RESEARCH ARTICLE

POLYP IMAGE SEGMENTATION USING DEEP LEARNING TECHNIQUES: RESUNET++ ARCHITECTURE

Tran Nguyen Quynh Tram

Faculty of Information Technology, HUFLIT

tramtnq@huflit.edu.vn

ABSTRACT— This study presents a novel polyp segmentation approach using ResUnet++. Trained on Kvasir-SEG and CVC-ClinicDB, ResUnet++ significantly outperforms traditional UNet and ResUnet. Its residual blocks and attention mechanisms enhance feature extraction, leading to improved segmentation in challenging cases. This highlights the potential of deep learning for advancing polyp segmentation and improving early colorectal cancer detection. Future research could explore further modifications or alternative architectures.

Keywords— Image segmentation, colonoscopy, deep learning, computer vision, health informatics

I. INTRODUCTION

Colon cancer[1] remains a pressing global health issue, ranking among the leading causes of cancer-related deaths. In 2022, there were an estimated 1.9 million new cases and 904,000 deaths worldwide. To accurately compare cancer rates across populations with varying age structures, age-standardized rates (ASRs) are crucial. The global ASR for colorectal cancer incidence in 2022 reached a concerning 18.4 per 100,000 individuals, highlighting the urgent need for enhanced prevention, early detection, and treatment strategies.

Early detection of colon cancer is crucial for improved patient outcomes. Regular colonoscopies are recommended to identify and remove polyps[2], which can develop into cancer. However, accurate polyp segmentation in medical images remains a challenging task due to their complex nature, diverse appearances, and the presence of noise and artifacts. Robust segmentation methods are essential for precise diagnosis and treatment planning.

This work presents a novel approach for segmenting polyp images using a modified ResUNet++[3] architecture and implement it using the TensorFlow framework. ResUNet++, a medical image segmentation architecture built upon the ResUNet[4] architecture, created by Debesh Jha and his team. It takes advantage of Residual Networks[5], Squeeze and Excitation blocks[6], Atrous Spatial Pyramidal Pooling (ASPP)[7], and attention blocks[8]. The primary goal of this research is to enhance the feature extraction capabilities of the deep learning model, thereby improving the accuracy of polyp segmentation for automated recognition and diagnostic support systems.

The paper is organized as follows:

I. Introduction: Provides a general overview of the research topic and its significance.

II. Related Work: Reviews existing methods for polyp image segmentation, highlighting their strengths and weaknesses.

III. Proposed Method: Details the proposed modified ResUNet++ architecture and its implementation.

IV. Experiments and Results: Presents experimental results, including performance metrics and comparisons with state-of-the-art methods.

V. Conclusion: Summarizes the key findings of research.

II. RELATED WORK

A. POLYP IMAGE SEGMENTATION

Image segmentation, a key computer vision task, involves classifying each pixel in an image. In medical image analysis, polyp segmentation aims to identify and delineate polyp regions in colonoscopy images. Accurate polyp segmentation is crucial for early detection and diagnosis of colorectal cancer.

B. TRADITIONAL METHOD AND DEEP LEARNING METHOD

In the field of medical image segmentation, traditional methods often include techniques such as thresholding, clustering, and contour or region-based methods. The emergence of deep learning has brought significant advancements to medical image segmentation with models such as convolutional neural networks (CNN) and their variants. U-Net, in particular, is a prominent example, introduced by Olaf Ronneberger[9] and colleagues in the paper "U-Net: Convolutional Networks for Biomedical Image Segmentation" (2015) depicted in (Figure 1).

This model uses a symmetric architecture with skip connections to improve information transmission during the learning process, allowing the network to segment objects with high accuracy, even in cases of limited data.

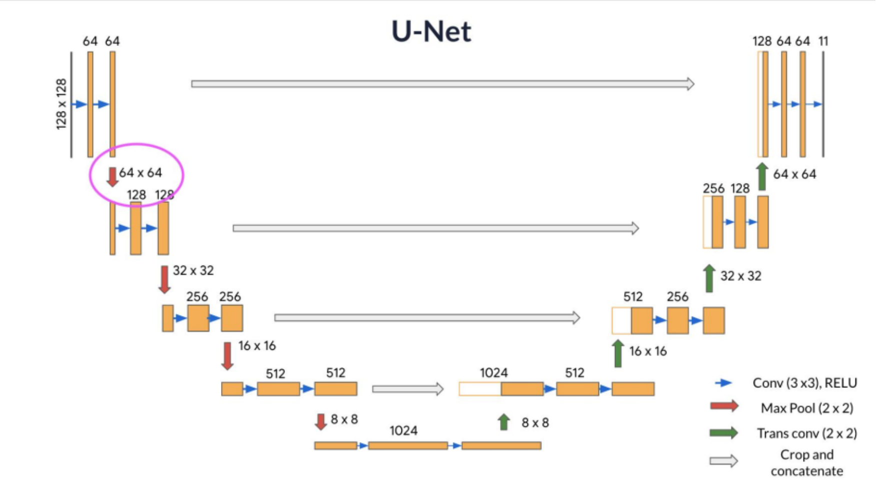

Fig. 1: *UNet Models Architecture*

Compared to traditional methods, deep learning provides a deeper understanding of image content by automatically learning complex features. This not only improves accuracy but also helps models better adapt to the diversity in real-world applications.

