Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

Polyp Identification from a Colonoscopy Image Using Semantic Segmentation Approach

Wahyu Hauzan Rafi, Mahmud Dwi Sulistiyo*, Sugondo Hadiyoso, Untari Novia Wisesty

Faculty of Informatics, Informatics, Telkom University, Bandung, Indonesia
Email: [1]wahyuhr@student.telkomuniversity.ac.id, [2]mahmuddwis@telkomuniversity.ac.id, [3]sugondo@telkomuniversity.ac.id, [4]untarinw@telkomuniversity.ac.id
Correspondence Author Email: mahmuddwis@telkomuniversity.ac.id
Submitted: **15/08/2023**; Accepted: **25/09/2023**; Published**: 27/09/2023**

Abstract−Colorectal Cancer (CRC) is a major contributor to cancer-related mortality worldwide, necessitating early detection and treatment of polyps to prevent cancer progression. A colonoscopy is a critical diagnostic procedure for identifying colon abnormalities and removing premalignant polyps. However, accurately segmenting polyps in colonoscopy images poses challenges due to their diverse appearance and indistinct boundaries. In this study, we investigate augmentation techniques to enhance polyp semantic segmentation using the U-Net model. Our analysis reveals that the most effective technique is found in sub-scenario 2.6.c with an input size of 320×320, striking a favorable balance between accuracy and efficiency. Additionally, we explore the benefits of larger input sizes, taking into account resource considerations. Moreover, we conduct further testing of the best augmentation technique identified in previous experiments with the SegNet model. The results show a 3.5% improvement in the dice coefficient and slightly better qualitative outcomes. However, it is important to note that this enhancement comes with a nearly fivefold increase in training time. Moving forward, our objective is to develop a unified model for segmenting diverse medical images, pushing the boundaries of polyp detection and medical imaging. This research provides valuable insights and lays the foundation for more advanced applications in polyp detection and medical image analysis.

Keywords: Semantic Segmentation; Polyps; Colonoscopy; U-Net; SegNet.

1. INTRODUCTION

Colorectal Cancer (CRC) stands as a significant contributor to cancer-related mortality in the United States, ranking third among the leading causes of such deaths in both men and women [1]. This form of cancer often originates from the development of polyps, abnormal growths that can affect various parts of the body, including the nasal passages and the digestive system. Detecting and addressing these polyps in their early stages is critical [2]. Colonoscopy, a diagnostic procedure that examines the large intestine (colon) for anomalies or changes, plays a pivotal role in the prevention of CRC. It enables healthcare professionals to identify and remove slow-growing precancerous polyps, thereby preventing their progression into CRC [3].

The effectiveness of colonoscopy in CRC examination and prevention is well-established, and early colonoscopy has contributed to a notable 30.4% reduction in CRC cases [4]. However, the precise segmentation of polyps remains a challenging task due to two primary reasons. Firstly, polyps exhibit significant variations in appearance, such as size, color, and texture, even when they belong to the same category. Secondly, the boundaries between polyps and the surrounding mucosa are often blurred and lack the distinct contrast required for segmentation in colonoscopy images [5]. These challenges have resulted in inaccurate polyp segmentation, leading to instances where a considerable number of polyp cases remain undetected during colonoscopy examinations [6]–[8].

In addressing these complexities and enhancing semantic segmentation, this investigation utilizes the U-Net as the primary architectural model and incorporates SegNet for comparative performance assessment. Both neural networks have been purposefully crafted for the task of image segmentation, particularly within the domain of medical imaging. It is noteworthy to mention that while U-Net is well-established for its capabilities in medical image segmentation, SegNet was originally developed with a focus on autonomous vehicle applications.

Efforts to achieve automatic polyp segmentation have been active over the past two decades, focusing on developing efficient methods and algorithms. Previous studies have concentrated on analyzing the color and texture characteristics of polyps through manually generated descriptor-based feature learning [9].

The field of medical image segmentation has undergone significant advancements, largely driven by the groundbreaking work of Ronneberger et al. [10], who introduced U-Net architecture. This neural network model, meticulously crafted for the specific demands of biomedical image segmentation, has played a pivotal role in achieving precise object delineation within medical images. Its distinctive contracting and expanding pathway design have established it as an indispensable tool for various medical imaging tasks. Notably, U-Net's influence on polyp segmentation within colonoscopy images has been substantial, with previous studies, such as the one conducted by Jha et al. [11], reporting noteworthy achievements. These studies harnessed variations of the U-Net model, like ResUNet, and applied them to datasets like Kvasir-SEG, where remarkable results were achieved. For example, Jha et al. [11] reported a Dice coefficient accuracy of 0.787763 and a Mean IoU of 0.777771 when employing ResUNet on the Kvasir-SEG dataset.

While U-Net has undeniably become a cornerstone in the realm of medical image segmentation, the SegNet architecture, introduced by Badrinarayanan et al. [12], offers an equally compelling approach to image segmentation. Originally tailored for autonomous vehicle applications, SegNet's deep convolutional encoder-decoder structure has found its way into the domain of medical imaging, including the critical task of polyp segmentation within

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

colonoscopy images. Its proficiency in capturing intricate image details and features has garnered significant attention, as evidenced by studies such as Lin et al. [13], which underscore its potential in medical image segmentation. In fact, Lin et al. [13] reported promising results in medical image segmentation tasks, achieving a Dice coefficient accuracy of 0.85 when applying the RefineU-Net, a U-Net variant enriched with progressive global feedback and residual attention-guided local refinement.

