 50, 101, 152) are used to compare different network depths, where shallower models such as ResNet-18 and 34 are lighter and faster, while deeper models like ResNet-50, 101, and 152 can capture more complex features but are at risk of overfitting. The number of epochs is set to 50 to ensure the model learns sufficiently without overfitting. Three types of optimizers tested include Adam (known for its stability and fast convergence), SGD (more robust for generalization), and RMSProp (suitable for handling gradient fluctuations in deep learning). Additionally, three learning rate levels (0.01, 0.001, and 0.0001) are chosen to examine the impact of learning speed on model convergence, where 0.001 is commonly used as a standard due to its balance between stability and efficiency. This model employs the Sigmoid activation function, which is suitable for binary segmentation tasks as it produces probability values between 0 and 1 [65]. The selection of these hyperparameters aims to determine the best combination that can enhance Mean Intersection over Union (mIoU) and Dice Similarity Coefficient (DSC) in brain tumor segmentation using the ResNet-UNet model.

Table 1. Hyperparameter tuning on the ResNet architectures

ResNet Architectures	18, 34, 50, 101, 152
Epoch	50
Optimizer	Adam, SGD, RMSProp
Learning Rate	0.01, 0.001, 0.0001
Activation Function	Sigmoid

2.5. Model Performance Evaluation

Evaluating the performance of a segmentation model is crucial to ensure its accuracy and reliability in identifying regions of interest, such as brain tumors in MRI scans. In deep learning-based segmentation tasks, various metrics are used to measure how well the model predicts segmentation masks compared to the ground truth [66]. Among the most commonly used evaluation metrics are Loss, Dice Similarity Coefficient (DSC), and Mean Intersection over Union (mIoU). Each of these metrics provides valuable insights into different aspects of the model's segmentation performance.

2.5.1. Loss Function

The loss function quantifies the difference between the predicted segmentation and the ground truth, guiding the optimization process during model training. In MRI segmentation tasks, common loss functions include Binary Cross-Entropy (BCE) Loss, Dice Loss, and Combination Losses (such as BCE + Dice Loss) [67]. Binary Cross-Entropy (BCE) Loss is often used in binary segmentation tasks, where each pixel is classified as either tumor or non-tumor. However, BCE loss can be less effective when dealing with imbalanced datasets where the background class significantly outweighs the tumor class [68]. To address this, Dice Loss is commonly employed, which is derived from the Dice Similarity Coefficient (DSC). Dice Loss is particularly useful for segmentation tasks as it directly optimizes the overlap between predicted and ground truth masks. Many segmentation models also use hybrid loss functions that combine BCE and Dice Loss, ensuring both pixel-wise accuracy and global shape consistency in segmentation. The loss function helps the model learn to refine segmentation predictions over multiple training iterations, reducing false positives and false negatives [69].

2.5.2. Dice Similarity Coefficient (DSC)

The Dice Similarity Coefficient (DSC), also known as the Dice Score, is a widely used metric to measure the overlap between the predicted segmentation and the ground truth [70]. The DSC formula is given by:

$$DSC = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{1}$$

where A represents the predicted segmentation mask and B represents the ground truth mask. The DSC value ranges from 0 to 1, where 1 indicates perfect segmentation (complete overlap) and 0 indicates no overlap at all. A higher Dice Score suggests that the model is accurately identifying tumor regions, while a lower score implies poor segmentation performance. Dice Similarity is particularly useful in medical image segmentation because it effectively accounts for class imbalances, ensuring that the model focuses on overlapping areas rather than just pixel-wise accuracy. Since MRI brain tumor segmentation typically involves small tumor regions compared to the entire brain structure, DSC is often a preferred metric in performance evaluation [71].

2.5.3. Mean Intersection over Union (mIoU)

Another critical evaluation metric in MRI segmentation is Mean Intersection over Union (mIoU), which measures the ratio of intersection to the union between predicted and ground truth segmentations [72]. The mIoU formula is:

$$mIoU = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

Similar to the Dice Score, IoU (also called the Jaccard Index) evaluates the overlap between segmentation masks but tends to be more sensitive to small errors in boundary predictions. mIoU is commonly used in multi-class segmentation tasks, where it averages IoU scores across all classes to provide an overall segmentation accuracy measure. Like DSC, the higher the mIoU score, the better the segmentation accuracy. In MRI brain tumor segmentation, mIoU is particularly useful for assessing how well the model can distinguish tumor boundaries from healthy brain tissue. However, since mIoU is generally more conservative than DSC (i.e., it produces lower scores for the same segmentation results), it is often used in combination with DSC to get a more comprehensive evaluation of segmentation performance [73].

3. Results and Discussion

This study aims to develop a brain tumor segmentation model based on ResNet-UNet that has been optimized through hyperparameter tuning to improve segmentation accuracy and efficiency.

Therefore, a hybrid model between UNet and several ResNet architectures (18, 34, 50, 101, and 152) are proposed in this study. Each ResNet architecture is then tuned with hyperparameters by changing the learning rate (0.01, 0.001, and 0.0001), activation function type (Sigmoid), and optimizer type (Adam, SGD, and RMSProp). Based on the results of hyperparameter tuning and model testing, the performance metrics are obtained as shown in the following table.

