Combination of VGG19 (Encoder) and U-Net (Decoder) for Colorectal Polyp Segmentation Image

Nuri Sutiyaningsih[1*], Putu Desiana Wulaning Ayu[2], Roy Rudolf Huizen[3]

[1]Magister Program, Department of Magister Information System, Institut Teknologi dan Bisnis STIKOM Bali, Indonesia

[2,3]Department of Magister Information System, Institut Teknologi dan Bisnis STIKOM Bali, Indonesia

*corr-author: 222012015@stikom-bali.ac.id

Abstract - Health involves the proper function of the body and organs, with colon polyps being a common issue. Doctors often face challenges in segmenting medical images, especially endoscopic images for polyp detection. The complexity and variation in the appearance of polyps make accurate identification challenging, and the subjective manual segmentation process can result in misdiagnosis or delayed treatment. This study examines the effectiveness of the combination of U-Net decoder model architecture and VGG19 encoder in segmentation of colon polyp images. This study uses a public dataset, namely Kvasir-Seg with a total of 1000 images of colon polyps. An innovative approach using VGG19 as encoder and U-Net as decoder improves colorectal polyp segmentation, achieving high performance with a Loss of 0.05, Accuracy 0.95, Precision 0.96, Recall 0.92, IoU 0.89, and Dice 0.94. Using optimal parameters such as Nadam Optimizer, 5 Fold Cross Validation, Learning Rate 0.0001, and 25 Epochs significantly improved performance, increasing the Dice Coefficient to 0.92 and IoU to 0.86 compared to previous studies. This study concludes that the proposed architecture is reliable for colon polyp segmentation. Future work should explore attention mechanisms or transformer-based models to enhance accuracy and efficiency.

Keywords: Colorectal polyp; segmentation image; U-Net; VGG19.

I. INTRODUCTION

Health is a holistic condition involving body and organ functions, influenced by lifestyle, environment, and access to healthcare [1]. Colon health, including its major parts, is vital for body balance by absorbing water and electrolytes. Early detection of adenomatous polyps helps reduce cancer risk, making knowledge of colon structure and function important for digestive health [2].

Colon health is vital for digestive well-being, with colon polyps as abnormal growths on the intestinal wall being a common issue [3] Although many polyps are benign, some can progress to colorectal cancer if untreated. Regular screening is crucial to prevent serious complications [4] risk factors for polyp development include a low-fiber diet, hereditary predisposition to colon cancer, and genetic disorders such as Lynch syndrome and familial adenomatous polyposis. Routine colonoscopy exams enable early detection, lowering colorectal cancer risk [5].

Doctors often struggle to accurately segment colon polyp images due to their complexity and variation, and manual methods risk misdiagnosis or delays. Thus, faster and more accurate automatic segmentation techniques are essential. [6,7]. Despite advances in technology and deep learning boosting diagnostic speed and accuracy, challenges persist in matching traditional methods' effectiveness, as well as in clinical implementation and integration into routine practice [7]. Conventional methods often struggle to accurately detect polyps, especially under poor lighting or with irregular shapes. Variations in polyp size, shape, texture, and image noise complicate segmentation. Thus, a more adaptive and robust approach is needed to enhance accuracy and reliability.

U-Net excels in polyp segmentation because it can recognize objects of different sizes and shapes and efficiently retains and combines spatial information through its symmetrical structure [8]. U-Net is widely applied not only in biomedical image segmentation but also in natural image segmentation. Typically, the encoder consists of several convolutional layers designed to extract and represent semantic features from low-level to high-level abstractions. The encoder can be replaced by pre-trained networks, such as VGG16, VGG19. VGG-19 extracts features using regular convolution and VGG net as the basic building blocks of the 19 layers [9].

Segmentation in microscopic imagery isolates important structures or objects, enabling focused analysis of areas like cells or organelles. Using deep learning to recognize patterns, it improves measurement of size, shape, and distribution for more effective study

[10]. Manual segmentation, especially in identifying Regions of Interest (ROI) in medical imaging, the process is often time-consuming and susceptible to bias. Therefore, establishing automated and standardized approaches is essential to enhance both accuracy and efficiency. particularly for clinical populations with diverse health conditions [11]. In the context of medical diffusion imaging, the ROI is defined as the localized region utilized for the analysis of network-level attributes. Given that manual ROI segmentation is time-intensive and vulnerable to bias, there is a growing need for automated, standardized solutions that can ensure accuracy and efficiency in studies involving patients with diverse medical profiles [12].

Rajak and Mirza's research, "Segmentation of Polyp Instruments using a U-Net Based Deep Learning Model," employed the U-Net framework to identify and segment polyps and endoscopic instruments from the KVASIR dataset. The study prioritizes high segmentation metrics but does not investigate alternative architectures like attention-based models. Their U-Net achieves a Dice Coefficient of 0.91, IoU of 0.85, and accuracy of 0.95, showing strong segmentation performance [13].

