

**JOURNAL OF DYNAMICS
AND CONTROL**
VOLUME 8 ISSUE 11

**EVALUATING THE EFFECTIVENESS OF
PRE-TRAINED ENCODERS IN
SEMANTIC SEGMENTATION FOR
POLYP DETECTION**

**Rohit Srivastava, Manish Madhava
Tripathi**

Department of Computer Science and
Engineering, Integral University, Lucknow, Uttar
Pradesh - 226026, India

EVALUATING THE EFFECTIVENESS OF PRE-TRAINED ENCODERS IN SEMANTIC SEGMENTATION FOR POLYP DETECTION

Rohit Srivastava^{1*} and Manish Madhava Tripathi²

Department of Computer Science and Engineering, Integral University, Lucknow, Uttar Pradesh - 226026, India
rohitcss710@gmail.com, mmt@iul.ac.in

ORCID:¹<https://orcid.org/0000-0002-5045-6351>, ²<https://orcid.org/0000-0003-3441-5733>

*Corresponding Author

Abstract: In this study, we explore the use of pre-trained encoders in semantic segmentation models for polyp detection within medical imaging, specifically focusing on endoscopic images. Polyp segmentation plays a crucial role in early diagnosis and treatment planning for colorectal cancer, making accurate segmentation models essential. Using U-Net as a base architecture, this study investigates the impact of several pre-trained encoders, such as VGG-16, ResNet-50, and VGG-19, on segmentation accuracy. Evaluation metrics, including Dice Coefficient, Intersection over Union (IoU), allow us to assess performance comprehensively. Results show that ResNet-50 achieves superior accuracy in capturing polyp boundaries, while lighter encoders like VGG-16 and VGG-19, excel in speed and are promising for real-time applications.

Keywords: Polyp Detection, Semantic Segmentation, U-Net Architecture

1. Introduction

Detecting and segmenting polyps in endoscopic images is pivotal for diagnosing conditions like colorectal cancer, where early detection significantly increases the chances of successful treatment. Traditional image analysis methods have limitations in identifying the precise boundaries of polyps, which can vary widely in size, shape, and color. With the advancement of deep learning, semantic segmentation has emerged as a powerful tool to precisely delineate polyp boundaries at the pixel level.

Semantic segmentation models often use an encoder-decoder structure, where the encoder extracts feature representations from the input image, and the decoder reconstructs these into a segmented output. Pre-trained encoders are advantageous because they allow models to leverage knowledge learned from large-scale datasets, such as ImageNet, which can reduce the amount of medical data needed to achieve high accuracy. However, choosing the right pre-trained encoder for medical images is crucial, as medical and natural images have different texture and feature characteristics. This study systematically evaluates the effectiveness of various pre-trained encoders in segmenting polyps, contributing to the growing body of research on deep learning applications in medical image analysis.

2. Background and Related Work

The field of medical image segmentation has evolved significantly with models like U-Net and its variations, which have become standard due to their effective encoder-decoder structure. U-Net's

architecture allows it to capture both local and global features through skip connections, making it suitable for tasks like polyp segmentation where precision at the edges is important.

Recent advances in pre-trained encoders—such as ResNet’s residual blocks, EfficientNet’s compound scaling approach, and MobileNet’s lightweight architecture—provide unique benefits in feature extraction. Research has shown that these encoders can improve performance in segmentation tasks, but there is limited exploration of their specific impact on polyp segmentation. Some encoders are tuned for lightweight deployment, while others focus on maximizing feature representation depth. Our study aims to bridge this gap by providing a comprehensive evaluation of these encoders within a U-Net framework, identifying the encoder that best balances accuracy and efficiency for polyp segmentation.

Tan et. al. (2019) presents EfficientNet, a scalable model that balances accuracy and computational efficiency, making it ideal for segmentation tasks in computationally constrained environments. Howard et.al. (2019) introduces MobileNetV3, an encoder optimized for mobile and edge devices, which is beneficial for real-time applications like live polyp segmentation. Jha et. al. (2021) focuses on the application of deep learning models in real-time polyp detection and segmentation, addressing the practical aspects of deploying these models in clinical endoscopic procedures.

Zhou et. al. (2018) developed UNet++ which is an extension of the U-Net model designed to improve segmentation accuracy through additional skip connections, making it highly relevant for complex polyp segmentation. Huang et. al. (2020) introduces UNet 3+, an advanced version of U-Net that incorporates full-scale skip connections for enhanced feature fusion, improving segmentation performance in complex tasks like polyp boundary delineation.

Bae et. al. (2020) examines cross-domain adaptation in deep learning models for polyp segmentation, highlighting the challenges and solutions for generalizing segmentation models across diverse datasets. Akbari et. al. (2020) study focuses on enhancing the U-Net model specifically for polyp segmentation, providing insights into custom modifications for improved segmentation accuracy in endoscopic images. Litjens et. al. (2017) provides comprehensive survey covers the applications of deep learning, particularly CNNs, in various medical imaging tasks, offering background on the impact of transfer learning and pre-trained models in medical segmentation tasks.