The Table 2 presents the performance metrics of different ResNet models (ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152) in brain tumor MRI segmentation tasks. The evaluation was conducted using three different optimizers (Adam, SGD, and RMSProp) with a learning rate of 0.01 and the Sigmoid activation function. The performance was assessed using four key metrics, such as Loss, Mean Intersection over Union (mIoU), Dice Similarity Coefficient (DSC), and Iterations per Second. These metrics provide insights into how well each combination of model depth and optimizer performs in the segmentation task. Among the three optimizers, Adam and RMSProp produced lower loss values compared to SGD, indicating that these optimizers were more effective in minimizing segmentation errors. Adam and RMSProp consistently achieved lower loss values (0.006 - 0.008), whereas SGD produced higher loss values (0.010 - 0.012), suggesting that Adam and RMSProp are better suited for this segmentation task. In terms of Mean IoU and Dice Similarity Coefficient (DSC), RMSProp achieved the highest values across all models, with mIoU reaching up to 0.888 and DSC reaching 0.917, signifying superior segmentation accuracy. Adam also performed well, with slightly lower values than RMSProp, while SGD had the lowest mIoU and DSC scores, indicating weaker segmentation performance.

Table 2. Performance Metrics of ResNet model with LR = 0.01 and AF = Sigmoid

Learning Rate (LR) = 0.01 and Activation Function (AF) = Sigmoid					
Optimizer	**Model**	**Loss**	**Mean IoU**	**Dice Coefficient**	**Iterations/ Second**
Adam	ResNet18	0.006	0.887	0.914	8.63
	ResNet34	0.007	0.882	0.913	6.62
	ResNet50	0.008	0.875	0.903	4.65
	ResNet101	0.007	0.873	0.903	3.34
	ResNet152	0.007	0.872	0.900	2.66
SGD	ResNet18	0.012	0.797	0.833	8.85
	ResNet34	0.012	0.778	0.812	6.79
	ResNet50	0.019	0.640	0.640	4.79
	ResNet101	0.019	0.632	0.632	3.46
	ResNet152	0.011	0.813	0.845	2.83
RMSProp	ResNet18	0.006	0.887	0.916	9.19
	ResNet34	0.006	0.889	0.917	7.01
	ResNet50	0.007	0.880	0.909	4.94
	ResNet101	0.007	0.884	0.913	3.59
	ResNet152	0.007	0.888	0.917	2.61

The results show that deeper models (ResNet50, ResNet101, and ResNet152) tend to have higher accuracy (higher mIoU and DSC) but operate at slower processing speeds (iterations per second). ResNet18 and ResNet34 consistently performed segmentation at higher speeds (e.g., 8.630 and 6.620 iterations per second using Adam), whereas ResNet152 was the slowest, with speeds dropping to 2.660 iterations per second. The deeper models, particularly ResNet50, ResNet101, and ResNet152, achieved the highest segmentation accuracy, with mIoU exceeding 0.870 and DSC reaching 0.917 when trained with RMSProp. This suggests that deeper architectures are more effective in capturing complex tumor structures but require more computational resources. Based on the results, ResNet152 with RMSProp achieved the highest segmentation performance, with mIoU of 0.888 and DSC of 0.917, indicating precise tumor boundary identification. However, this came at the cost of lower inference speed (2.610 iterations per second). If computational efficiency is a priority, ResNet34 or ResNet50 with RMSProp may offer a good trade-off between accuracy and processing speed, with mIoU around 0.889 and DSC around 0.917 while maintaining a moderate number of iterations per second.

The Table 3 presents the performance metrics of different ResNet models (ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152) for brain tumor MRI segmentation, using a learning rate of 0.001 and the Sigmoid activation function. The models were evaluated with three different optimizers (Adam, SGD, and RMSProp), and their performance was measured using four key metrics, such as Loss, Mean Intersection over Union (mIoU), Dice Similarity Coefficient (DSC), and Iterations per Second. These metrics provide insights into the model's segmentation accuracy, efficiency, and computational performance. Among the three optimizers, Adam and RMSProp consistently achieved lower loss values, indicating their effectiveness in minimizing segmentation errors. Adam showed stable performance across all ResNet models, with loss values ranging from 0.006 to 0.007, while RMSProp produced similar values, particularly excelling in deeper models like ResNet50 and ResNet152. On the other hand, SGD had significantly higher loss values, ranging from 0.019 to 0.033, with lower segmentation accuracy, suggesting that it is less suitable for this task when using a learning rate of 0.001. In terms of Mean IoU (mIoU) and Dice Similarity Coefficient (DSC), RMSProp achieved the highest values across all models, with mIoU reaching 0.901 and DSC reaching 0.926 when used with ResNet152. Adam also performed well, achieving similar mIoU and Dice scores (0.897 - 0.904 for mIoU and 0.926 - 0.931 for DSC), making it a strong alternative to RMSProp. Meanwhile, SGD significantly underperformed, with mIoU values as low as 0.614 and Dice scores ranging from 0.623 to 0.681, indicating weaker segmentation accuracy compared to Adam and RMSProp.

Table 3. Performance Metrics of ResNet model with LR = 0.001 and AF = Sigmoid

Learning Rate (LR) = 0.001 and Activation Function (AF) = Sigmoid					
Optimizer	**Model**	**Loss**	**Mean IoU**	**Dice Coefficient**	**Iterations/ Second**
Adam	ResNet18	0.006	0.897	0.926	9.12
	ResNet34	0.007	0.897	0.926	6.93
	ResNet50	0.006	0.887	0.915	4.90
	ResNet101	0.006	0.904	0.931	3.55
	ResNet152	0.007	0.884	0.913	1.79
SGD	ResNet18	0.033	0.676	0.676	9.34
	ResNet34	0.034	0.614	0.639	7.16
	ResNet50	0.031	0.623	0.623	5.05
	ResNet101	0.024	0.726	0.757	3.52
	ResNet152	0.025	0.681	0.681	2.85
RMSProp	ResNet18	0.006	0.902	0.930	9.19
	ResNet34	0.007	0.882	0.910	7.03
	ResNet50	0.007	0.895	0.922	4.96
	ResNet101	0.006	0.872	0.901	3.59
	ResNet152	0.006	0.901	0.926	2.61