The study conducted by M. Yeung, E. Sala, C. B. Schönlieb, and L. Rundo, titled "Focus U-Net: A Novel Dual Attention-Gated CNN for Polyp Segmentation During Colonoscopy", presents a new model known as Focus U-Net which integrates spatial and channel attention mechanisms for selective learning of polyp features, along with architectural improvements. The model demonstrates strong performance with a Dice Coefficient of 0.93 and IoU of 0.87 but may struggle with difficult polyps in poor image conditions and could face generalization challenges when trained on limited datasets [14].

In the study "A Comprehensive Study on Colorectal Polyp Segmentation with ResU-Net++, Conditional Random Field and Test-Time Augmentation," D. Jha, P. Helén Smedsrud, D. Johansen, T. de Lange. introduced ResU-Net++, built on ResU-Net, enhanced with Conditional Random Field and Test-Time Augmentation to boost segmentation accuracy. While the model performed well, the study lacked detailed analysis of metrics like sensitivity and specificity for a fuller evaluation. The model achieved a Dice Coefficient of 0.85 and an IoU of 0.83 on training data [15].

Previous research by Rishav Kumar Rajak and Ashar Beg Mirza using U-Net for polyp segmentation achieved a Dice Coefficient of 0.91, IoU of 0.85, and accuracy of 0.95 but faced challenges with small polyps and noise sensitivity. This study improves performance by combining VGG19 as an encoder with U-Net as a decoder, leveraging VGG19's deep feature extraction and U-Net's precise segmentation. While methods like Focus U-Net and ResUNet++ perform well, Focus U-Net struggles with high image noise, and ResUNet++ may degrade on larger or diverse datasets. The VGG19-U-Net combination synergizes strong feature extraction and segmentation, overcoming these limitations for better accuracy and reliability.

In medical imaging, especially for colon polyp endoscopy, accurate segmentation is challenging due to the complex and varied appearance of polyps. Effective segmentation is crucial to understand the structure and function of these lesions. This research aims to develop a more efficient and accurate automatic segmentation method for polyp detection, addressing the following problem formulation:

1) *Research Problem 1:* What are the results of the analysis of the architectural effectiveness of the U-Net decoder model and VGG19 encoder in the segmentation of polyp images in the colon?

2) *Research Problem 2:* What are the results of the analysis of the influence of optimizer and evaluation metrics on the results of polyp segmentation and polyp disease segmentation performance in the U-Net decoder and VGG19 encoder models?

This study presents an evaluation of a hybrid deep learning architecture that integrates VGG19 as the encoding network and U Net as the decoding component to enhance the segmentation precision of colon polyps. The experimental design involves assessing several optimization algorithms and systematically tuning hyperparameters such as learning rate, batch size, and epoch configuration to achieve optimal performance. Furthermore, regularization methods including dropout and batch normalization are applied to mitigate overfitting and strengthen the model's generalization capability. By combining VGG19's feature extraction capacity with U Net's structural segmentation accuracy, the proposed approach seeks to provide a more reliable and efficient framework for computer aided medical diagnosis.

II. METHOD

The proposed method employs a deep learning model to perform colon polyp segmentation. The workflow includes data collection from Kaggle, preprocessing, and data splitting using K Fold Cross Validation. The optimized U Net architecture is subsequently trained and evaluated.

A. Dataset

The dataset used in this study is Kvasir-SEG, consisting of 1,000 endoscopic images and 1,000 corresponding mask images of colon polyps. This balanced dataset supports the development and analysis of polyp detection and segmentation models, enabling deeper research in colon polyp diagnosis and treatment [16]. An example of the endoscopic polyp images employed in this study is presented in Fig. 1.

B. Data Preprocessing

Data preprocessing for polyp image segmentation involves standardizing image dimensions and removing irrelevant details to optimize model training. Images are resized to 256 by 256 pixels for consistency and reduced processing complexity. Cropping eliminates unnecessary elements such as text or patient information before further processing [17] Dropout is a regularization method that prevents overfitting by randomly deactivating neurons during training. In this study, dropout layers are integrated within the network architecture to improve generalization and reduce model dependency on specific neurons, ensuring more stable learning performance.a dropout rate of 0.5 was used, meaning half of the neurons in a layer are turned off each iteration. This forces the model to distribute learning across neurons. An illustration of the preprocessing pipeline utilized in this study is presented in Fig. 2.

C. Data Split (K-Fold)

The dataset is partitioned into training and validation subsets to minimize overfitting and enable an objective performance evaluation. K-Fold cross-validation helps the model perform well across various data subsets and generalize to new data. This method is especially useful for small datasets, maximizing data utilization for training and testing. [18] By averaging results across several folds, cross-validation minimizes variance and yields a more stable performance assessment. Dividing data into K folds prevents the model from training on just one subset, helping avoid overfitting and improving generalization. In this study, experiments use K=3 and K=5, meaning the dataset is split into 3 or 5 parts respectively.

D. Model Encoder

In the VGG19 structure, there is an additional layer of convolution that allows the model to extract features with a higher level of complexity [19]. Fig. 3 shows the VGG19 model used as the encoder in this study.