Furthermore, the emergence of novel U-Net variants, such as the DoubleU-Net proposed by Jha et al. [11], and the RefineU-Net introduced by Lin et al. [13], has further diversified the landscape of medical image segmentation techniques. These variants incorporate progressive global feedback mechanisms, residual attention guidance, and local refinement strategies, all of which contribute to heightened performance and accuracy of U-Net-based models across a spectrum of medical image segmentation tasks. Their contributions have significantly broadened the scope of U-Net and related architectural models, establishing them as indispensable components of the burgeoning field of medical image analysis. As this study embarks on the challenging task of polyp identification within colonoscopy images through semantic segmentation, it endeavors to provide an exhaustive comparative analysis, shedding light on the strengths and limitations of these architectural models in the context of precise polyp identification.

The challenge arises from the imprecise segmentation of polyps in Colonoscopy images due to the reasons outlined above. Consequently, this research aims to (1) achieve precise polyp segmentation in Colonoscopy images and (2) implement semantic segmentation techniques for Colonoscopy images. The primary objective of this study is to implement and analyze the application of the U-Net and SegNet model for semantic segmentation, specifically for identifying polyp appearances in Colonoscopy images with precision. To gain a comprehensive understanding of the research outcomes, quantitative results are presented in Figure 6, while qualitative results can be observed in Figure 8. This study contributes to the advancement of polyp detection and medical imaging through its comprehensive analysis and implementation of cutting-edge semantic segmentation techniques.

2. RESEARCH METHODOLOGY

2.1 Research Stages

In this study, a system is designed to segment polyps in Colonoscopy images. The system uses the U-Net architecture model as the proposed solution and SegNet as a comparison. In general, the design of the system to be built is as shown in Figure 1.

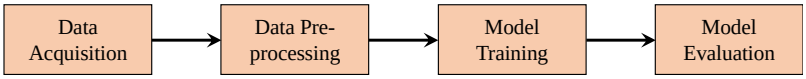


Figure 1. Overview of the research methodology.

Overall, there are four stages to the system, namely:

a. Data Acquisition, which is the process of obtaining Colonoscopy images obtained from the Kvasir-SEG dataset.
b. Data Pre-Processing is the process of processing image data so that it can be used in training the U-Net and SegNet model.
c. Model Training, which is the process of training the U-Net and SegNet model with previously processed image data.
d. Model Evaluation, which is the process of measuring the performance of the proposed system.

This research process begins with collecting Colonoscopy image data obtained from the Kvasir-SEG [2]. After the image data is obtained, the next step is to divide the dataset into three parts, namely training, validation, and test data. After the dataset is divided into three parts, data pre-processing will be carried out which aims to prepare the data for a better model training process. The training and validation stages of the model will be repeated continuously using training data and validation data until a certain number of repetitions. After the model has been trained, an evaluation will be carried out with previously prepared test data. The output of the entire process is labeled image data. For more details, please see Figure 2.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

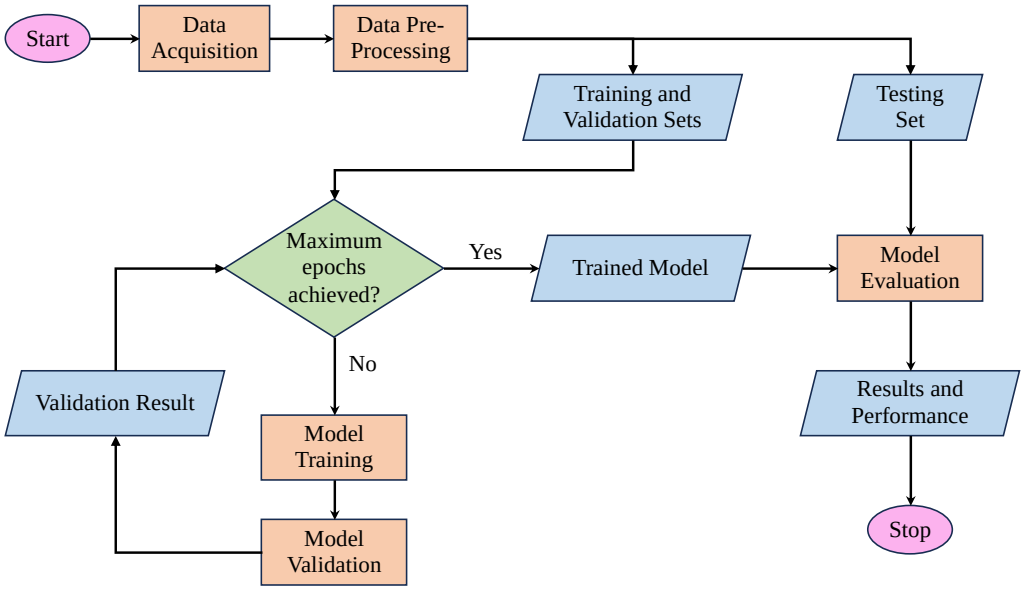


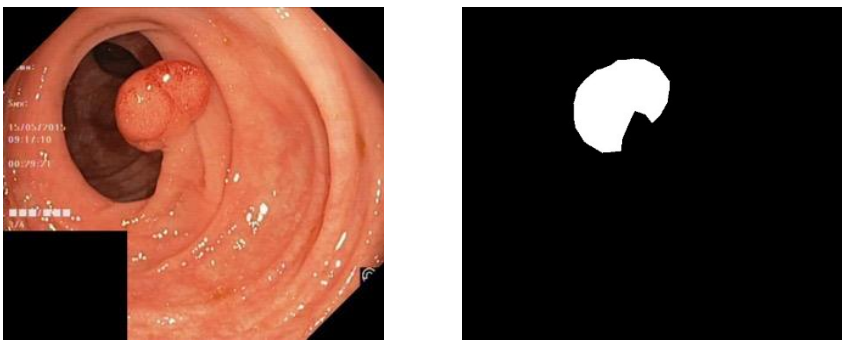
Figure 2. Flowchart diagram in the process of system modeling.