As observed in previous experiments, deeper models (ResNet50, ResNet101, and ResNet152) tend to achieve higher segmentation accuracy but operate at slower inference speeds (iterations per second). ResNet18 and ResNet34 had the highest processing speeds, with RMSProp and Adam achieving up to 9.190 and 9.120 iterations per second, respectively. Meanwhile, ResNet152 exhibited the slowest inference speed, dropping to 2.610 iterations per second, which indicates that deeper models require significantly more computational resources. However, deeper models such as ResNet50, ResNet101, and ResNet152 demonstrated superior segmentation accuracy, with mIoU exceeding 0.890 and DSC reaching 0.926 when using RMSProp or Adam. This confirms that deeper networks are more effective in capturing intricate tumor structures, albeit at the cost of processing speed. The results suggest that ResNet152 with RMSProp achieved the best segmentation performance, with an mIoU of 0.901 and DSC of 0.926, making it the most precise model in identifying tumor boundaries. However, this performance came with reduced inference speed (2.610 iterations per second). ResNet50 and ResNet101 with RMSProp or Adam provide a good balance between accuracy and computational efficiency, offering high Dice scores (above 0.920) while maintaining moderate processing speed.

The Table 4 presents the performance metrics of different ResNet models (ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152) for brain tumor MRI segmentation, using a learning rate of 0.0001 and the Sigmoid activation function. The models were evaluated with three different optimizers (Adam, SGD, and RMSProp), and their performance was measured using Loss, Mean Intersection over Union (mIoU), Dice Similarity Coefficient (DSC), and Iterations per Second. These metrics provide a comprehensive evaluation of the model's segmentation accuracy, computational efficiency, and processing speed. Among the three optimizers, Adam and RMSProp demonstrated significantly better segmentation performance compared to SGD, which produced notably higher loss values and lower accuracy metrics. Adam maintained loss values between 0.007 and 0.009, achieving a Mean IoU (mIoU) of up to 0.904 and a Dice Coefficient (DSC) of up to 0.928 when used with ResNet50. Similarly, RMSProp achieved comparable results, with loss values as low as 0.006, an mIoU of up to 0.902, and a DSC of 0.928. These results suggest that Adam and RMSProp are more suitable for brain tumor segmentation tasks, as they ensure better accuracy and lower error rates. In contrast, SGD performed significantly worse, with loss values reaching 0.108 for ResNet18 and 0.103 for ResNet152. Additionally, its mIoU values remained very low, ranging between 0.075 and 0.528, indicating poor segmentation quality. The Dice Coefficient for SGD did not exceed 0.528, which is considerably lower than the scores achieved by Adam and RMSProp. These results confirm that SGD struggles to converge effectively with a learning rate of 0.0001, making it a suboptimal choice for this task.

Table 4. Performance Metrics of ResNet model with LR = 0.0001 and AF = Sigmoid

Learning Rate (LR) = 0.0001 and Activation Function (AF) = Sigmoid					
Optimizer	**Model**	**Loss**	**Mean IoU**	**Dice Coefficient**	**Iterations/ Second**
Adam	ResNet18	0.007	0.877	0.908	9.01
	ResNet34	0.007	0.901	0.926	6.74
	ResNet50	0.007	0.904	0.932	4.65
	ResNet101	0.007	0.891	0.921	3.33
	ResNet152	0.010	0.858	0.888	2.65
SGD	ResNet18	0.108	0.304	0.304	8.86
	ResNet34	0.106	0.075	0.075	6.79
	ResNet50	0.106	0.055	0.062	4.92
	ResNet101	0.113	0.348	0.348	3.62
	ResNet152	0.103	0.528	0.528	2.80
RMSProp	ResNet18	0.006	0.897	0.925	8.82
	ResNet34	0.007	0.873	0.904	6.64
	ResNet50	0.006	0.901	0.928	4.69
	ResNet101	0.007	0.889	0.920	3.37
	ResNet152	0.006	0.902	0.928	2.38

The results also highlight the impact of model depth on segmentation performance. Deeper models (ResNet50, ResNet101, and ResNet152) consistently achieved higher segmentation accuracy, with mIoU surpassing 0.900 and Dice scores above 0.920 when using Adam or RMSProp. This indicates that deeper networks are more effective in capturing complex tumor structures [74]. However, deeper models also come with a trade-off in terms of computational efficiency, as observed in the iterations per second metric. ResNet18 and ResNet34 exhibited the highest processing speeds, with Adam and RMSProp achieving up to 9.01 and 9.19 iterations per second, respectively. Meanwhile, ResNet152 was the slowest, dropping to 2.38 iterations per second, emphasizing the increased computational burden of deeper architectures. The results suggest that ResNet50 and ResNet152 with RMSProp achieved the highest segmentation performance, with mIoU values of 0.902 and Dice Coefficients of 0.928. ResNet50 with Adam also performed exceptionally well, reaching an mIoU of 0.904 and a DSC of 0.928, making it an optimal choice for accurate tumor segmentation. However, if computational efficiency is a priority, ResNet34 with Adam or RMSProp offers a strong balance between accuracy and speed, achieving an mIoU of 0.901 while maintaining a moderate processing speed of 6.74 to 6.49 iterations per second.