VGG19 has 19 layers: 16 convolutional and 3 fully connected. It takes a 224×224 RGB image as input. Early layers include two convolutional layers with 64 filters and 3×3 kernels, followed by 2×2 max pooling. Next, two convolutional layers with 128 filters and 3×3 kernels are applied, followed by another max pooling. Convolutional layers with ReLU activation extract complex features at multiple abstraction levels, essential for detailed image analysis [20] VGG19's many convolutional layers enable the model to extract complex features from images, which is crucial for image segmentation tasks that require recognizing fine details and intricate patterns.

E. Model Decoder

In the U-net structure starts with the input image, with a predetermined size of 256×256×3 for the RGB image. Fig. 4 shows the U-Net model used as a decoder in this study.

Fig. 1 Polyp endoscopy dataset

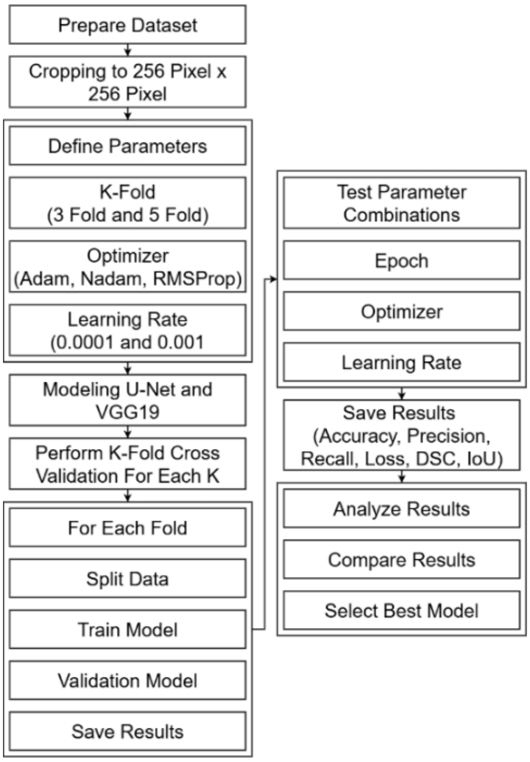


Fig. 2 Pipeline preprocessing

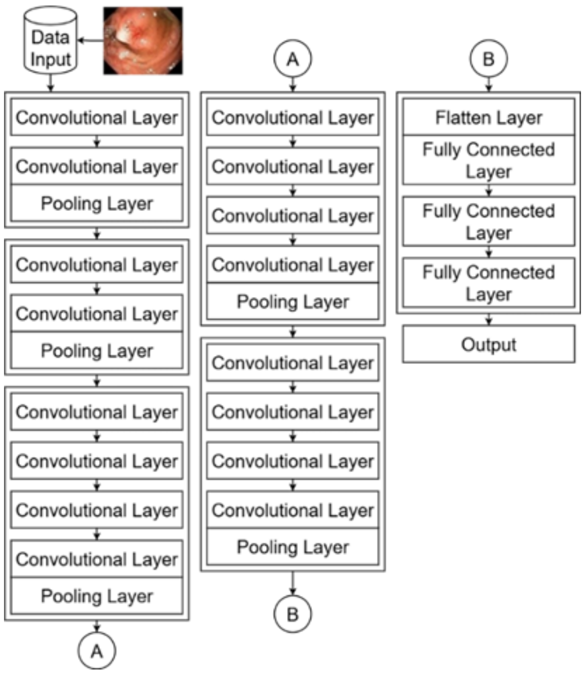


Fig. 3 Model VGG19

The segmentation process begins with an input image of size 256×256×3 (RGB). Each encoding block contains two convolutional layers using 3×3 filters with padding to maintain spatial dimensions, and uses the ReLU activation function. The feature maps are then downsampled using 2×2 max pooling, reducing their dimensions by half at each level. The bottleneck section connects the encoder and decoder paths with two convolutional layers and ReLU activations. The decoder path mirrors the encoder structure, using upsampling followed by convolutional layers to reconstruct the spatial resolution.At the final stage, a 1×1 convolution with a sigmoid activation function generates a binary mask, segmenting the colon polyp regions from the endoscopic image [21]. U-Net with VGG19 as the encoder enhances segmentation by extracting deeper features, improving robustness to lighting and noise. Its symmetrical architecture captures multi-scale information for accurate detection of varied polyp shapes. VGG19's convolutional depth strengthens feature learning, while U-Net ensures precise segmentation. The model uses a batch size of 8, runs for 25 epochs, and applies a learning rate of 0.0001 for stable training. Binary cross-entropy loss evaluates prediction accuracy and directs the learning process effectively.

F. Evaluation Optimizer

1) Adam Optimizer: Adam learning rate optimization is a technique in machine learning that adaptively adjusts the learning rate for each model parameter. It does this by estimating the first moment (mean) and second moment (uncentered variance) of the gradients, enabling efficient and effective training [22].

2) Nadam Optimizer: Nadam is a modified version of the Adam optimizer that incorporates Nesterov Accelerated Gradient (NAG). It improves Adam by replacing the traditional momentum component with the current Nesterov momentum vector, enabling faster and more stable convergence during training [23] Nadam is an optimization algorithm that combines Nesterov accelerated gradient (NAG) with Adam. It integrates Nesterov momentum into Adam's adaptive learning process, enhancing parameter updates for better training performance [24].