2.2 Dataset

The dataset used in this research is Colonoscopy image data available in the Kvasir-SEG dataset [2]. The image data provided by the Kvasir-SEG dataset is a three-channel (RGB) image as in Figure 3, which has a .jpg format and has two classes of objects to be classified, namely polyps and non-polyps.

The dataset utilized in this study is the Kvasir-SEG dataset. This dataset is based on the Kvasir dataset [14], which is a multi-class dataset designed for the detection and classification of gastrointestinal (GI) diseases. The Kvasir-SEG dataset was curated and validated by gastroenterology experts from Vestre Viken Health Trust in Norway. To address the high incidence of CRC, the "polyp" class was selected from the Kvasir dataset [14], resulting in the creation of the Kvasir-SEG dataset [2].

As depicted in Figure 3, pixels representing polyp tissue and regions of interest (ROI) are represented in white, while the background, devoid of positive pixels, is in black. The Kvasir-SEG dataset is structured into two folders: one for images and another for ground truth data. Each folder contains 1000 images. Image files from colonoscopy with corresponding ground truth data share the same naming conventions. Image files are encoded using JPEG compression, and researchers can easily access the dataset for research purposes at https://datasets.simula.no/kvasir-seg/ [2].



(a) Colonoscopy Image (b) Ground Truth

Figure 3. Sample of Kvasir-SEG dataset: Input image and its corresponding ground truth.

2.3 Data Augmentation

Obtaining and annotating medical data is challenging [2]. Training Deep Learning models on these datasets is challenging because most existing datasets have few samples. One possible way to overcome this challenge is to use data augmentation techniques that increase the number of instances during training. Therefore, we split the datasets into training, validation, and testing sets. Then we apply different data augmentation methods for each scenario. The methods include center crop, random crop, horizontal flip, vertical flip, scale augmentation, random rotation, cutout, and brightness augmentation.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

2.3 Model Architecture

2.4.1 U-Net

U-Net is a semantic segmentation technique originally proposed for medical image segmentation. U-Net was first introduced in a paper, U-Net: Convolutional Network for Biomedical Image Segmentation. The architectural model is quite simple: an encoder (for down sampling) and a decoder (for up sampling). As seen in Figure 4, it looks like the letter U which is why it is called U-Net.

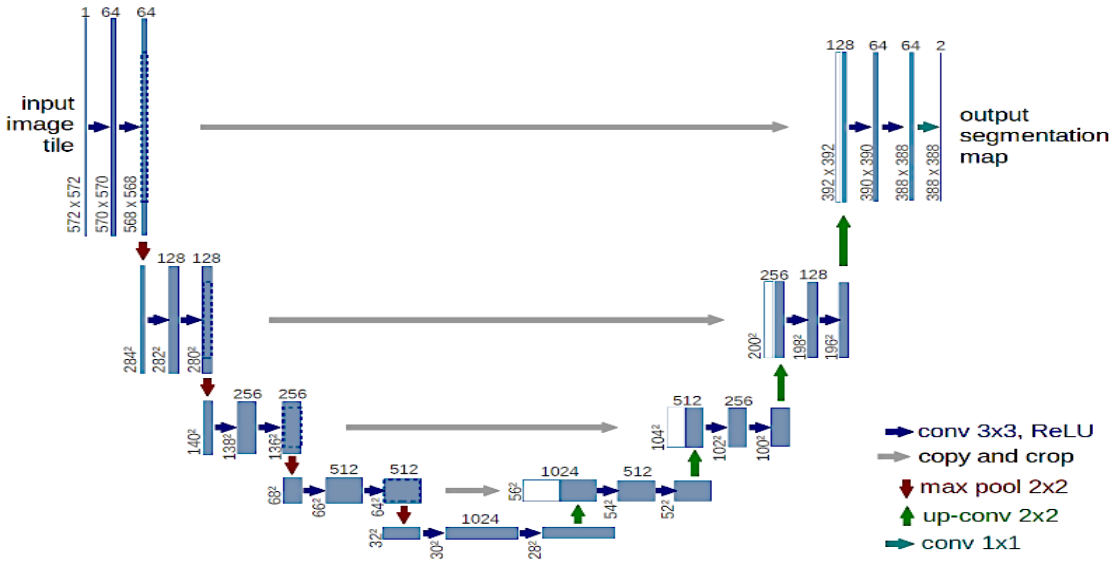


Figure 4. U-Net Architecture (captured from [10]).

In Figure 4, the U-Net architecture can be dissected into two distinctive components: the contraction path on the left and the expansion path on the right. The contraction path employs a sequence of two 3×3 unpadded convolutional layers, each followed by a Rectified Linear Unit (ReLU) activation function, and a 2×2 max-pooling layer with a stride of 2. During each downsampling step, the number of feature channels is doubled. Moving to the expansion path, an upsampling operation is followed by a 2×2 convolution layer, known as “up-convolution”, which reduces the number of feature channels by half. This path also involves feature integration from the contraction path and two additional 3×3 convolution layers, each accompanied by a ReLU activation function. Finally, a 1×1 convolutional layer is employed to map the 64 components of the feature vector to the desired class labels [10]. The U-Net architecture has found application in various domains, including endoscopic image segmentation for polyp detection in the digestive tract [11], [13], [15], [16], colon object delineation [17], laryngeal leukoplakia detection [18], and surgical instrument identification [19].