The evaluation of hyperparameter tuning in brain tumor MRI segmentation demonstrates that parameters such as learning rate, optimizer selection, and activation function significantly influence model performance evaluation metrics like Loss Function, Dice Similarity Coefficient (DSC), and Mean Intersection over Union (mIoU) [75]. The results from different settings indicate how adjustments in these hyperparameters impact segmentation accuracy, model stability, and computational efficiency. One of the most crucial factors affecting model performance is the learning rate. A higher learning rate (0.01) results in unstable convergence, higher loss values, and lower segmentation accuracy. This is evident from the poor performance of models trained with SGD at a learning rate of 0.01, which exhibited higher loss values (above 0.012) and lower mIoU and Dice scores. As the learning rate was reduced to 0.001 and 0.0001, segmentation accuracy improved, particularly when combined with Adam and RMSProp optimizers. For instance, ResNet50 and ResNet152 with RMSProp at 0.0001 achieved lower loss values (0.006), an mIoU of 0.902, and a Dice score of 0.928, indicating that a lower learning rate stabilizes model training and enhances segmentation precision. The choice of optimizer also plays a critical role in determining how well the model can minimize segmentation errors [76]. The results indicate that Adam and RMSProp consistently outperformed SGD across all learning rate settings, delivering lower loss values, higher mIoU, and superior Dice scores. For example, at 0.0001 learning rate, ResNet50 and ResNet152 with RMSProp achieved the best segmentation accuracy, while SGD struggled with high loss values (0.103) and poor mIoU (as low as 0.075). This suggests that SGD is not well-suited for MRI segmentation tasks, as it struggles with convergence at lower learning rates. In contrast, Adam and RMSProp, which dynamically adjust learning rates, proved to be more effective for brain tumor segmentation.

Regarding activation functions, all models in the experiments utilized Sigmoid, which is commonly used in binary segmentation tasks like tumor vs. non-tumor classification. However, the results suggest that learning rate and optimizer selection have a greater impact on model accuracy than the choice of activation function itself. Since Sigmoid is prone to vanishing gradient problems in deep networks, alternative activation functions like ReLU or Leaky ReLU could potentially improve segmentation results in deeper architectures such as ResNet101 and ResNet152 [77]. Model depth also significantly influences segmentation performance. The deeper models (ResNet50, ResNet101, and ResNet152) consistently achieved higher segmentation accuracy, particularly when trained with Adam or RMSProp at lower learning rates. For instance, ResNet152 with RMSProp at 0.0001 attained the highest segmentation accuracy with an mIoU of 0.902 and DSC of 0.928, suggesting that deeper networks are better at capturing complex tumor structures. However, these models also come with a trade-off in terms of computational efficiency [78]. The iterations per second metric indicates that ResNet18 and ResNet34 process images faster (above 9.0 iterations per second), while deeper models like ResNet152 operate significantly slower (2.38 iterations per second), highlighting the increased computational burden of using deeper networks.

Based on the analysis, the optimal configuration for MRI brain tumor segmentation involves using a low learning rate (0.0001), Adam or RMSProp optimizer, and deeper ResNet architectures (ResNet50 or ResNet152). SGD is the least effective optimizer for this task, as it produces higher loss values and lower segmentation accuracy. While deeper models provide superior segmentation results, they require more computational resources, making them less practical for real-time applications. Therefore, for high-precision segmentation tasks, ResNet50 or ResNet152 with RMSProp or Adam at a learning rate of 0.0001 is the best choice. However, if computational efficiency is a priority, ResNet34 with Adam or RMSProp offers a strong balance between accuracy and speed. These findings highlight the importance of hyperparameter tuning in optimizing medical image segmentation models, demonstrating that the right combination of learning rate, optimizer, and model depth can significantly enhance segmentation accuracy and efficiency. However, this research still does not consider the imbalance of the LGG dataset used in brain tumor segmentation. This data imbalance can cause the model to tend to be more accurate in identifying non-tumor areas but fail to recognize tumor boundaries well [79]. Therefore, in future research, an approach using a balanced

sampling strategy is very important in improving the performance of brain tumor segmentation using UNet-ResNet, especially on datasets that have an imbalanced class distribution such as LGG.

4. Conclusion

This study successfully developed a brain tumor segmentation model based on ResNet-UNet, optimized through hyperparameter tuning to enhance segmentation accuracy and computational efficiency. The evaluation of different ResNet architectures (ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152) under various hyperparameter settings revealed that deeper models generally achieve better segmentation accuracy, while shallower models provide faster inference speeds. Among the tested hyperparameters, learning rate, optimizer selection, and activation function played a crucial role in model performance. The results demonstrate that lower learning rates (0.0001 and 0.001) led to improved segmentation accuracy, as they stabilized model convergence and minimized segmentation errors. The choice of optimizer significantly impacted model performance, with Adam and RMSProp consistently outperforming SGD across all learning rate settings. Models trained with SGD exhibited higher loss values, lower Dice Similarity Coefficient (DSC), and Mean Intersection over Union (mIoU), making it a less suitable optimizer for this task. Meanwhile, RMSProp achieved the highest segmentation accuracy, with an mIoU of 0.902 and DSC of 0.928 when applied to ResNet152, while Adam showed competitive results with similar scores. Additionally, the activation function (Sigmoid) was found to be effective for binary segmentation tasks, but alternative functions such as ReLU or Leaky ReLU could be explored to further enhance performance in deeper architectures.

The study also highlights the trade-off between segmentation accuracy and computational efficiency. Deeper models (ResNet50, ResNet101, and ResNet152) achieved superior accuracy in detecting tumor structures, but they required higher computational resources and exhibited slower inference speeds. In contrast, shallower models (ResNet18 and ResNet34) provided faster inference but lower accuracy, making them more suitable for real-time applications with computational constraints. The optimal balance between accuracy and efficiency was found in ResNet34 and ResNet50 with RMSProp or Adam, which maintained high Dice scores (above 0.920) while preserving a reasonable inference speed. While this study has demonstrated the effectiveness of hyperparameter tuning in brain tumor segmentation, several areas remain open for future research. One potential improvement is the exploration of alternative activation functions, such as ReLU, Leaky ReLU, or Swish, which may mitigate vanishing gradient problems in deep ResNet architectures. Additionally, future studies could investigate adaptive learning rate strategies, such as cyclical learning rates or learning rate warm-up techniques, to further optimize model convergence and prevent underfitting or overfitting. Another important area for future research is the integration of more advanced optimization techniques, including gradient accumulation or adaptive gradient methods like AdaBelief or NovoGrad, which may enhance segmentation accuracy while maintaining computational efficiency. Furthermore, incorporating data augmentation and transfer learning techniques could improve model generalization, particularly for handling small or imbalanced datasets in medical imaging.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Acknowledgment: The authors would like to acknowledge the Department of Medical Technology, Institut Teknologi Sepuluh Nopember, for the facilities and support in this research. The authors also gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under project scheme of the Publication Writing and IPR Incentive Program (PPHKI) 2025.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Y. Yang *et al*., "Segmentation method of magnetic resonance imaging brain tumor images based on improved UNet network," *Translational Cancer Research*, vol. 13, no. 3, pp. 1567–1583, 2024, https://doi.org/10.21037/tcr-23-1858.