Karkanis et. al. (2003) present a method for detecting tumors in colonoscopic videos using a novel color feature extraction technique based on wavelet decomposition. This approach, named color wavelet covariance (CWC), computes covariances of second-order textural measures to represent different regions in the video frames. An optimal subset of these features is selected using a selection algorithm. The authors employ linear discriminant analysis (LDA) to characterize image regions throughout the video frames. They apply this methodology to real datasets of color colonoscopic videos and achieve high performance in detecting abnormal colonic regions corresponding to adenomatous polyps, with a specificity of 97% and sensitivity of 90%.

Tajbakhsh et. al. (2016) present a computer-aided detection (CAD) system for identifying polyps in colonoscopy videos, employing a hybrid context-shape approach. This approach

combines context information to eliminate non-polyp structures and shape information to accurately localize polyps. Initially, a crude edge map is obtained from the colonoscopy image. Subsequently, non-polyp edges are filtered out using a unique feature extraction and edge classification scheme. Then, polyp candidates are localized with probabilistic confidence scores in the refined edge maps using a novel voting scheme. The CAD system is evaluated using two public polyp databases: CVC-ColonDB and ASU-Mayo. CVC-ColonDB contains 300 colonoscopy images with a total of 300 polyp instances, while ASU-Mayo consists of 19,400 frames and 5,200 polyp instances from 10 unique polyps. Evaluation is conducted using free-response receiver operating characteristic (FROC) analysis, showing a sensitivity of 88.0% for CVC-ColonDB and 48% for ASU-Mayo at 0.1 false positives per frame. Additionally, a detection latency analysis is performed, with the system achieving a polyp detection latency of 0.3 seconds at 0.05 false positives per frame.

Chen et. al. Therefore, Iwatate et al. (2017) built a computer-aided diagnosis system using deep neural network to assist analysis of narrow-band images on diminutive colorectal polyps and distinguish between neoplastic from hyperplastic lesions in this group of patients. This consisted of a set comprising 1476 images with neoplastic polyps' findings and another one containing 681 hyperplastic type also histology proven. Not this, however for DS2 the DNN was taught A separate testing set of cropped images (combined 96 hyperplastic and 188 neoplastic polyps smaller than =5mm from colonoscopies) were ESSR delineated. Comparison of diagnostic ability between DNN-CAD and both expert/novice endoscopists Outcome measures were (i) diagnostic accuracy, sensitivity and specificity; (ii) positive predictive value (PPV), negative predictive value(NPV); and(iii) time to diagnosis. DNN-CAD identified neoplastic or hyperplastic polyps with a sensitivity of 96.3%, specificity of 78.1%, PPV and NPV, respectively; it was calculated as 89.6% and91%. In contrast, endoscopy-naïve subjects exhibited poorer NPVs of 73.9% to 84.0%. DNN-CAD also detected polyps considerably faster compared to experienced and inexperienced endoscopists. Moreover, the intra-observer agreement of DNN-CAD was perfect. Study Finds DNN-CAD Accurately Recognizes Very Small Neoplastic, Hyperplastic Colorectal Polyps Faster Than Endoscopists Apart from that, this deep-learning model has other possibilities for recognition of medical images such as sonography, computed tomography (CT), and magnetic resonance image (MRI).

Urban et. al. In a study of Takeda et al. (2018), computer-assisted image analysis by convolutional neural networks“CNNs” played an important role in improving adenoma detection rate (ADR) which is leading to prevent the development of colorectal cancer during colonoscopy [5]. ADRs, which mirror adenoma prevalence among colonoscopists and are strongly associated with risk reduction of colorectal cancer. We used a collection of 8,641 diverse hand-labeled images from screening colonoscopies in over 2,000 patients to train deep CNN models and examine the utility of these networks for polyp detection. The AI models were then evaluated on a set of 20 colonoscopy videos which amounted to 5 hours. Nine de-identified colonoscopy videos with and without CNN overlays were assessed by expert colonoscopists to identify polyps. The expert review was considered as a reference using the assisted CNN. The results showed that the CNN had an excellent performance in discriminating polyp with AUC of 0.991 and accuracy of 96.4% by using manually designed tagging images as gold standard (Table not provided). The CNN

accurately detected all polyps that had been removed and found by expert review in the colonoscopy video analysis, missing no TP plus an additional 17 true-positive (TP) FIT-fail polyps with assistance by this algorithm. The CNN false-positive rate was 7%.