2.4.2 SegNet

SegNet is a deep fully convolutional neural network architecture. SegNet is designed to be an efficient architecture for semantic segmentation. SegNet was first introduced in the paper, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. This architecture model consists of an encoder and a decoder. The encoder in SegNet is topologically identical to the convolutional layer in VGG16. While the key component of SegNet is the decoder which consists of a hierarchy of decoders that corresponds to each encoder [12]. An overview of the SegNet architecture can be seen in Figure 5.

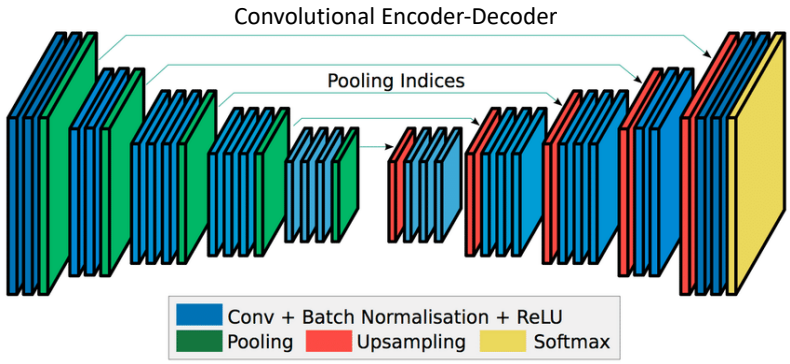


Figure 5. SegNet Architecture (captured from [12]).

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

2.5 Experiment Setup and Configuration

All model implementations were carried out utilizing the Keras Framework, with TensorFlow 2.9.0 serving as the backend. The experiments were executed on an Nvidia GTX 1650Ti GPU. The dataset partitioning involved allocating 80% for training, 10% for validation, and another 10% for testing purposes. The Dice coefficient function served as the uniform loss function across all networks, alongside the default parameters of the Nadam optimizer. Batch size configuration was set to 1, and the learning rate was established at 1e-4. Training iterations spanned 50 epochs, with the incorporation of ReduceLROnPlateau.

To evaluate the trained model, the Intersection over Union (IoU) metric is used. This metric calculates the similarity between the prediction result (A) and the ground truth (B) which corresponds to the equation below:

$$IoU(A,B) = \frac{A \cap B}{A \cup B} \quad (1)$$

$$Dice\ coefficient(A,B) = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (2)$$

2.6 Experimental Scenario

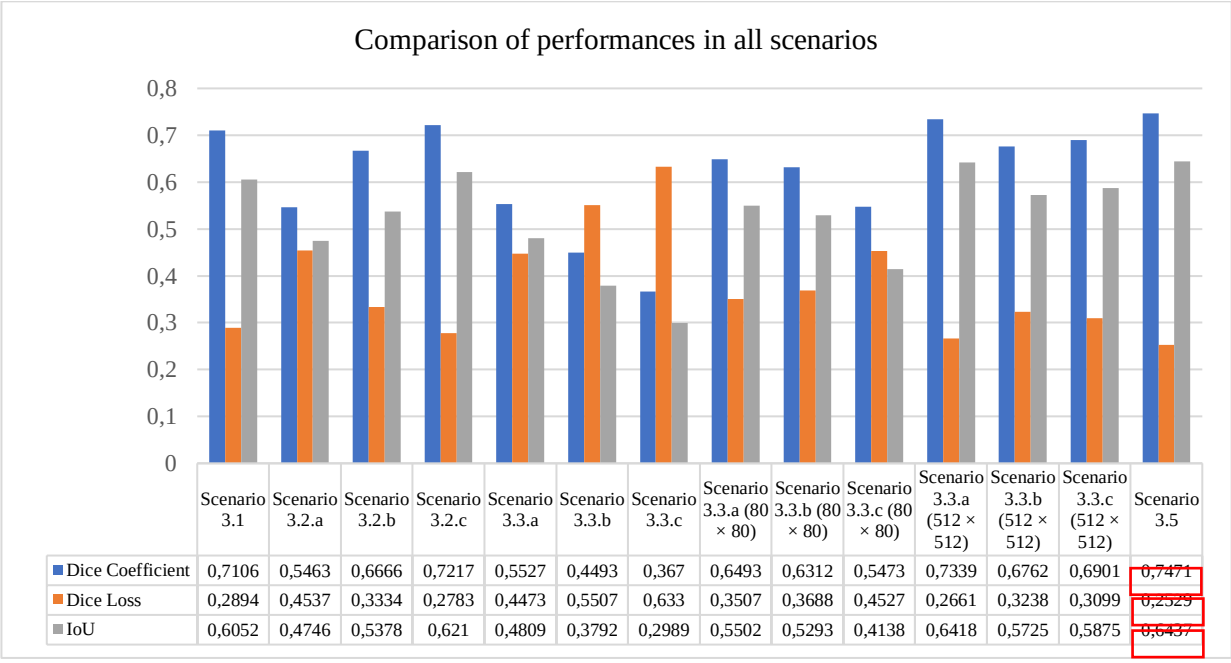
This research will perform several test scenarios. Generally, there will be six types of scenarios. Initially, we will train the model without applying data augmentation to assess its performance before incorporating data augmentation techniques. Second, perform data augmentation on the training data set only. Third, perform data augmentation on all data sets (training, validation, test). For the first, second, and third scenarios, the input size is 320×320 pixels, and the U-Net model is used. Fourth, we used an input size of 80×80. Fifth, we used an input size of 512×512. For the fourth and fifth scenarios, we used the best augmentation combination from the second and third scenarios to see if the input size affects the model performance. In conclusion, we will employ the optimal combination of augmentation technique and input size to train the SegNet model.