[2] A. Lerner *et al*., "Gliomas in adults: Guidance on investigations, diagnosis, treatment and surveillance," *Clinical Medicine*, vol. 24, no. 5, p. 100240, 2024, https://doi.org/10.1016/j.clinme.2024.100240.

[3] D. Rastogi *et al*., "Deep learning-integrated MRI brain tumor analysis: feature extraction, segmentation, and Survival Prediction using Replicator and volumetric networks," *Scientific Reports*, vol. 15, no. 1, 2025, https://doi.org/10.1038/s41598-024-84386-0.

[4] D. Veiga-Canuto *et al*., "Comparative Multicentric Evaluation of Inter-Observer Variability in Manual and Automatic Segmentation of Neuroblastic Tumors in Magnetic Resonance Images," *Cancers*, vol. 14, no. 15, p. 3648, 2022, https://doi.org/10.3390/cancers14153648.

[5] Md. E. Rayed, S.M. S. Islam, S. I. Niha, J. R. Jim, Md. M. Kabir, and M.F. Mridha, "Deep learning for medical image segmentation: State-of-the-art advancements and challenges," *Informatics in medicine unlocked*, vol. 47, p. 101504, 2024, https://doi.org/10.1016/j.imu.2024.101504.

[6] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging," *Cancers*, vol. 15, no. 16, p. 4172, 2023, https://doi.org/10.3390/cancers15164172.

[7] Y. Feng, J. Li, and X. Zhang, "Research on Segmentation of Brain Tumor in MRI Image Based on Convolutional Neural Network," *BioMed Research International*, vol. 2022, no. 1, pp. 1–9, 2022, https://doi.org/10.1155/2022/7911801.

[8] L. P.-Coelho, "How Artificial Intelligence Is Shaping Medical Imaging Technology: a Survey of Innovations and Applications," *Bioengineering*, vol. 10, no. 12, p. 1435, 2023, https://doi.org/10.3390/bioengineering10121435.

[9] P. Zheng, X. Zhu, and W. Guo, "Brain tumour segmentation based on an improved U-Net," *BMC Medical Imaging*, vol. 22, no. 1, 2022, https://doi.org/10.1186/s12880-022-00931-1.

[10] F. Ullah *et al*., "Brain Tumor Segmentation from MRI Images Using Handcrafted Convolutional Neural Network," *Diagnostics*, vol. 13, no. 16, pp. 2650–2650, 2023, https://doi.org/10.3390/diagnostics13162650.

[11] M. Aggarwal, A. K. Tiwari, M. P. Sarathi, and A. Bijalwan, "An early detection and segmentation of Brain Tumor using Deep Neural Network," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, 2023, https://doi.org/10.1186/s12911-023-02174-8.

[12] J. Walsh, A. Othmani, M. Jain, and S. Dev, "Using U-Net network for efficient brain tumor segmentation in MRI images," *Healthcare Analytics*, vol. 2, p. 100098, 2022, https://doi.org/10.1016/j.health.2022.100098.

[13] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. J. Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Computers in Biology and Medicine*, vol. 152, p. 106405, 2023, https://doi.org/10.1016/j.compbiomed.2022.106405.

[14] M. Buda, A. Saha, and M. A. Mazurowski, "Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm," *Computers in Biology and Medicine*, vol. 109, pp. 218–225, 2019, https://doi.org/10.1016/j.compbiomed.2019.05.002.

[15] O. Diaz *et al*., "Data preparation for artificial intelligence in medical imaging: A comprehensive guide to open-access platforms and tools," *Physica Medica*, vol. 83, pp. 25–37, 2021, https://doi.org/10.1016/j.ejmp.2021.02.007.

[16] D. Liu *et al*., "Imaging-Genomics in Glioblastoma: Combining Molecular and Imaging Signatures," *Frontiers in Oncology*, vol. 11, 2021, https://doi.org/10.3389/fonc.2021.699265.

[17] M. F. Ahamed *et al*., "A review on brain tumor segmentation based on deep learning methods with federated learning techniques," *Computerized Medical Imaging and Graphics*, vol. 110, p. 102313, 2023, https://doi.org/10.1016/j.compmedimag.2023.102313.

[18] B. S. Alemu, S. Feisso, E. A. Mohammed, and A. O. Salau, "Magnetic resonance imaging-based brain tumor image classification performance enhancement," *Scientific African*, vol. 22, p. e01963, 2023, https://doi.org/10.1016/j.sciaf.2023.e01963.

[19] S. Chen, S. Zhao, and Q. Lan, "Residual Block Based Nested U-Type Architecture for Multi-Modal Brain Tumor Image Segmentation," *Frontiers in Neuroscience*, vol. 16, 2022, https://doi.org/10.3389/fnins.2022.832824.