RMSProp: RMSProp is a widely used stochastic optimization algorithm in deep learning. However, recent studies indicate that it may fail to converge to the optimal solution, even in simple convex problem [25] RMSProp modifies Adagrad by accumulating gradients using an exponential weighted average, which allows it to focus on the most recent gradient information while ignoring older history [26].

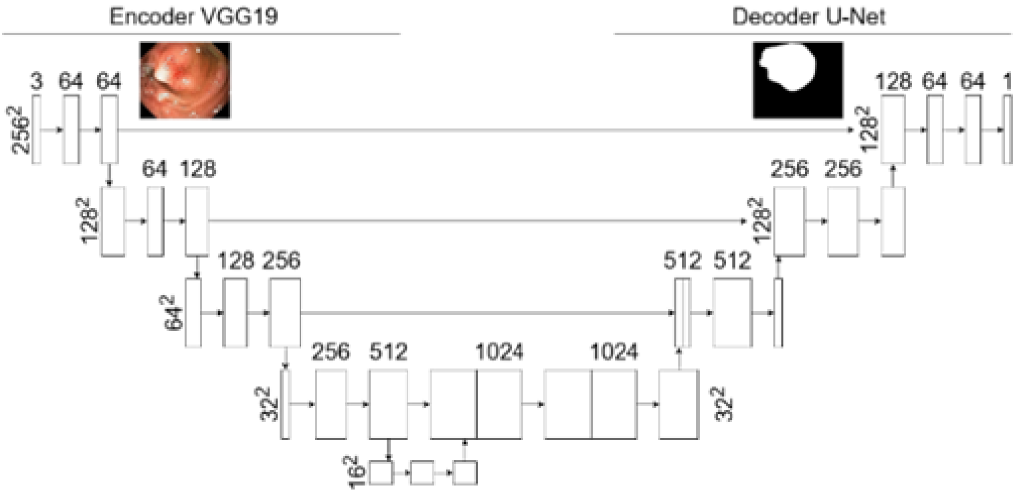


Fig. 4 Model U-Net

G. Matrix Evaluation

1) Dice Coefficient: Dice Coefficient measures similarity between two data sets, commonly used to evaluate image segmentation. It calculates the overlap between the predicted and actual areas relative to their total size, indicating segmentation accuracy. The formula for calculating the Dice Coefficient is presented in (1).

$$\frac{2 \times TP}{2 \times TP + FP + FN} \quad (1)$$

The value of the Dice Coefficient ranges between 0 and 1, where 1 indicates perfect equality and 0 indicates no equality. This metric is particularly useful in image segmentation tasks because it is more sensitive to small areas that are mispredicted, thus providing a more accurate evaluation of the quality of segmentation. Dice Coefficients are often used in the medical field, such as organ segmentation in MRI images or CT scans, to ensure that the model's predictions are as accurate as possible with the actual condition.

2) Intersection over Union (IoU): Intersection over Union (IoU), also called the Jaccard Index, is a metric to evaluate segmentation or object detection accuracy. It measures the overlap between the predicted area and the ground truth by dividing their intersection by their union. The formula for calculating IoU is presented in (2).

$$\frac{TP}{TP + FP + FN} \quad (2)$$

The IoU value ranges from 0 to 1, where 1 means perfect overlap between prediction and ground truth, and 0 means no overlap. It is crucial in image segmentation and object detection as it measures how accurately the model predicts the object's location and shape. A higher IoU indicates better model performance [27].

3) Recall: Also known as sensitivity or true positive rate, it measures how many of the total positive cases a model has successfully identified. This is important when the cost of false negatives is high. The formula for calculating Recall is presented in (3).

$$\frac{TP}{TP + FN} \quad (3)$$

4) Precision: Measures how much of the positive prediction is true-true positive. This is important when the cost of false positives is high. The formula for calculating Recall is presented in (4).

$$\frac{TP}{TP + FP} \quad (4)$$

5) Accuracy: It measures how much of the total predictions are true, both positive and negative. This is the most commonly used metric, but it can be misleading if the dataset is unbalanced. The formula for calculating Recall is presented in (5).

$$\frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

TP true positive (positive label predicted as the true label), false positive FP (negative label but predicted as positive label), TN true negative (negative data predicted true), and negative FN false (positive label but predicted as negative label) [28].