For the second, third, fourth, and fifth scenarios, each will have the following sub-scenarios:

a. Data augmentation in the form of random crops, horizontal flips, and random rotation.
b. All the data augmentations are mentioned in subsection 2.3.
c. Data augmentation in the form of (1) center crop; (2) 3 random crops with 60×60 crop size; (3) horizontal flip; (4) vertical flip; (5) scale augmentation with scale ranges of 512×768, 540×82, and 720×1024 with 60×60 crop size for each scale range; (6) 5 random rotations; (7) 3 cutouts; (8) gray scaling.

3. RESULT AND DISCUSSION

This section comprehensively presents and contrasts the training outcomes of the U-Net model across the diverse data augmentation scenarios outlined earlier. For an in-depth understanding of the quantitative results of each scenario, please refer to Figure 6, while Figure 8 provides a visual representation of the qualitative performance, facilitating a holistic evaluation of the model's performance.



	Scenario 3.1	Scenario 3.2.a	Scenario 3.2.b	Scenario 3.2.c	Scenario 3.3.a	Scenario 3.3.b	Scenario 3.3.c	Scenario 3.3.a (80 × 80)	Scenario 3.3.b (80 × 80)	Scenario 3.3.c (80 × 80)	Scenario 3.3.a (512 × 512)	Scenario 3.3.b (512 × 512)	Scenario 3.3.c (512 × 512)	Scenario 3.5
Dice Coefficient	0,7106	0,5463	0,6666	0,7217	0,5527	0,4493	0,367	0,6493	0,6312	0,5473	0,7339	0,6762	0,6901	0,7471
Dice Loss	0,2894	0,4537	0,3334	0,2783	0,4473	0,5507	0,633	0,3507	0,3688	0,4527	0,2661	0,3238	0,3099	0,2529
IoU	0,6052	0,4746	0,5378	0,621	0,4809	0,3792	0,2989	0,5502	0,5293	0,4138	0,6418	0,5725	0,5875	0,6437

Figure 6. Comparison of performances in all scenarios, measured in Dice Coefficient, Dice loss, and IoU metrics. For the Dice Coefficient and IoU metrics, the higher the better; Meanwhile, for the Dice Loss metric, the lower the better. Red boxes show the best scores in each performance metric row.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

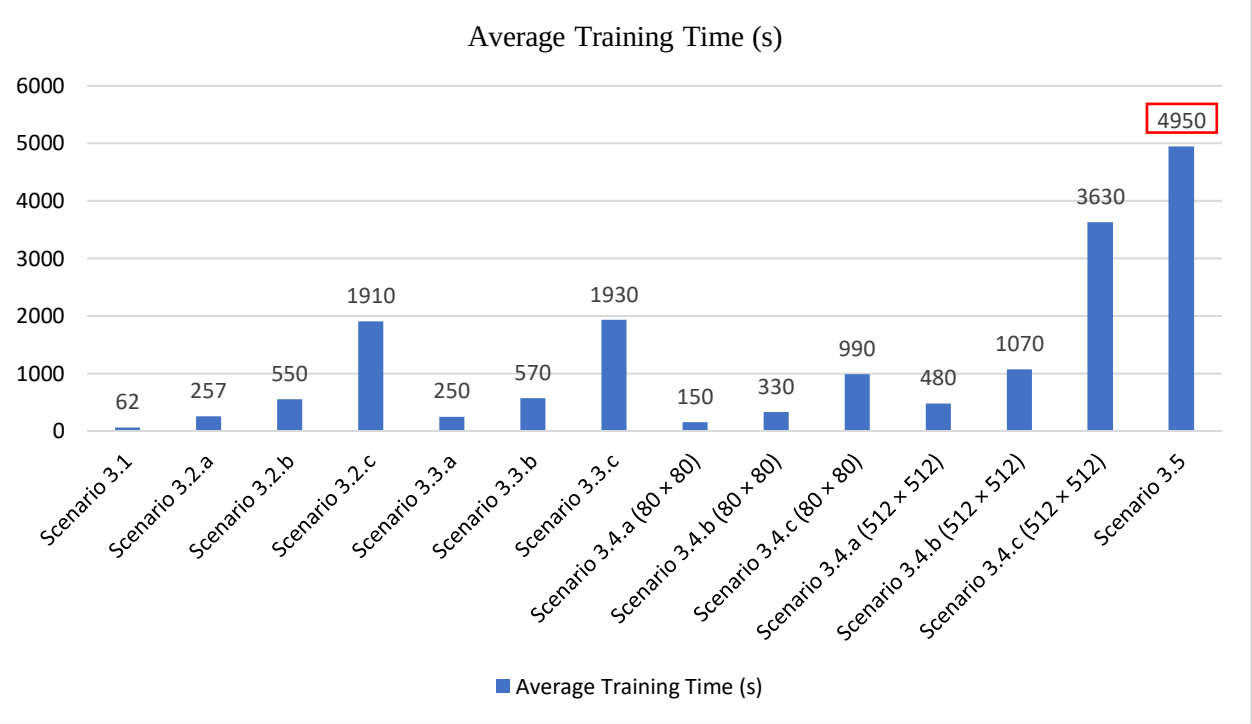


Figure 7. Average training time in all scenarios. The red box shows the longest time required for training phase in the corresponding scenario.