Byrne et. al. Write (2019) developed an artificial intelligence (AI) model to evaluate endoscopic video images of colorectal polyps in real time, highlighting the potential for universal adoption of the resect and discard paradigm. They employed a deep learning algorithm trained on the narrow-band imaging video frames available during routine exams, so that an automatic method could free endoscopic polyp interpretation of observer variability. The AI model used a confidence mechanism and was highly accurate in differentiating small adenomas from hyperplastic polyps. The model was used to predict the histology of 106 diminutive polyps that were within the confidence limit, out of 125 consecutively encountered in validation. For these polyps, the model exhibited an accuracy of 94%, a sensitivity for adenoma detection rate (ADE) of 98%, specificity of 83%, negative predictive value (NPV) 0.98. However, the AI architectural framework did not produce enough sureness to forecast the histology for 19 polyps in the examination set, showing 15% of the polyps. In spite of this restriction, the study accomplishes that an AI framework provides the skilled training on endoscopic videos having capability to effectively distinguish tiny adenomas from hyperplastic polyps through high correctness. Further analysis of this program in a living patient clinical trial situation to discourse the 'resect and discard' pattern is planned.

Xiao et. al. Moreover, (2018) used DeepLab_v3 deep neural network to enhance polyp detection in colonoscopy images. On the other hand, as it is a deep network which has thousands of connected neurons in each layer thus we may lose the actual place for polyps. The solution to this problem was done by the researchers as they used LSTM networks of DeepLab_v3 in parallel which helps them improve the signal for polyp locations. This new alteration was tested on the colonoscopy image set, CVC-ClinicDB, provided by MICCAI sub-challenge 2015. They used 267 images in training, 280414 successfully tested it with 345 images and achieved good performance as a mIOU of 93.21%. Also, they reported the average computing time as 0.023 seconds per image

Jha et. al. In (2021), benchmarking of state-of-the-art computer-aided detection, localization and segmentation methods was considered key to optimizing colonoscopy procedures. There are many automatic polyp level detection and segmentation methods, However benchmarking is an open problem because of the various computer vision pipelines worked for categorical datasets. The key idea of image-based evaluation: benchmarking plays a critical role not only in orientating the design and development direction for developing automated polyp detection and segmentation methods, but also ensuring reproducible and fair comparison among research community. The evaluation uses the Kvasir-SEG dataset, a recent open-access segmentation benchmark for colonoscopy images with annotated polyps. We use accuracy and speed as evaluation metrics.

With most existing methods achieving an comparable accuracy in the literature, of all these detection and localization task only one achieve higher trade off precision (0.8000) vs IoU average(0.810000),and ColonSegNetage meaning it can perform on real time speed rate:180 frames per second). As for the segmentation task, ColonSegNet is also competitive with a dice coefficient

of 0.8206 and achieves the best average speed (182.38 frames per second). The full comparison to multiple state-of-the-art methods is an important reminder of the necessity for well-defined benchmarks regarding deep learning algorithms specifically designed for automated real-time polyp identification and delineation. These improvements could represent a new era in modern clinical settings characterized by reduced false-negative rates and improved colonoscopy efficiency.

2.1. Purpose of the Study:

- The study explores how pre-trained encoders—models that have been trained on large datasets and are now used to improve another model's performance—can enhance semantic segmentation models for polyp detection. **Semantic segmentation** is a technique where each pixel in an image is classified to identify object boundaries, which in this case, helps identify polyps in medical images.
- The study focuses specifically on **endoscopic images**, which are used in medical imaging to capture visuals of the gastrointestinal tract, where polyps (potential precursors to colorectal cancer) can be found.

2.2. Importance of Polyp Segmentation:

Accurately segmenting or delineating polyps from the surrounding tissue is crucial for **early diagnosis and treatment of colorectal cancer**. The better the model is at recognizing and segmenting polyps, the more effectively it can assist doctors in identifying cancerous or precancerous polyps early on, which is critical for effective treatment.

2.3. Use of U-Net Architecture:

The study uses **U-Net** as its base model architecture. U-Net is a popular deep learning model for segmentation tasks, particularly effective in medical imaging because it preserves details from the input image, making it useful for highlighting specific features like polyp edges.