[20] B. Sariturk and D. Z. Seker, "A Residual-Inception U-Net (RIU-Net) Approach and Comparisons with U-Shaped CNN and Transformer Models for Building Segmentation from High-Resolution Satellite Images," *Sensors*, vol. 22, no. 19, p. 7624, 2022, https://doi.org/10.3390/s22197624.

[21] S. Arshad, T. Amjad, A. Hussain, I. Qureshi, and Q. Abbas, "Dermo-Seg: ResNet-UNet Architecture and Hybrid Loss Function for Detection of Differential Patterns to Diagnose Pigmented Skin Lesions," *Diagnostics*, vol. 13, no. 18, p. 2924, 2023, https://doi.org/10.3390/diagnostics13182924.

[22] L. Zhao, H. Zhang, X. Sun, Z. Ouyang, C. Xu, and X. Qin, "Application of ResUNet-CBAM in Thin-Section Image Segmentation of Rocks," *Information*, vol. 15, no. 12, p. 788, 2024, https://doi.org/10.3390/info15120788.

[23] Q. Zhang, G. Geng, P. Zhou, Q. Liu, Y. Wang, and K. Li, "Link Aggregation for Skip Connection–Mamba: Remote Sensing Image Segmentation Network Based on Link Aggregation Mamba," *Remote Sensing*, vol. 16, no. 19, p. 3622, 2024, https://doi.org/10.3390/rs16193622.

[24] R. Yousef *et al*., "U-Net-Based Models towards Optimal MR Brain Image Segmentation," *Diagnostics*, vol. 13, no. 9, p. 1624, 2023, https://doi.org/10.3390/diagnostics13091624.

[25] X. Zhang, S. Yang, Y. Jiang, Y. Chen, and F. Sun, "FAFS-UNet: Redesigning skip connections in UNet with feature aggregation and feature selection," *Computers in Biology and Medicine*, vol. 170, p. 108009, 2024, https://doi.org/10.1016/j.compbiomed.2024.108009.

[26] D. Rastogi *et al*., "Deep learning-integrated MRI brain tumor analysis: feature extraction, segmentation, and Survival Prediction using Replicator and volumetric networks," *Scientific Reports*, vol. 15, no. 1, 2025, https://doi.org/10.1038/s41598-024-84386-0.

[27] W. Zafar *et al*., "Enhanced TumorNet: Leveraging YOLOv8s and U-Net for Superior Brain Tumor Detection and Segmentation Utilizing MRI Scans," *Results in Engineering*, vol. 24, p. 102994, 2024, https://doi.org/10.1016/j.rineng.2024.102994.

[28] C. Li, W. Chen, and Y. Tan, "Point-Sampling Method Based on 3D U-Net Architecture to Reduce the Influence of False Positive and Solve Boundary Blur Problem in 3D CT Image Segmentation," *Applied Sciences*, vol. 10, no. 19, p. 6838, 2020, https://doi.org/10.3390/app10196838.

[29] A. Santone, R. D. Vivo, L. Recchia, M. Cesarelli, and F. Mercaldo, "A Method for Retina Segmentation by Means of U-Net Network," *Electronics*, vol. 13, no. 22, p. 4340, 2024, https://doi.org/10.3390/electronics13224340.

[30] T. Singha, D.-S. Pham, and A. Krishna, "Improved Short-term Dense Bottleneck network for efficient scene analysis," *Computer Vision and Image Understanding*, vol. 235, p. 103795, 2023, https://doi.org/10.1016/j.cviu.2023.103795.

[31] J. Nodirov, A. B. Abdusalomov, and T. K. Whangbo, "Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images," *Sensors*, vol. 22, no. 17, p. 6501, 2022, https://doi.org/10.3390/s22176501.

[32] L. Gui, X. Gu, F. Huang, S. Ren, H. Qin, and C. Fan, "Road Extraction from Remote Sensing Images Using a Skip-Connected Parallel CNN-Transformer Encoder-Decoder Model," *Applied Sciences*, vol. 15, no. 3, p. 1427, 2025, https://doi.org/10.3390/app15031427.

[33] Y.-T. Chen, L. Chang, and J.-H. Wang, "Full-Scale Aggregated MobileUNet: An Improved U-Net Architecture for SAR Oil Spill Detection," *Sensors*, vol. 24, no. 12, p. 3724, 2024, https://doi.org/10.3390/s24123724.

[34] G. M. Kumar and E. Parthasarathy, "Development of an enhanced U-Net model for brain tumor segmentation with optimized architecture," *Biomedical Signal Processing and Control*, vol. 81, pp. 104427, 2023, https://doi.org/10.1016/j.bspc.2022.104427.

[35] M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," *Applied Sciences*, vol. 12, no. 18, p. 8972, 2022, https://doi.org/10.3390/app12188972.

[36] L. Alzubaidi *et al*., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, 2021, https://doi.org/10.1186/s40537-021-00444-8.

[37] L. H. Shehab, O. M. Fahmy, S. M. Gasser, and M. S. El-Mahallawy, "An efficient brain tumor image segmentation based on deep residual networks (ResNets)," *Journal of King Saud University - Engineering Sciences*, vol. 33, no. 6, pp. 404–412, 2021, https://doi.org/10.1016/j.jksues.2020.06.001.

[38] Y. Gowda, D. R. V. Manjunath, Dr. Shubha B, Dr. Punya Prabha, Prof. Aishwarya N, and None ManuH M, "Deep Learning Technique for automatic Liver and Liver Tumor Segmentation in CT Images," *Journal of Liver Transplantation*, p. 100251, 2024, https://doi.org/10.1016/j.liver.2024.100251.

[39] L. Atika, S. Nurmaini, R. U. Partan, and E. Sukandi, "Image Segmentation for Mitral Regurgitation with Convolutional Neural Network Based on UNet, Resnet, Vnet, FractalNet and SegNet: A Preliminary Study," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 141, 2022, https://doi.org/10.3390/bdcc6040141.