Deep learning has significantly advanced medical image segmentation. Early models like Fully Convolutional Networks (FCNs) by Long et al. (2015) [10] and DeepLab series[11] paved the way for pixel-level segmentation. Mask R-CNN[12] further enhanced the field by combining object detection with instance segmentation. More recently, Vision Transformers (ViTs)[13] have shown impressive results in medical image segmentation, leveraging the power of self-attention mechanisms. These advancements have led to significant improvements in accuracy and generalization, opening new avenues for medical image analysis.

However, deep learning requires a large amount of training data and computational resources, posing technical and cost challenges.

C. RESUNET

ResUnet is a semantic segmentation model, combines the encoder-decoder structure of U-Net with residual blocks from ResNet. The ResUnet uses Residual Units as basic building block instead of plain convolutional block. Residual units consists of two 3x3 convolutional block and a identity mapping, which connects the input and output of the Residual units. The convolutional block consists of: 1 Batch Normalization layer, 1 ReLU Activation layer and 1 Convolutional layer.

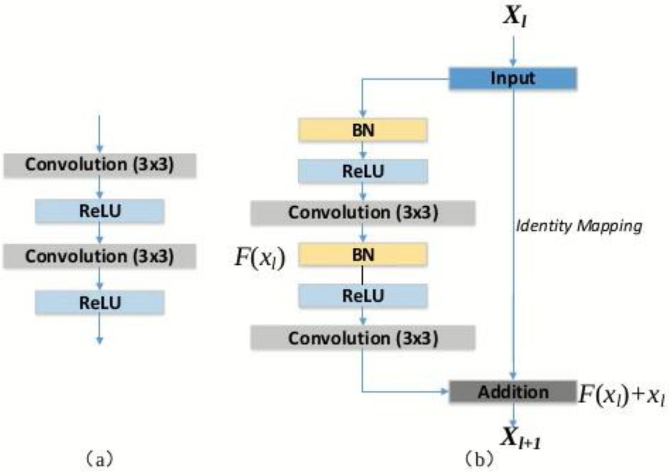

Fig. 2: *ResUnet Basic Building Block*

The Residual block in the ResUNet architecture serves a critical role in facilitating efficient feature propagation and addressing the vanishing gradient problem. Incorporating shortcut connections, allows for the smooth flow of gradients during back propagation, enhancing optimization and enabling the training of deeper networks.

The ResUnet consists of three part: Encoding, Bridge and Decoding. In Encoding units, instead of using pooling operation to downsample feature map size, a stride of 2 is applied to the first convolution block to reduce the feature map by half. Before each encoding unit, there is an up-sampling of feature maps form lower level and a concatenation with the feature maps form the corresponding encoding path. At last a 1x1 Convolution is applied with Sigmoid activation to obtain a desired segmentation map.

ResUNet stands as a formidable architecture in semantic segmentation, blending the strengths of U-Net and ResNet. Armed with this understanding, researchers can leverage ResUNet for diverse image segmentation tasks, empowering advancements in computer vision and medical imaging.

III. PROPOSED METHOD

A. RESUNET++

Debesh Jha and his team introduced the ResUnet++ architecture in 2019 at the 21st IEEE International Symposium on Multimedia (ISM). ResUnet++, a substantial advancement in medical image segmentation, offers enhanced accuracy and robustness compared to its predecessors. Its innovative approach has garnered widespread recognition, making it a popular choice for various medical image analysis applications.