III. RESULT AND DISCUSSION

The U Net model with a VGG19 encoder in this study was implemented by adjusting several hyperparameters, including the number of epochs, learning rate, loss function, and optimizer. This hyperparameter adjustment aimed to determine the most effective configuration for colon polyp segmentation. To ensure the reliability and stability of model performance, K Fold cross validation with 3 and 5 folds was applied during the training and validation stages with the number of training epochs ranging from 1 to 25, and optimizers including Adam, Nadam, and RMSProp. Learning rates were also compared, with a dropout rate set at 0.5 to randomly deactivate half of the neurons during training, promoting better generalization. K-Fold Cross-Validation divides the data into K parts, training the model K times with one fold for validation each time. Using 3 folds balances training time and validation, while 5 folds offer more robust evaluation and reduce variance in performance estimates.The advantage of using k fold cross validation is to reduce the risk of overfitting by ensuring the model is tested on invisible data. Provides more stable and reliable performance estimates compared to single train-test split. Compares U-Net performance using 3-Fold and 5-Fold cross-validation for polyp segmentation. The 5-Fold model performs better, with lower Loss at 0.23 and higher Accuracy at 0.90, along with improved Precision (0.74), Recall (0.72), IoU (0.60), and Dice Coefficient (0.75). This shows 5-Fold training improves generalization and segmentation accuracy. Epochs represent full training cycles using the entire dataset to update weights. Testing from 1 to 25 epochs helps identify when the model starts overfitting and find the optimal epoch count for best performance without excessive training. This monitoring provides insights into training progress and convergence. That as epochs increase, U-Net's performance improves significantly. At Epoch 5, Loss is 0.23 and Accuracy is 0.90. By Epoch 15, Loss drops to 0.09 and Accuracy rises to 0.94. At Epoch 25, Loss reaches 0.05 and Accuracy 0.95. Precision improves from 0.77 to 0.96, Recall from 0.76 to 0.92, while IoU and Dice coefficients also increase, reaching 0.89 and 0.94 respectively, demonstrating better segmentation accuracy over time.

This study tested three optimizers: Adam, Nadam, and RMSProp. Adam adapts the learning rate and often achieves good results quickly, making it widely used. Nadam combines Adam with Nesterov momentum for faster, more stable convergence on complex data. RMSProp addresses vanishing gradients by adjusting learning rates based on gradient averages, improving training stability. Testing multiple optimizers helps identify the best fit for the dataset and model architecture. Each optimizer's unique traits impact performance, guiding optimal selection. It can be concluded that Nadam Optimizer shows the best performance among the three Optimizers tested, with a Loss value of 0.09 Accuracy 0.94 Precision 0.92 Recall 0.88 IoU Coefficient 0.82 and Dice Coefficient 0.90 Meanwhile, Adam and RMSProp Optimizers have lower performance compared to Nadam, with Loss values of 0.13 and 0.23 respectively and other metric values that are also lower. From these results, Hyperparameter adjustments were then made on the model with the Nadam Optimizer by conducting experiments on the Learning Rate value. It can be concluded that the Learning Rate value of 0.0001 indicates the best performance in the U-Net model, with a Loss value of 0.09 Accuracy 0.94 Precision 0.92 Recall 0.88 IoU Coefficient 0.82 and Dice Coefficient 0.90 Meanwhile, at the Learning Rate 0.001 the model's performance has decreased significantly, with a Loss value of 0.53 Accuracy 0.64 Precision 0.39 Recall 0.83 IoU Coefficient 0.31 and Dice Coefficient 0.46 From these results, It can be concluded that the adjustment of Hyperparameter in the Learning Rate greatly affects the performance of the model. To find out the required Epoch value until this model converges, it is done by checking the number of Epochs with a limit of up to 25. Fig. 6 to 7 show the plot between the number of epochs and the performance of the matrix in the U-Net model of the results of this Tuning Hyperparameter. By applying the method proposed in this study, namely the combination of VGG19 Encoder and U-Net decoder, the best results were achieved. With a Dice Coefficient of 0.94 and an IoU Coefficient of 0.89, this method showed excellent performance in polyp segmentation, surpassing the previous two methods.

The training loss drops sharply until around epoch 10, then gradually decreases, indicating the model is nearing convergence. The validation loss also declines but stabilizes after epoch 10, suggesting possible slight overfitting. The final training loss reaches 0.05.

The training accuracy improves significantly up to epoch 10, reaching around 95%, then stabilizes, indicating the model has reached optimal learning. The validation accuracy also increases but with slower progress and noticeable fluctuations, suggesting challenges in generalization and a possibility of overfitting. Despite this, the model achieves strong training performance, with a final accuracy of 0.95.

The training Precision consistently improves throughout training, indicating increasingly accurate positive predictions. In contrast, the validation Precision fluctuates more and lacks a steady upward trend, sometimes decreasing at certain epochs. Although validation Precision does not reach the same level as training, the model demonstrates strong training performance overall. The final Precision value is 0.96.

The training Recall steadily improves throughout training, indicating the model's growing ability to detect all relevant positives in the training data. Meanwhile, the validation Recall fluctuates more and lacks a consistent upward trend, with occasional drops at certain epochs. Although validation Recall improves somewhat, it does not reach the training Recall level. This gap suggests possible overfitting, where the model performs well on training data but struggles to generalize. Overall, the model shows strong training performance. The final Recall value is 0.92. Fig. 5 presents the IoU Coefficient results over 25 epochs.

Fig. 5 shows IoU values over 25 epochs, with the X-axis representing epochs and the Y-axis showing IoU scores. The training IoU (blue line) steadily improves, indicating the model's growing accuracy in segmenting relevant areas in the training data. The validation IoU (orange line) also improves but at a slower pace and stabilizes below the training level. This suggests the model performs well but does not fully generalize to unseen data. Overall, the model demonstrates strong segmentation performance. The final IoU value is 0.89. Fig. 6 presents the Dice Coefficient results over 25.