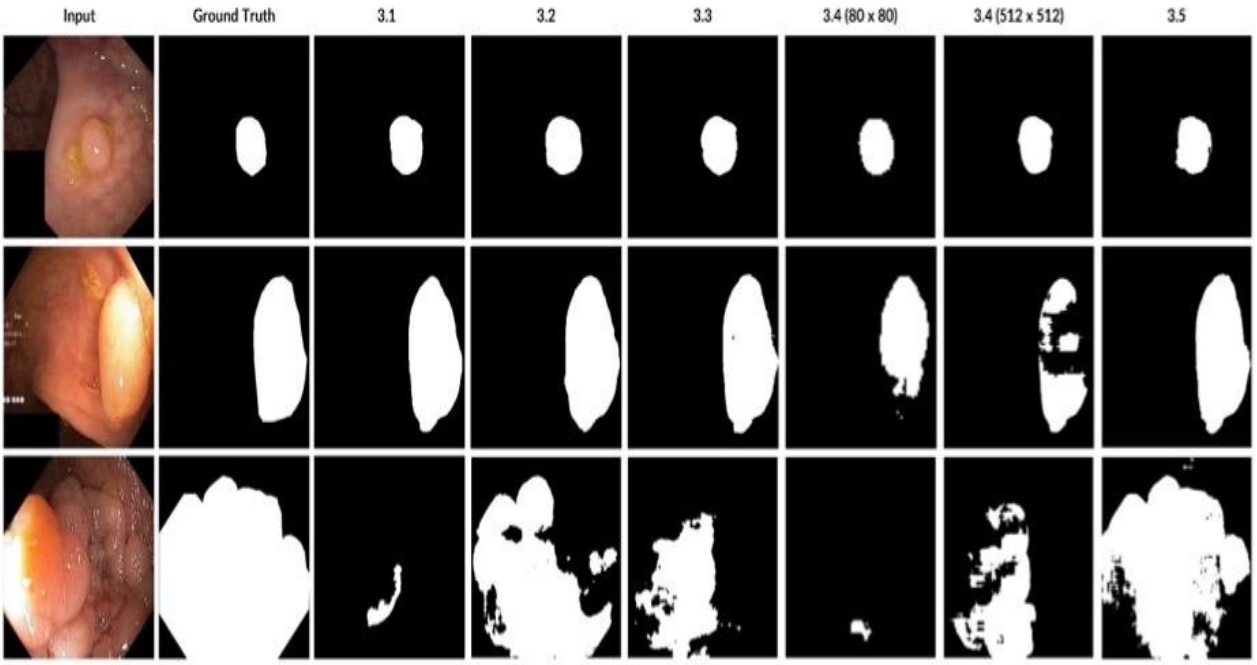


Figure 8. Qualitative results by the models in all scenarios.

3.1 Comparison on Original Kvasir-SEG Dataset on U-Net

In the upcoming experiment, the U-Net model will undergo training using non-augmented data, employing an input size of 320×320. The progression of the model's learning can be observed in Figure 9, illustrating that its peak performance is attained by the 49th epoch and remains stable thereafter. Consequently, we have extended the training duration to 50 epochs for our subsequent experimentation. The specific training results are itemized in Figure 6.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

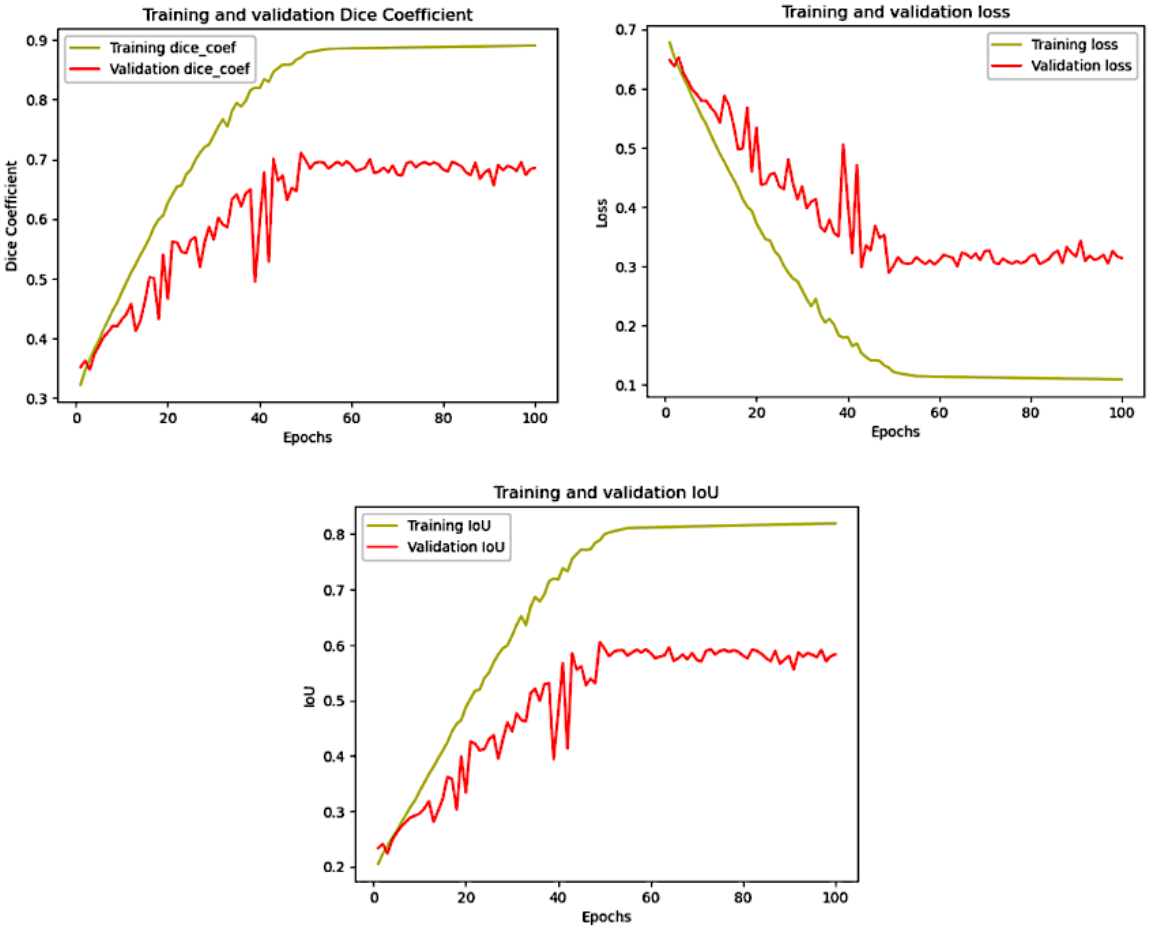


Figure 9. Depicting training progress in every epoch, measured in Dice Coefficient, Dice loss, and IoU metrics.