Rooted in the ResUNet architecture, ResUnet++ combines the strengths of deep residual learning and U-Net .It builds upon the U-Net architecture and incorporates several key components to enhance performance. It was proposed to address certain limitations of U-Net and further enhance the accuracy and efficiency of medical image segmentation. These components can be seen in Figure 1 (Image taken from [3]):

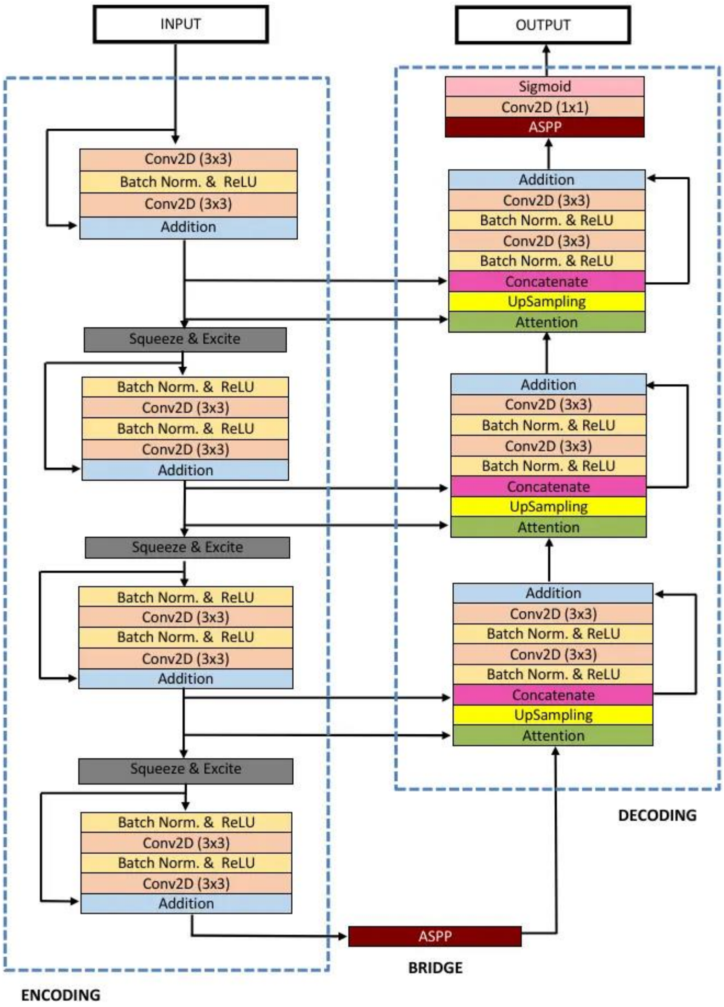

Fig. 3: *Block diagram of the proposed ResUnet++ architecture*

The proposed ResUNet++ architecture, is based on the Deep Residual U-Net (ResUNet, also takes advantage of the residual blocks, the squeeze and excitation block, ASPP, and the attention block. The key features of ResUnet++ will be discuss in the next part.

B. IMPLEMENTATIONS DETAIL

1. STEM BLOCK

Unlike U-Net, ResUNet++ starts with a modified stem block that reduces the spatial dimensions of the input image. It begins with a 3×3 convolution layer, followed by a batch normalization and ReLU activation function. Then, it was followed again by a 3×3 convolution layer and a shortcut connection consisting of a 1×1 convolution layer with batch normalization. At last, it is followed by the squeeze and excitation attention mechanism to improve the features.

2. RESIDUAL BLOCK

Residual connections are employed in the encoder path. These connections bypass one or more layers and help in addressing the vanishing gradient problem, leading to deeper networks.

3. SQUEEZE-AND-EXCITATION UNIT

Squeeze-and-Excitation (SE) units are a type of attention mechanism introduced to improve the performance of deep neural networks. To improve their channel interdependencies, the Squeeze and Excitation Network introduces a novel channel-wise attention mechanism for CNNs (Convolutional Neural Networks), called Squeeze and Excitation Block. The Squeeze and excitation blocks provide dynamic channel-wise recalibration, enhancing the model's ability to focus on significant features. The structure of the block is shown in Figure 2:

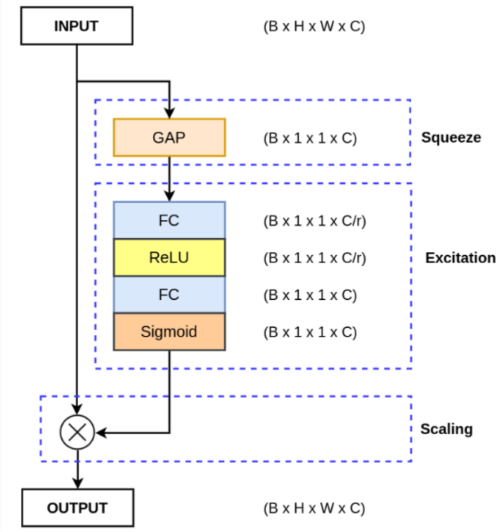

Fig. 4: *Squeeze-and-Excitation (SE) Block*

The network adds a parameter that re-weights each channel accordingly to become more sensitive to significant features while ignoring irrelevant ones. The convolution operator generates a feature map with a different number of channels, treating all the channels equally. This means that every single channel is equally important, and this may not be the best way. The Squeeze and Excitation attention mechanism adds a parameter to each channel that re-scales them independently.