Fig. 6 shows the Dice Coefficient over 25 epochs, with the X-axis as epochs and the Y-axis as Dice values. The training Dice (blue line) steadily improves, indicating the model's increasing effectiveness in segmenting relevant objects. The validation Dice (orange line) rises more slowly and stabilizes, showing consistent but slightly lower performance on unseen data. The final Dice Coefficient reaches 0.94, demonstrating excellent segmentation accuracy. Below are the test segmentation results using the VGG19 encoder and U-Net decoder.

Fig. 7 shows the image displays three elements: (a) Original Mask (raw data), (b) True Mask (ground truth), and (c) ROI (model prediction). Achieving a 90% Dice Coefficient, the model shows excellent agreement and high accuracy in polyp segmentation. This U-Net model supports medical education with clear visuals and aids research by analyzing large colonoscopy datasets to identify patterns and risk factors, helping improve prevention, treatment, and screening. Table I compares U-Net's segmentation performance.

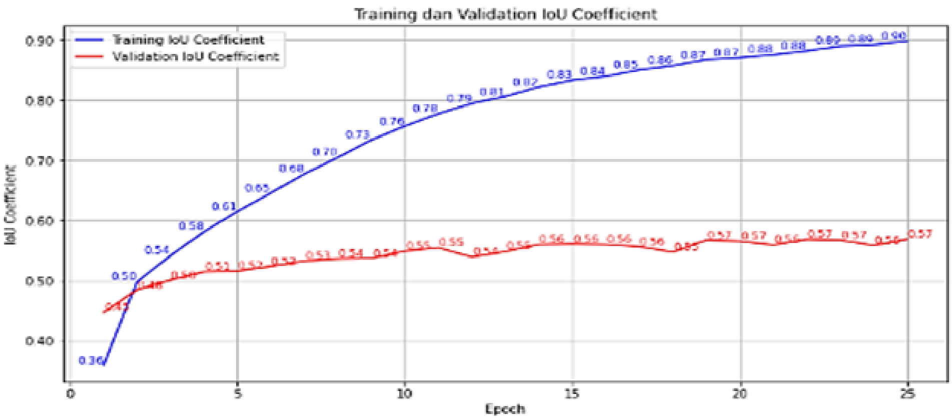


Fig. 5 IoU Coefficient test graph on matrix evaluation

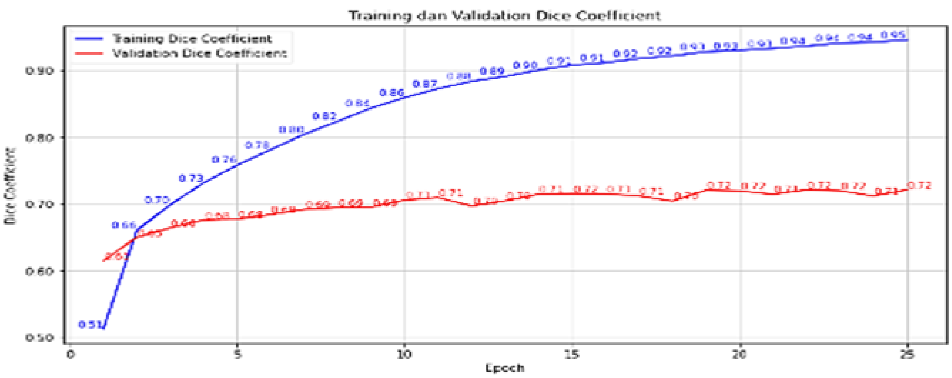
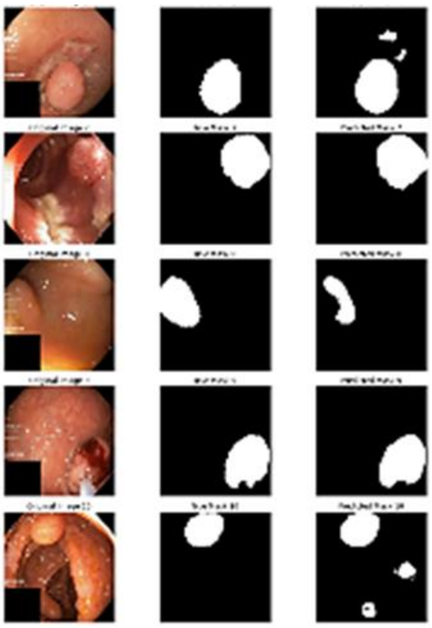


Fig. 6 Dice Coefficient test graph on matrix evaluation

TABLE I
COMPARISON OF MODEL PERFORMANCE RESULTS

Comparison of Performance Results			
Author	**Method**	**Dice Coefficient**	**IoU Coefficient**
Rishav Kumar Rajak, Ashar Beg Mirza [13].	U-Net	0,79	0,68
Devika Rajasekar, Girish Theja, Manas Ranjan Prusty, Suchismita Chinara [29].	AdaptU-Net	0,92	0,86
Proposed Method	U-Net and VGG-19 with optimizer	0,94	0,89



(a) (b) (c)

Fig. 7 Segmentation results from tests using VGG 19 encoder and U-net decoder: (a) original image (b) masking (c) ROI proposed method

Performance comparison reveals that using VGG19 as encoder with U-Net decoder improves colon polyp segmentation versus standard U-Net. Rajak and Mirza's U-Net on KVASIR achieved Dice 0.79 and IoU 0.68, while the current study's VGG19 integration enhances feature extraction, boosting accuracy and performance [13].