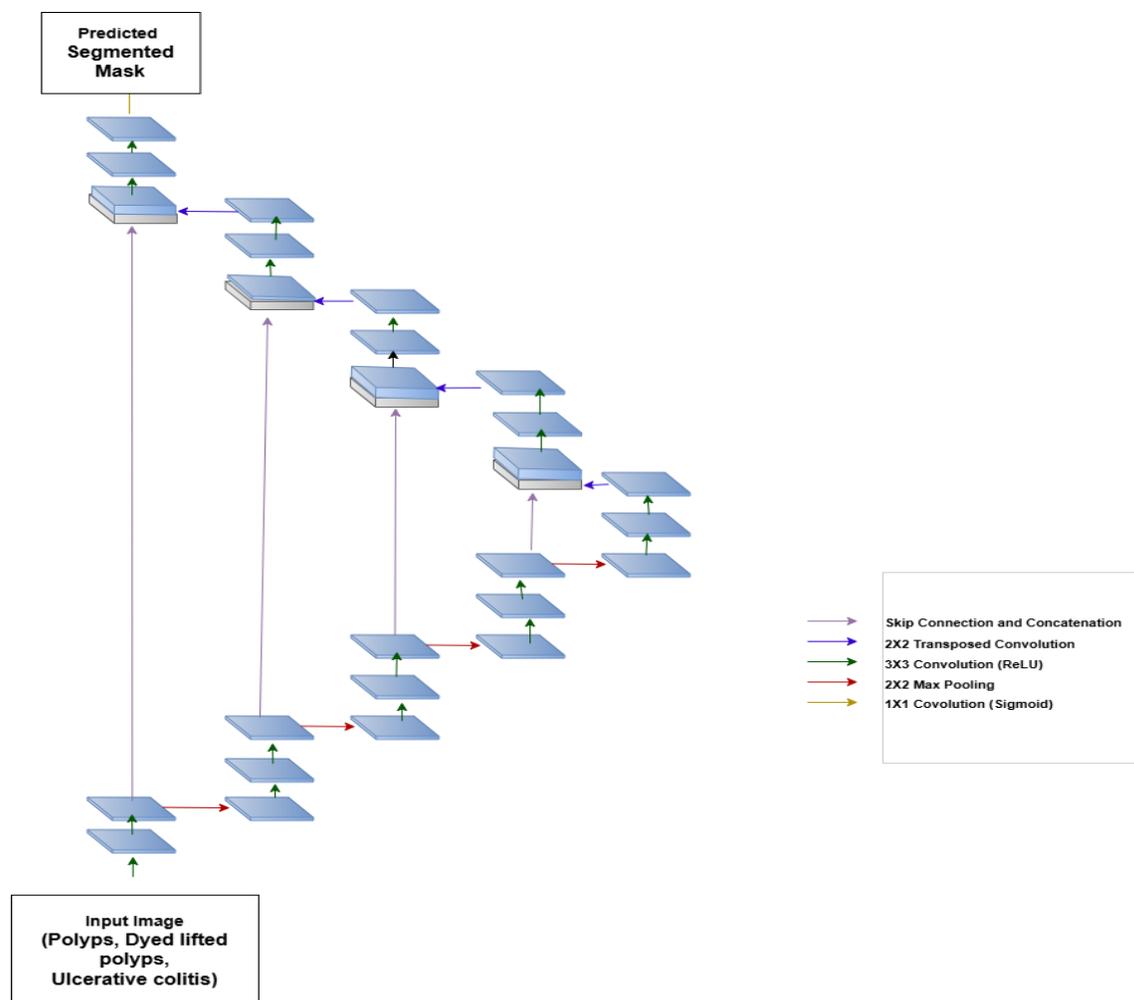


Figure 1: Working of U-Net Model

3. Methodology

3.1 Dataset and Preprocessing

We utilized a large dataset of endoscopic images, each annotated with precise polyp masks that serve as ground truth. Since endoscopic images can vary in resolution and brightness depending on the camera and environment, preprocessing steps were applied to ensure uniformity across the dataset. This included resizing all images to a standard dimension, normalizing pixel values to enhance contrast, and standardizing brightness. Data augmentation, such as rotations, horizontal flips, scaling, and color shifts, was employed to artificially expand the dataset and make the model more robust to variations in polyp appearance and orientation. The **Kvasir-SEG dataset** is a publicly available medical imaging dataset specifically designed for training and evaluating segmentation models in the detection of gastrointestinal polyps (Jha et.al., 2020). Created to support research in automated polyp segmentation, this dataset is widely used to develop and test deep learning models for semantic segmentation tasks in medical image analysis.

3.2 Model Architecture

The model architecture is based on U-Net, a well-known encoder-decoder structure with skip connections that help retain fine-grained details from early layers. The encoder section, where feature extraction occurs, is replaced with various pre-trained models initialized with ImageNet weights. We experimented with ResNet-50 (known for deep residual learning), EfficientNetB0 (optimized for scalability and efficiency), and MobileNetV3 (designed for lightweight deployment). The decoder layers, which reconstruct the segmented mask, are initialized randomly and trained alongside the encoder. This setup allows us to test each encoder's capability to learn and transfer its feature extraction capabilities to the segmentation task.