[40] S. A. Hasanah, A. A. Pravitasari, A. S. Abdullah, I. N. Yulita, and M. H. Asnawi, "A Deep Learning Review of ResNet Architecture for Lung Disease Identification in CXR Image," *Applied sciences*, vol. 13, no. 24, p. 13111, 2023, https://doi.org/10.3390/app132413111.

[41] S. A. Samad and J. Gitanjali, "Augmenting DenseNet: Leveraging Multi-Scale Skip Connections for Effective Early-Layer Information Incorporation," *IEEE Access*, vol. 12, pp. 141344–141360, 2024, https://doi.org/10.1109/access.2024.3460830.

[42] Z. Liu *et al*., "Deep learning-based brain tumor segmentation: a survey," *Complex & Intelligent Systems*, vol. 9, pp. 1001-1026, 2022, https://doi.org/10.1007/s40747-022-00815-5.

[43] G. Wu, X. Ji, G. Yang, Y. Jia, and C. Cao, "Signal-to-Image: Rolling Bearing Fault Diagnosis Using ResNet Family Deep-Learning Models," *Processes*, vol. 11, no. 5, p. 1527, 2023, https://doi.org/10.3390/pr11051527.

[44] W. Cai, M. Li, G. Jin, Q. Liu, and C. Lu, "Comparison of Residual Network and Other Classical Models for Classification of Interlayer Distresses in Pavement," *Applied Sciences*, vol. 14, no. 15, p. 6568, 2024, https://doi.org/10.3390/app14156568.

[45] J. Li *et al*., "Detection of bruising in pear with varying bruising degrees and formation times by using SIRI technique combining with texture feature-based LS-SVM and ResNet-18-based CNN model," *Postharvest Biology and Technology*, vol. 223, p. 113434, 2025, https://doi.org/10.1016/j.postharvbio.2025.113434.

[46] H. E. Kim, A. C.-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging*, vol. 22, no. 1, 2022, https://doi.org/10.1186/s12880-022-00793-7.

[47] S. R. S. Chakravarthy, N. Bharanidharan, C. Vinothini, V. V. Kumar, T. R. Mahesh, and S. Guluwadi, "Adaptive Mish activation and ranger optimizer-based SEA-ResNet50 model with explainable AI for multiclass classification of COVID-19 chest X-ray images," *BMC Medical Imaging*, vol. 24, no. 1, 2024, https://doi.org/10.1186/s12880-024-01394-2.

[48] Q. Zhang, "A novel ResNet101 model based on dense dilated convolution for image classification," *SN Applied Sciences*, vol. 4, no. 1, 2021, https://doi.org/10.1007/s42452-021-04897-7.

[49] Y. Pande and J. Chaki, "Brain tumor detection across diverse MR images: An automated triple-module approach integrating reduced fused deep features and machine learning," *Results in Engineering*, vol. 25, p. 103832, 2024, https://doi.org/10.1016/j.rineng.2024.103832.

[50] Q. Li, X. Xu, J. Guan, and H. Yang, "The Improvement of Faster-RCNN Crack Recognition Model and Parameters Based on Attention Mechanism," *Symmetry*, vol. 16, no. 8, p. 1027, 2024, https://doi.org/10.3390/sym16081027.

[51] V. Joshi *et al*., "Accurate deep neural network inference using computational phase-change memory," *Nature Communications*, vol. 11, no. 1, p. 2473, 2020, https://doi.org/10.1038/s41467-020-16108-9.

[52] M. A. K. Raiaan *et al*., "A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks," *Decision Analytics Journal*, vol. 11, p. 100470, 2024, https://doi.org/10.1016/j.dajour.2024.100470.

[53] F. Z. E.-Hassani, M. Amri, N.-E. Joudar, and K. Haddouch, "A New Optimization Model for MLP Hyperparameter Tuning: Modeling and Resolution by Real-Coded Genetic Algorithm," *Neural Processing Letters*, vol. 56, no. 2, 2024, https://doi.org/10.1007/s11063-024-11578-0.

[54] S. Mezzah and A. Tari, "Practical Hyperparameters Tuning of Convolutional Neural Networks for EEG Emotional Features Classification," *Intelligent Systems with Applications*, vol. 18, p. 200212, 2023, https://doi.org/10.1016/j.iswa.2023.200212.

[55] A. Paviglianiti, V. Randazzo, S. Villata, G. Cirrincione, and E. Pasero, "A Comparison of Deep Learning Techniques for Arterial Blood Pressure Prediction," *Cognitive Computation*, vol. 14, no. 5, pp. 1689–1710, 2021, https://doi.org/10.1007/s12559-021-09910-0.

[56] N. Bhushan, S. Mekhilef, K. S. Tey, M. Shaaban, M. Seyedmahmoudian, and A. Stojcevski, "Dynamic K-Decay Learning Rate Optimization for Deep Convolutional Neural Network to Estimate the State of Charge for Electric Vehicle Batteries," *Energies*, vol. 17, no. 16, p. 3884, 2024, https://doi.org/10.3390/en17163884.

[57] J.-S. Hwang, S.-S. Lee, J.-W. Gil, and C.-K. Lee, "Determination of Optimal Batch Size of Deep Learning Models with Time Series Data," *Sustainability*, vol. 16, no. 14, p. 5936, 2024, https://doi.org/10.3390/su16145936.

[58] I. Kandel and M. Castelli, "The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset," *ICT Express*, vol. 6, no. 4, p. 312-315, 2020, https://doi.org/10.1016/j.icte.2020.04.010.

[59] C. Xu, P. Coen-Pirani, and X. Jiang, "Empirical Study of Overfitting in Deep Learning for Predicting Breast Cancer Metastasis," *Cancers*, vol. 15, no. 7, p. 1969, 2023, https://doi.org/10.3390/cancers15071969.