Devika Rajasekar. Proposed AdaptUNet for polyp segmentation on the same dataset, using skip connections to merge low-level details with high-level context. This approach improved accuracy, achieving a Dice Coefficient of 0.92 and IoU of 0.86 [29].

This study's method, combining VGG19 encoder with U-Net decoder, achieved superior results with a Dice Coefficient of 0.94 and IoU of 0.89, outperforming the previous methods in polyp segmentation accuracy.

IV. CONCLUSION

Based the complexity and variation in colon polyp endoscopic images, accurate segmentation is vital for isolating relevant structures and improving understanding. This study concludes that developing an automatic segmentation method that is both efficient and effective is crucial for enhancing polyp disease analysis. This study concludes that the development of efficient and accurate automated segmentation methods is essential to improve colon polyp image analysis. U-Net decoder and VGG19-based models show good performance with low loss values (0.05) and high evaluation metrics (0.95 Accuracy; Precision 0.96; Recall 0.92; IoU 0.89; Dice 0.94). The selection of parameters such as the Nadam optimizer, 5-Fold cross-validation, learning rate of 0.0001, and 25 epochs was proven to improve segmentation results. This hyperparameter tuning notably improved the Dice Coefficient compared to earlier studies [29] and [30]. The proposed model effectively segments polyps, with optimized parameters improving results. Several suggestions can be made for future research. One key concern is the risk of overfitting, particularly when using limited datasets, which may result in models that perform well during training but poorly on unseen data. Upcoming research should focus on improving segmentation accuracy through the design and optimization of advanced deep learning models. such as attention mechanisms, which have the potential to enhance both the accuracy and efficiency of colon polyp image segmentation.

ACKNOWLEDGEMENT

The author would like to thank ITB STIKOM Bali for the opportunity given to conduct this research and the colleagues who provided a lot of support and motivation.

REFERENCES

[1] K. Danna and R. W. Griffin, "Health and well-being in the workplace: A review and synthesis of the literature," J Manage, vol. 25, no. 3, pp. 357–384, 1999, doi: 10.1177/014920639902500305.

[2] S. K. T. Seraglio, M. Schulz, L. V. Gonzaga, R. Fett, and A. C. O. Costa, "Current status of the gastrointestinal digestion effects on honey: A comprehensive review," Sep. 30, 2021, Elsevier Ltd. doi: 10.1016/j.foodchem.2021.129807.

[3] Y. Lafau, Z. Azmi, and A. Calam, "Sistem Pakar Mendiagnosis Polip Usus Pada Manusia Menggunakan Metode Certainty Factor," vol. 3, no. 5, pp. 602–609, 2024, [Online]. Available: https://ojs.trigunadharma.ac.id/index.php/jsi

[4] U. Athiyah, I. Muhimmah, and E. Marfianti, "Ekstraksi Ciri Polip dan Pendarahan Berdasarkan Citra Endoskopi Kolorektal," Jurnal Informatika: Jurnal Pengembangan IT (JPIT), vol. 03, no. 01, 2018.

[5] D. Dornblaser, S. Young, and A. Shaukat, "Colon polyps: Updates in classification and management," Jan. 01, 2024, Lippincott Williams and Wilkins. doi: 10.1097/MOG.0000000000000988.

[6] Y. Tudela, M. Majó, N. de la Fuente, A. Galdran, A. Krenzer, F. Puppe, A. Yamlahi, T. Nuong Tran, B. J. Matuszewski, K. Fitzgerald, C. Bian, J. Pan, S. Liu, G. Fernández-Esparrach, A. Histace, J. Bernal., "A complete benchmark for polyp detection, segmentation and classification in colonoscopy images," Front Oncol, vol. 14, 2024, doi: 10.3389/fonc.2024.1417862.

[7] J. A. Yani, and U. Athiyah, "Pengenalan Polip Usus Menggunakan Neural Network". Jurnal, I. Komputer, 2024.

[8] K. Wickstrom, M. Kampffmeyer, and R. Jenssen, "Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps," Medical Image Analysis, vol. 60, Feb. 2020, doi: 10.1016/j.media.2019.101619.

[9] S. Hosseinzadeh Kassani, P. Hosseinzadeh Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Automatic Polyp Segmentation Using Convolutional Neural Networks," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer, 2020, pp. 290–301. doi: 10.1007/978-3-030-47358-7_29.

[10] J. P. Horwath, D. N. Zakharov, R. Mégret, and E. A. Stach, "Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images," NPJ Comput Mater, vol. 6, no. 1, Dec. 2020, doi: 10.1038/s41524-020-00363-x.

[11] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," Nat Commun, vol. 15, no. 1, Dec. 2024, doi: 10.1038/s41467-024-44824-z.