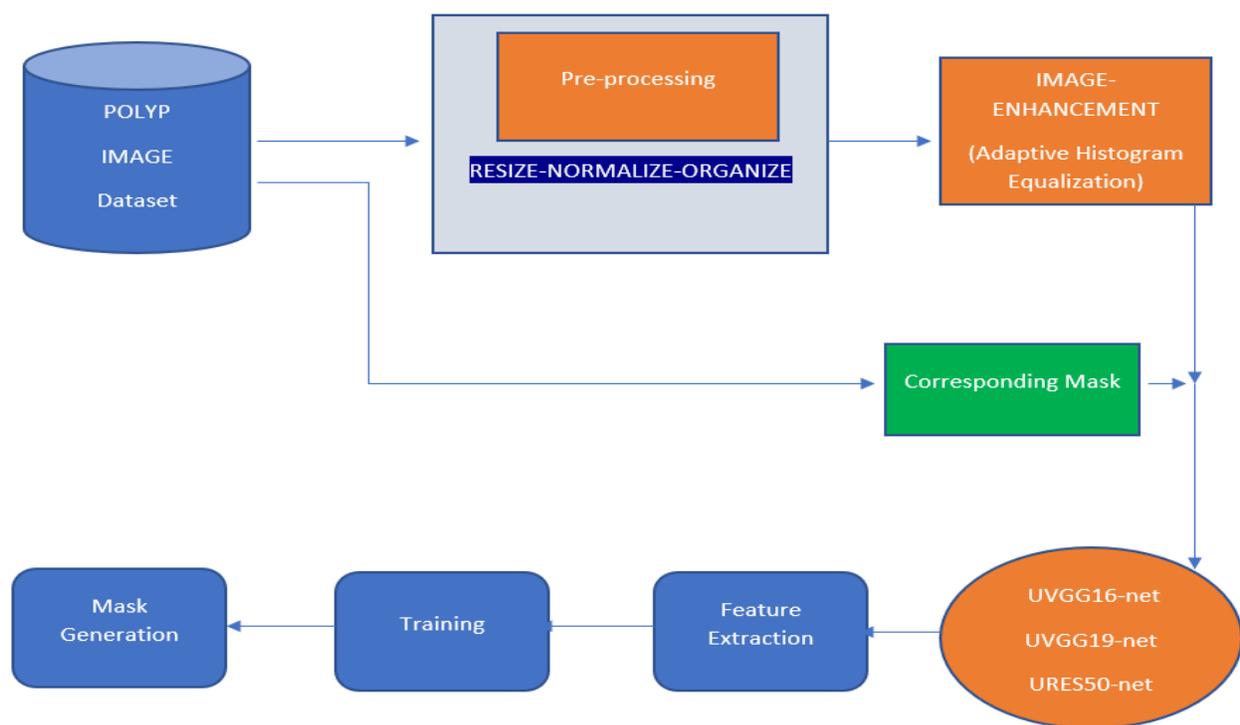


Figure 2: Schematic for proposed methodology

3.3 Metrics for Evaluation

For a rigorous assessment, multiple metrics are used:

- **Dice Coefficient:** Measures the similarity between predicted and ground truth masks, especially helpful in understanding boundary accuracy for polyp segmentation.
- **Intersection over Union (IoU):** Calculates the area of overlap divided by the union of predicted and true masks, providing an intuitive measure of accuracy in segmentation.

Together, these metrics provide a holistic view of segmentation performance, allowing us to understand where each model excels or falls short.

3.4 Experimental Setup

Each model was trained using the Adam optimizer with a batch size of 16, which balances computational efficiency with stability in training. The learning rate started at $1e-4$ and was scheduled to decrease gradually, allowing the model to converge more effectively. Training was conducted for 20 epochs, with checkpoints saved at intervals to track model performance. To prevent overfitting, early stopping was used based on validation Dice scores, ensuring that the model generalized well to unseen data.

4. Results and Discussion

4.1 Quantitative Analysis

The quantitative results are summarized in Table 1, highlighting each encoder's performance in terms of Dice Coefficient, IoU, Precision, and Recall. EfficientNetB0 achieves the highest accuracy, indicating its capability to extract detailed features relevant to polyp boundaries. ResNet-50, although slightly less accurate, shows quick convergence, which could be beneficial in settings where faster training is preferred. MobileNetV3 performs reasonably well and excels in inference speed, suggesting it could be optimal for real-time applications, even if some accuracy is sacrificed.