3.2. Comparison on Augmented Kvasir-SEG Train Dataset on U-Net

The quantitative outcomes pertaining to our second scenario, involving the application of data augmentation exclusively to the training dataset, have been consolidated and presented in Figure 6. Our findings reveal a marginal enhancement in performance within sub-scenario 2.6.c when contrasted with non-augmented data. Nevertheless, from a qualitative perspective, it is evident that sub-scenario 2.6.c exhibits superior segmentation capabilities, particularly for multiple polyp objects within a single image, as visually depicted in Figure 8 (3.1 and 3.2).

3.3. Comparison on Augmented Kvasir-SEG Dataset on U-Net

Figure 8 (3.3) provides an overview of the qualitative outcomes concerning the U-Net model's performance on the Augmented Kvasir-SEG Dataset. In this particular scenario, sub-scenario 2.6.a yielded a Dice coefficient of 0.5527, signifying a 27.5% reduction in performance compared to scenario 3.2. Furthermore, the subsequent sub-scenarios within this scenario exhibit a consistent decline in performance. Hence, our decision is to adopt the data augmentation technique featured in scenario 3.2 due to its superior performance characteristics.

3.4. Comparison on Augmented Kvasir-SEG Train Dataset on U-Net with Different Input Size

We have conducted experiments utilizing input sizes of 80×80 and 512×512 to investigate potential performance variations associated with different input dimensions. Figure 6 illustrates that the adoption of an 80×80 input size results in inferior performance across all sub-scenarios when compared to the 320×320 input size. Additionally, the qualitative assessment, as depicted in Figure 8 (3.4 (80×80)), corroborates that smaller input sizes lead to suboptimal performance.

Furthermore, we have documented both qualitative and quantitative outcomes pertaining to the 512×512 input size scenario, as presented in Figure 6 and Figure 8 (3.4 (512×512)). These results underscore that this particular scenario achieves the most favorable average quantitative performance among the three sub-scenarios. However, it is noteworthy that upon evaluating the qualitative results for sub-scenario 2.6.a, which showcases the highest quantitative performance within this scenario, there is a discernible disparity compared to the qualitative outcomes observed in scenario 3.2.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

3.5. Comparison on Augmented Kvasir-SEG Train Dataset on SegNet

The aim of this experimental endeavor was to meticulously assess the performance of the U-Net model. This evaluation entailed the utilization of the most effective augmented data from the preceding five scenarios and entailed a comparative analysis against the SegNet model. The quantitative analysis, meticulously presented in Figure 6, exhibited a significant enhancement, showcasing a noteworthy 3.5% improvement in the dice coefficient in direct comparison with scenario 3.3 sub-scenario 2.6.c. Moreover, the detailed examination, as illustrated in Figure 8 (3.5), provided valuable insights into the model's capabilities, revealing discernible advancements in results, particularly when addressing images featuring multiple polyps.

3.6 Discussion

Figure 9 in the study elucidates the training progression of the U-Net model, which was trained on the Kvasir-SEG dataset without the application of augmentation techniques. Convergence was proficiently attained at the 49th epoch, subsequently warranting the utilization of 50 epochs for subsequent experiments. Within scenario 3.2, Figure 6 showcases a commendable level of competitiveness between sub-scenario c and scenario 3.1. Nevertheless, a meticulous comparison of the qualitative results, as depicted in Figures 6 (3.1) and (3.2), distinctly highlights the U-Net model's exceptional performance when data augmentation is applied during training. Consequently, an endeavor was made to extend the application of augmentation to all datasets, encompassing training, validation, and testing sets. As anticipated, this extension resulted in suboptimal performance in contrast to scenario 3.1 (Figure 6). It is noteworthy that the judicious application of augmentation exclusively to the training set ensures an unbiased evaluation on the validation and test sets.

Subsequently, in the ensuing scenario, the U-Net model was subject to evaluation employing augmentation techniques from scenario 3.2, while considering disparate input sizes (80×80 and 512×512). The adoption of smaller input sizes (80×80) yielded relatively inferior performance (Figure 6) attributable to the curtailed spatial information, notably affecting the model's capacity to capture intricate details, particularly in the case of diminutive polyp structures. Conversely, scenario 3.4 (512×512) unveiled more favorable average quantitative results in comparison to scenario 3.2. However, it is imperative to acknowledge that larger input sizes introduce heightened computational complexity and elevated memory requisites. Moreover, a meticulous assessment of the qualitative results, as discerned in Figure 8 (3.4 (512×512)), reaffirmed the supremacy of scenario 3.2's performance.

The final scenario heralded a 3.5% enhancement in quantitative results in relation to the best-performing scenario among the preceding five (scenario 3.2), concurrently yielding slightly superior qualitative results. Nonetheless, it is paramount to underscore that these advancements were achieved at the expense of a near fivefold augmentation in training duration (see Figure 7). In pursuit of comprehensive clarity, the qualitative outcomes of all scenarios are thoughtfully juxtaposed within Figure 8.