[60] S. F. Ahmed *et al*., "Deep learning modelling techniques: current progress, applications, advantages, and challenges," *Artificial Intelligence Review*, vol. 56, p. 13521–13617, 2023, https://doi.org/10.1007/s10462-023-10466-8.

[61] Z. Hong and C. P. Yue, "Efficient-Grad: Efficient Training Deep Convolutional Neural Networks on Edge Devices with Grad ient Optimizations," *ACM Transactions on Embedded Computing Systems*, vol. 21, no. 2, pp. 1–24, 2022, https://doi.org/10.1145/3504034.

[62] Y. Shao *et al*., "An Improved BGE-Adam Optimization Algorithm Based on Entropy Weighting and Adaptive Gradient Strategy," *Symmetry*, vol. 16, no. 5, p. 623, 2024, https://doi.org/10.3390/sym16050623.

[63] X. Zhao, L. Wang, Y. Zhang, X. Han, Muhammet Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artificial Intelligence Review*, vol. 57, no. 4, 2024, https://doi.org/10.1007/s10462-024-10721-6.

[64] M. M. Noel, S. Bharadwaj, V. M.-Nakarajan, P. Dutta, and G. Bessie, "Biologically inspired oscillating activation functions can bridge the performance gap between biological and artificial neurons," *Expert Systems with Applications*, vol. 266, p. 126036, 2024, https://doi.org/10.1016/j.eswa.2024.126036.

[65] B. C. Sweetline, C. Vijayakumaran, and A. Samydurai, "Overcoming the Challenge of Accurate Segmentation of Lung Nodules: A Multi-crop CNN Approach," *Deleted Journal*, vol. 37, no. 3, pp. 988–1007, 2024, https://doi.org/10.1007/s10278-024-01004-1.

[66] E. S. Biratu, F. Schwenker, Y. M. Ayano, and T. G. Debelee, "A Survey of Brain Tumor Segmentation and Classification Algorithms," *Journal of Imaging*, vol. 7, no. 9, p. 179, 2021, https://doi.org/10.3390/jimaging7090179.

[67] Q. D. Nguyen and H.-T. Thai, "Crack segmentation of imbalanced data: The role of loss functions," *Engineering Structures*, vol. 297, p. 116988, 2023, https://doi.org/10.1016/j.engstruct.2023.116988.

[68] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, 2022, https://doi.org/10.1016/j.compmedimag.2021.102026.

[69] P. Aliniya, M. Nicolescu, M. Nicolescu, and G. Bebis, "Improved Loss Function for Mass Segmentation in Mammography Images Using Density and Mass Size," *Journal of Imaging*, vol. 10, no. 1, p. 20, 2024, https://doi.org/10.3390/jimaging10010020.

[70] K. H. Zou *et al*., "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004, https://doi.org/10.1016/S1076-6332(03)00671-8.

[71] P. Wang, Q. Yang, Z. He, and Y. Yuan, "Vision transformers in multi-modal brain tumor MRI segmentation: A review," *Meta-Radiology*, vol. 1, no. 1, p. 100004, 2023, https://doi.org/10.1016/j.metrad.2023.100004.

[72] Muhamad *et al*., "Multi-modal deep learning approaches to semantic segmentation of mining footprints with multispectral satellite imagery," *Remote Sensing of Environment*, vol. 318, p. 114584, 2024, https://doi.org/10.1016/j.rse.2024.114584.

[73] M. Yeung, L. Rundo, N. Yang, E. Sala, C.-B. Schönlieb, and G. Yang, "Calibrating the Dice Loss to Handle Neural Network Overconfidence for Biomedical Image Segmentation," *Journal of Digital Imaging*, vol. 36, no. 2, pp. 739–752, 2022, https://doi.org/10.1007/s10278-022-00735-3.

[74] V. Kumar, C. Prabha, P. Sharma, N. Mittal, S. S. Askar, and M. Abouhawwash, "Unified deep learning models for enhanced lung cancer prediction with ResNet-50–101 and EfficientNet-B3 using DICOM images," *BMC Medical Imaging*, vol. 24, no. 1, 2024, https://doi.org/10.1186/s12880-024-01241-4.

[75] M. Ashimgaliyev, B. Matkarimov, A. Barlybayev, R. Yi, and A. Zhumadillayeva, "Accurate MRI-Based Brain Tumor Diagnosis: Integrating Segmentation and Deep Learning Approaches," *Applied Sciences*, vol. 14, no. 16, p. 7281, 2024, https://doi.org/10.3390/app14167281.

[76] E. Hassan, M. Y. Shams, N. A. Hikal, and S. Elmougy, "The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study," *Multimedia Tools and Applications*, vol. 82, pp. 16591–16633, 2022, https://doi.org/10.1007/s11042-022-13820-0.

[77] S. Kumar, H. Kumar, G. Kumar, S. P. Singh, A. Bijalwan, and M. Diwakar, "A methodical exploration of imaging modalities from dataset to detection through machine learning paradigms in prominent lung disease diagnosis: a review," *BMC Medical Imaging*, vol. 24, no. 1, 2024, https://doi.org/10.1186/s12880-024-01192-w.

[78] N. Cinar, A. Ozcan, and M. Kaya, "A hybrid DenseNet121-UNet model for brain tumor segmentation from MR Images," *Biomedical Signal Processing and Control*, vol. 76, p. 103647, 2022, https://doi.org/10.1016/j.bspc.2022.103647.

[79] P. Wang and A. C. S. Chung, "Relax and focus on brain tumor segmentation," *Medical Image Analysis*, vol. 75, p. 102259, 2022, https://doi.org/10.1016/j.media.2021.102259.