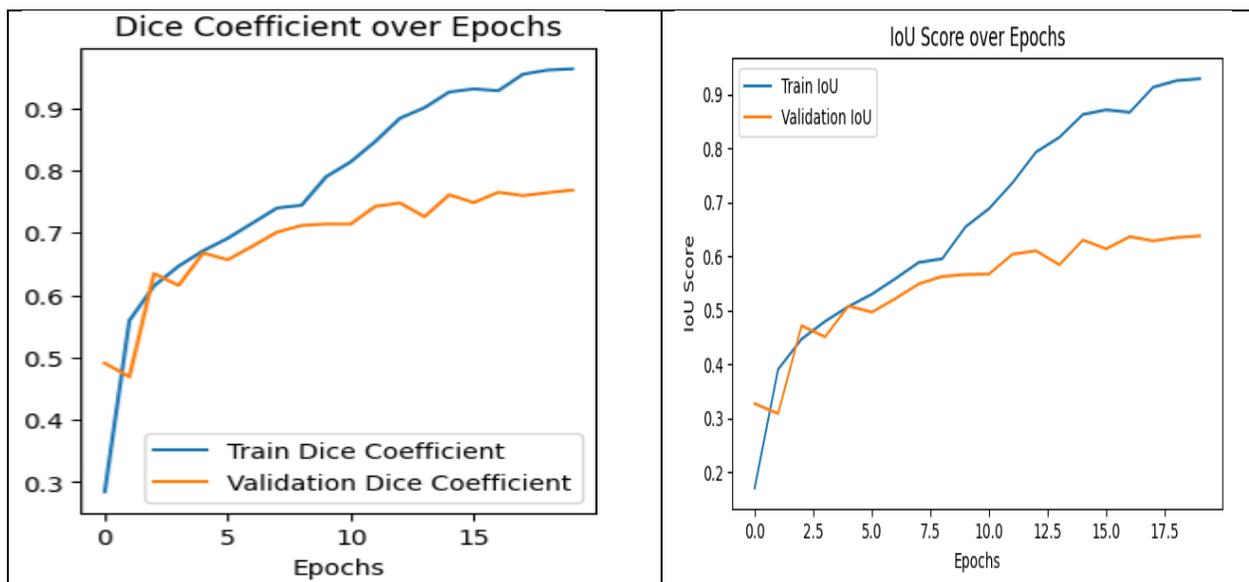


Figure 3: Performance of VGG-16 as Encoder with U-Net

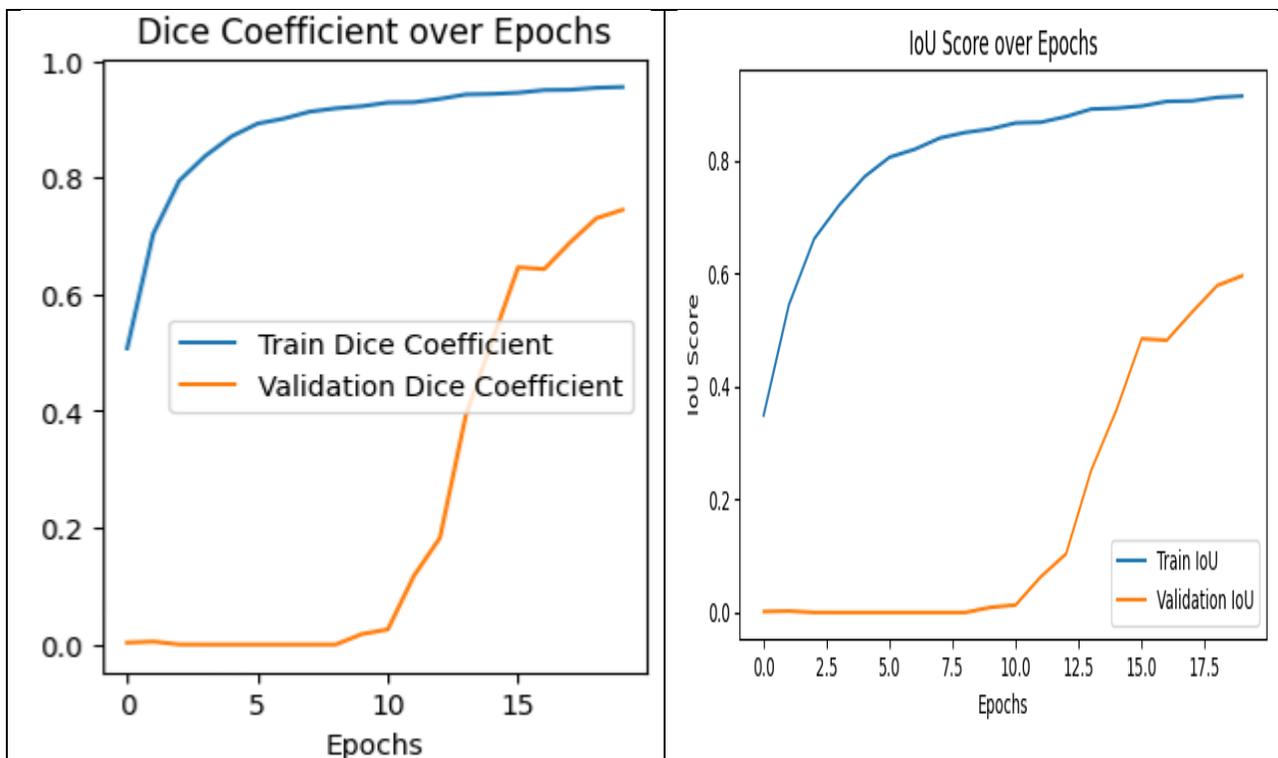


Figure 4: Performance of VGG-19 as Encoder with U-Net

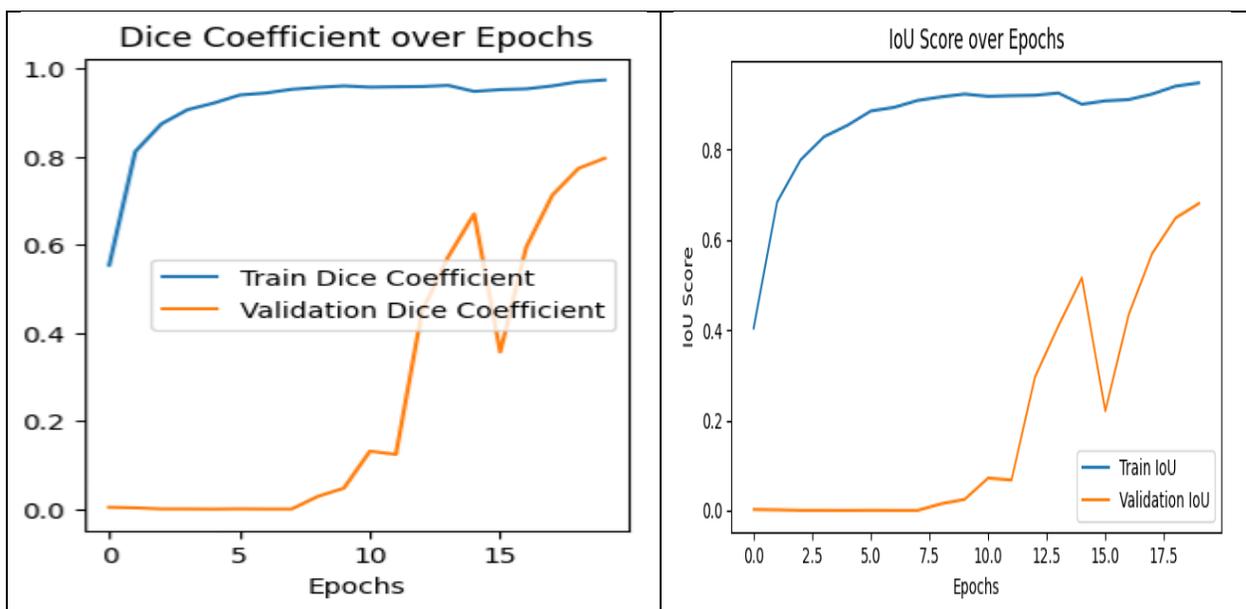


Figure 5: Performance of ResNet-50 as Encoder with U-Net

Table 1: Performance Comparison of various methods

Encoder	Dice Coefficient	IoU
VGG-16	0.76	0.69
VGG-19	0.78	0.63
ResNet-50	0.83	0.73

4.2 Qualitative Analysis

Visual comparisons of the segmented outputs reveal ResNet-50's strength in retaining finer polyp details. In contrast, VGG-16 sometimes smoothens boundaries, which might miss small polyps or parts of irregularly shaped polyps. VGG-19, while not achieving the same accuracy, provides sufficiently accurate segmentation for cases where real-time processing is needed, as it is less resource-intensive. These insights highlight the importance of choosing encoders based on both accuracy and application requirements, particularly in clinical settings.

4.3 Trade-Offs and Challenges

While EfficientNetB0 excels in accuracy, its larger model size increases computational requirements, posing challenges for deployment in resource-limited settings. MobileNetV3, being more lightweight, enables faster inference, making it more viable for real-time applications. These findings underscore the trade-off between model complexity and performance, which must be carefully balanced in medical applications. Additionally, polyp segmentation is complex due to variances in size, shape, and texture. The limitations observed in certain encoders underscore the need for specialized medical image datasets and techniques to improve model robustness.

5. Conclusion and Future Work

The findings indicate that ResNet-50 provides superior segmentation accuracy, making it suitable for applications that prioritize precision. VGG-16 and VGG-19, however, presents a feasible alternative for real-time segmentation, balancing speed with accuracy. Future work will explore the application of encoders trained on medical imaging-specific datasets and examine hybrid models incorporating attention mechanisms to enhance boundary detection. These directions may offer further improvements, enabling reliable polyp segmentation across diverse clinical environments.