[12] A. M. Azor, D. J. Sharp, A. E. Jolly, N. J. Bourke, and P. J. Hellyer, "Automation and standardization of subject-specific region-of-interest segmentation for investigation of diffusion imaging in clinical populations," PLoS One, vol. 17, no. 12 December, Dec. 2022, doi: 10.1371/journal.pone.0268233.

[13] A. Mirza and R. K. Rajak, "Segmentation of Polyp Instruments using UNet based deep learning model," Nordic Machine Intelligence, vol. 1, no. 1, pp. 44–46, Nov. 2021, doi: 10.5617/nmi.9145.

[14] M. Yeung, E. Sala, C. B. Schönlieb, and L. Rundo, "Focus U-Net: A novel dual attention-gated CNN for polyp segmentation during colonoscopy," Comput Biol Med, vol. 137, Oct. 2021, doi: 10.1016/j.compbiomed.2021.104815.

[15] D. Jha, P. Helén Smedsrud, D. Johansen, T. de Lange., "A Comprehensive Study on Colorectal Polyp Segmentation with ResUNet++, Conditional Random Field and Test-Time Augmentation," Jul. 2021, [Online]. Available: http://arxiv.org/abs/2107.12435

[16] D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, H. D. Johansen., "Kvasir-SEG: A Segmented Polyp Dataset," Nov. 2019, [Online]. Available: http://arxiv.org/abs/1911.07069

[17] N. Sharma, S. Gupta, D. Gupta, P. Gupta , S. Juneja, A. Shah, A. Shaikh., "UMobileNetV2 model for semantic segmentation of gastrointestinal tract in MRI scans," PLoS One, vol. 19, no. 5 May, May 2024, doi: 10.1371/journal.pone.0302880.

[18] L. Mardiana, D. Kusnandar, and N. Satyahadewi, "Analisis Diskriminan Dengan K Fold Cross Validation Untuk Klasifikasi Kualitas Air Di Kota Pontianak," 2022.

[19] M. A. Rajab, F. A. Abdullatif, and T. Sutikno, "Classification of grapevine leaves images using VGG-16 and VGG-19 deep learning nets," Telkomnika (Telecommunication Computing Electronics and Control), vol. 22, no. 2, pp. 445–453, 2024, doi: 10.12928/TELKOMNIKA.v22i2.25840.

[20] W. I. Kusumawati and A. Z. Noorizki, "Perbandingan Performa Algoritma VGG16 Dan VGG19 Melalui Metode CNN Untuk Klasifikasi Varietas Beras," Journal of Computer, Electronic, and Telecommunication, vol. 4, no. 2, Dec. 2023, doi: 10.52435/complete.v4i2.387.

[21] P. D. Wulaning Ayu and G. A. Pradipta, "U-Net Tuning Hyperparameter for Segmentation in Amniotic Fluid Ultrasonography Image," in 2022 4th International Conference on Cybernetics and Intelligent System, ICORIS 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICORIS56080.2022.10031294.

[22] O. Hospodarskyy, V. Martsenyuk, N. Kukharska, A. Hospodarskyy, and S. Sverstiuk, "Understanding the Adam Optimization Algorithm in Machine Learning," 2024.

[23] Z. Zhixuan and H. Zaien, "Research and Application of Rectified-NAdam Optimization Algorithm in Data Classification," American Journal of Computer Science and Technology, vol. 4, no. 4, p. 106, 2021, doi: 10.11648/j.ajcst.20210404.13.

[24] A. Khozaimi and W. Firdaus Mahmudy, "New insight in cervical cancer diagnosis using convolution neural network architecture," IAES International Journal of Artificial Intelligence (IJ-AI), vol. 13, no. 3, pp. 3092–3100, 2024, doi: 10.11591/ijai.v.

[25] Liu, Jinlan, Xu, Dongpo, Zhang, M. Huisheng, Danilo., "On hyper-parameter selection for guaranteed convergence of RMSProp," Cogn Neurodyn, 2022.

[26] P. Fan, Y. Diao, F. Li, W. Zhao, and Z. Chen, "SRSegNet: Super-resolution-assisted small targets polyp segmentation network with combined high and low resolution," Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 3, Mar. 2024, doi: 10.1016/j.jksuci.2024.101981.

[27] R. Islam, R. S. Akash, M. A. Hossen Rony, and M. Z. Hasan, "SAMU-Net: A dual-stage polyp segmentation network with a custom attention-based U-Net and segment anything model for enhanced mask prediction," Array, vol. 24, Dec. 2024, doi: 10.1016/j.array.2024.100370.

[28] P. D. W. Ayu, G. A. Pradipta, R. R. Huizen, E. S. W. Kadek, and I. G. E. Artana, "Combining CNN Feature Extractors and Oversampling Safe Level SMOTE to Enhance Amniotic Fluid Ultrasound Image Classification," International Journal of Intelligent Engineering and Systems, vol. 17, no. 1, pp. 251–262, 2024, doi: 10.22266/ijies2024.0229.24.

[29] D. Rajasekar, G. Theja, M. R. Prusty, and S. Chinara, "Efficient colorectal polyp segmentation using wavelet transformation and AdaptUNet: A hybrid U-Net," Heliyon, vol. 10, no. 13, Jul. 2024, doi: 10.1016/j.heliyon.2024.e33655.