4. CONCLUSION

In summary, our comprehensive investigation into augmentation strategies for polyp semantic segmentation has yielded valuable insights and promising outcomes. Particularly, Scenario 3.2, involving the augmentation of the training dataset with diverse techniques, such as (1) Center crop; (2) 3 random crops with 60×60 crop size; (3) horizontal flip; (4) vertical flip; (5) scale augmentation with scale ranges of 512×768, 540×82, and 720×1024 with 60×60 crop size for each scale range; (6) 5 random rotations; (7) 3 cutouts; and (8) gray scaling, has exhibited noteworthy performance metrics with a dice coefficient of 0.7217, IoU of 0.6210, and dice loss of 0.2783. Moreover, Scenario 3.2 employing an input size of 320×320 has emerged as a standout performer, striking an admirable equilibrium between accuracy and computational efficiency. While the potential advantages of larger input dimensions have been acknowledged, prudent consideration of resource implications remains paramount. Looking ahead, we are enthusiastic about our forthcoming contributions to the progression of medical image segmentation, aimed at the development of a unified model proficient in capturing intricate details across diverse medical images. This research serves as a pivotal foundation, paving the way for advanced and potent applications in the realm of polyp detection and medical imaging.

REFERENCES

[1] R. L. Siegel et al., "Colorectal cancer statistics, 2020," CA Cancer J Clin, vol. 70, no. 3, pp. 145–164, May 2020, doi: 10.3322/caac.21601.
[2] D. Jha et al., "Kvasir-SEG: A Segmented Polyp Dataset," 2020, pp. 451–462. doi: 10.1007/978-3-030-37734-2_37.
[3] D. A. Lieberman et al., "Use of Colonoscopy to Screen Asymptomatic Adults for Colorectal Cancer," New England Journal of Medicine, vol. 343, no. 3, pp. 162–168, Jul. 2000, doi: 10.1056/NEJM200007203430301.
[4] P. F. Pinsky and R. E. Schoen, "Contribution of Surveillance Colonoscopy to Colorectal Cancer Prevention," Clinical Gastroenterology and Hepatology, vol. 18, no. 13, pp. 2937-2944.e1, Dec. 2020, doi: 10.1016/j.cgh.2020.01.037.
[5] D.-P. Fan et al., "PraNet: Parallel Reverse Attention Network for Polyp Segmentation," 2020, pp. 263–273. doi: 10.1007/978-3-030-59725-2_26.

Building of Informatics, Technology and Science (BITS)
Volume 5, No 2, September 2023 Page: 423−431
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v5i2.4083

[6] S. N. Hong et al., "The Effect of the Bowel Preparation Status on the Risk of Missing Polyp and Adenoma during Screening Colonoscopy: A Tandem Colonoscopic Study," Clin Endosc, vol. 45, no. 4, p. 404, 2012, doi: 10.5946/ce.2012.45.4.404.

[7] A. Leufkens, M. van Oijen, F. Vleggaar, and P. Siersema, "Factors influencing the miss rate of polyps in a back-to-back colonoscopy study," Endoscopy, vol. 44, no. 05, pp. 470–475, May 2012, doi: 10.1055/s-0031-1291666.

[8] S. B. Ahn, D. S. Han, J. H. Bae, T. J. Byun, J. P. Kim, and C. S. Eun, "The Miss Rate for Colorectal Adenoma Determined by Quality-Adjusted, Back-to-Back Colonoscopies," Gut Liver, vol. 6, no. 1, pp. 64–70, Jan. 2012, doi: 10.5009/gnl.2012.6.1.64.

[9] S. Ameling, S. Wirth, D. Paulus, G. Lacey, and F. Vilarino, "Texture-Based Polyp Detection in Colonoscopy," 2009, pp. 346–350. doi: 10.1007/978-3-540-93860-6_70.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[11] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, "DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation," in 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), IEEE, Jul. 2020, pp. 558–564. doi: 10.1109/CBMS49503.2020.00111.

[12] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Trans Pattern Anal Mach Intell, vol. 39, no. 12, pp. 2481–2495, Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.

[13] D. Lin, Y. Li, T. L. Nwe, S. Dong, and Z. M. Oo, "RefineU-Net: Improved U-Net with progressive global feedbacks and residual attention guided local refinement for medical image segmentation," Pattern Recognit Lett, vol. 138, pp. 267–275, Oct. 2020, doi: 10.1016/j.patrec.2020.07.013.

[14] K. Pogorelov et al., "Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection," in Proceedings of the 8th ACM on Multimedia Systems Conference, 2017, pp. 164–169.

[15] B. Murugesan, K. Sarveswaran, S. M. Shankaranarayana, K. Ram, J. Joseph, and M. Sivaprakasam, "Psi-Net: Shape and boundary aware joint multi-task deep network for medical image segmentation," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Jul. 2019, pp. 7223–7226. doi: 10.1109/EMBC.2019.8857339.

[16] H. Toda et al., "Shape Recovery of Polyp Using Blood Vessel Detection and Matching Estimation by U-Net," in 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI), IEEE, Jul. 2019, pp. 450–453. doi: 10.1109/IIAI-AAI.2019.00098.

[17] N. Ibtehaz and M. S. Rahman, "MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation," Neural Networks, vol. 121, pp. 74–87, Jan. 2020, doi: 10.1016/j.neunet.2019.08.025.

[18] B. Ji et al., "A multi-scale recurrent fully convolution neural network for laryngeal leukoplakia segmentation," Biomed Signal Process Control, vol. 59, p. 101913, May 2020, doi: 10.1016/j.bspc.2020.101913.

[19] S. M. Kamrul Hasan and C. A. Linte, "U-NetPlus: A Modified Encoder-Decoder U-Net Architecture for Semantic and Instance Segmentation of Surgical Instruments from Laparoscopic Images," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Jul. 2019, pp. 7205–7211. doi: 10.1109/EMBC.2019.8856791.