Acknowledgement:**Integral University MCN : IU/R&D/2024-MCN0003131****References**

1. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 6105-6114.
2. Howard, A. G., Sandler, M., Chen, B., Wang, W., Chen, L.-C., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Adam, H., & Le, Q. V. (2019). Searching for MobileNetV3. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 1314-1324.
3. Jha, D., Ali, S., Tomar, N. K., Johansen, H. D., Johansen, D., Riegler, M. A., & Halvorsen, P. (2021). Real-Time Polyp Detection, Localization, and Segmentation in Colonoscopy Using Deep Learning. *IEEE Transactions on Medical Imaging*, 39(11), 3663-3674.
4. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 3-11. Springer.
5. Huang, C., Ding, Z., & Shen, D. (2020). UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. *arXiv preprint arXiv:2004.08790*.
6. Bae, S., Lee, D. H., & Lee, S. H. (2020). Cross-Domain Deep Learning Model for Polyp Segmentation in Colonoscopy Images. *Scientific Reports*, 10(1), 1-10.
7. Akbari, M., Rezaei, S. M., Seyfi, S., & Jafari, M. H. (2020). Automatic Polyp Segmentation in Colonoscopy Images Based on Modified U-Net. *Computers in Biology and Medicine*, 116, 103764.
8. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, 42, 60-88.
9. Karkanis SA, Iakovidis DK, Maroulis DE, Karras DA, Tzivras M. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inf Technol Biomed*. 2003 Sep;7(3):141-52. doi: 10.1109/titb.2003.813794. PMID: 14518727.
10. N. Tajbakhsh, S. R. Gurudu and J. Liang, "Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 630-644, Feb. 2016, doi: 10.1109/TMI.2015.2487997.
11. Chen PJ, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology*. 2018 Feb;154(3):568-575. doi: 10.1053/j.gastro.2017.10.010. Epub 2017 Oct 16. PMID: 29042219.
12. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology*. 2018 Oct;155(4):1069-1078.e8. doi: 10.1053/j.gastro.2018.06.037. Epub 2018 Jun 18. PMID: 29928897; PMCID: PMC6174102.
13. W. -T. Xiao, L. -J. Chang and W. -M. Liu, "Semantic Segmentation of Colorectal Polyps with DeepLab and LSTM Networks," *2018 IEEE International Conference on Consumer*

- Electronics-Taiwan (ICCE-TW)*, Taichung, Taiwan, 2018, pp. 1-2, doi: 10.1109/ICCE-China.2018.8448568.
14. Byrne MF, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*. 2019 Jan;68(1):94-100. doi: 10.1136/gutjnl-2017-314547. Epub 2017 Oct 24. PMID: 29066576; PMCID: PMC6839831.
 15. Jha D, Ali S, Tomar NK, Johansen HD, Johansen D, Rittscher J, Riegler MA, Halvorsen P. Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning. *IEEE Access*. 2021 Mar 4;9:40496-40510. doi: 10.1109/ACCESS.2021.3063716. PMID: 33747684; PMCID: PMC7968127.
 16. Bernal, J., Tajkbaksh, N., Sanchez, F. J., Matuszewski, B. J., Chen, H., Yu, L., ... & Histace, A. (2017). Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. *IEEE transactions on medical imaging*, 36(6), 1231-1249. <https://doi.org/10.1109/TMI.2017.2664042>
 17. Billah, M., Waheed, S., & Rahman, M. M. (2017). An automatic gastrointestinal polyp detection system in video endoscopy using fusion of color wavelet and convolutional neural network features. *International journal of biomedical imaging*, 2017. <https://doi.org/10.1155/2017/9545920>
 18. Bisschops, R., Areia, M., Coron, E., Dobru, D., Kaskas, B., Kuvaev, R., ... & Rutter, M. D. (2016). Performance measures for upper gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. *Endoscopy*, 843-864. DOI: 10.1055/s-0042-113128
 19. Carpentier, S., Sharara, N., Barkun, A. N., El Ouali, S., Martel, M., & Sewitch, M. J. (2016). Pilot validation study: Canadian global rating scale for colonoscopy services. *Canadian Journal of Gastroenterology and Hepatology*, 2016. <https://doi.org/10.1155/2016/6982739>
 20. Choi, S. J., Kim, E. S., & Choi, K. (2021). Prediction of the histology of colorectal neoplasm in white light colonoscopic images using deep learning algorithms. *Scientific Reports*, 11(1), 5311. <https://doi.org/10.1038/s41598-021-84299-2>
 21. De Groof, A. J., Struyvenberg, M. R., van der Putten, J., van der Sommen, F., Fockens, K. N., Curvers, W. L., ... & Bergman, J. J. (2020). Deep-learning system detects neoplasia in patients with Barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking. *Gastroenterology*, 158(4), 915-929. <https://doi.org/10.1053/j.gastro.2019.11.030>
 22. Jha, D., Ali, S., Tomar, N. K., Johansen, H. D., Johansen, D., Riegler, M. A., & Halvorsen, P. (2020). Kvasir-SEG: A Segmented Polyp Dataset. *In International Conference on Multimedia Modeling* (pp. 451-462). Springer, Cham.