4. ATTENTION UNIT (ATTENTION GATE)

The architecture introduces attention gates to the skip connections, allowing the model to focus on specific features more prominently.

Attention mechanisms have been incorporated into various deep learning architectures to improve their performance. The attention mechanism prioritizes which regions of the neural network demand more attention. Additionally, the attention mechanism lowers the computational cost of converting each polyp image's information into a fixed-dimension vector. The attention mechanism's primary benefits are its simplicity, adaptability to varying input sizes, and ability to improve the quality of features that improve outcomes.

In ResUnet++, The attention block is used in the decoder part of ResUNet++. It provides spatial attention to features using the skip connection features from the encoder, helping to enhance feature representation. This attention block allows the network to dynamically adjust the focus on different regions of the image, potentially improving segmentation accuracy.

5. ATROUS SPATIAL PYRAMIDAL POOLING (ASPP)

The Atrous Spatial Pyramidal Pooling (ASPP) is a module introduced in DeepLabv3[16] to capture multi-scale information in images. It has been incorporated into various architectures, including ResUnet++, to improve segmentation performance.

ASPP block adds dilated convolution to the network. These dilated convolutional layers help to increase the receptive field of the convolutional kernel, thereby capturing features at different scales.

In the content of ResUnet++, ASPP is utilized in two places: first, at the end of the encoder, i.e., between the encoder and decoder network, and second, at the end of the decoder network. This enables the model to effectively process features at different scales, improving its ability to segment objects of varying sizes and shapes. To capture multi-scale contextual information, ASPP is used in the last layer before the final output. By capturing both fine-grained details and broader contextual information, ASPP helps ResUnet++ to achieve more accurate and robust segmentation results.

6. DECODER PATH

The decoder path of ResUNet++ employs up-convolution followed by a series of convolutions and is equipped with long-range skip connections to gather multi-scale contextual information. It begins with an attention block, followed by an upsampling block. Next, the upsampled feature is concatenated with the feature from the encoder, i.e., a skip connection. At last, we have a ResNet block.

The decoder is used for upsampling the feature from the previous block and learning how to generate the semantic feature representation to generate a required segmentation mask.

IV. EXPERIMENTS AND RESULTS

A. DATASET

We use 2 datasets in this project: Kvasir-SEG[14] and CVC-ClinicDB [15]. The Kvasir-SEG dataset, a comprehensive resource for polyp segmentation, offers a collection of 1000 colonoscopy images meticulously annotated with polyp regions. Its diverse polyp morphology, challenging image quality, and large size make it an ideal benchmark for evaluating segmentation models. In another hand, CVC-ClinicDB, a curated dataset for polyp segmentation, offers 612 high-quality colonoscopy images. Its focus on real-world clinical scenarios and superior image quality make it valuable for training and evaluating segmentation models. The dataset was partitioned into 80% for training, 10% for validation, and 10% for testing, ensuring a comprehensive evaluation of the model's performance.

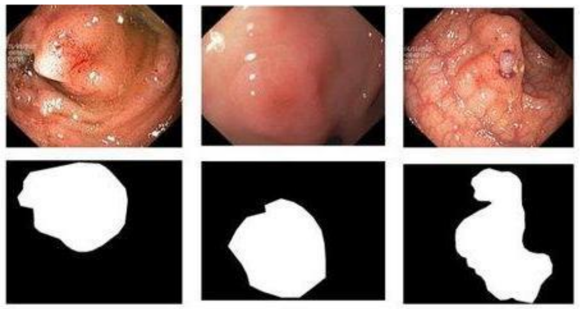

Fig. 3: *Kvasir-SEG*

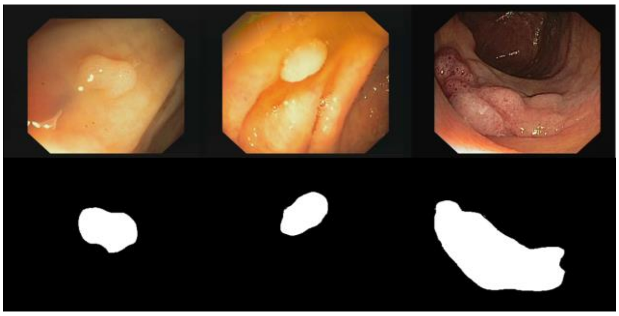

Fig. 4: *CVC-ClinicDB*

Kvasir-SEG and CVC-ClinicDB are valuable datasets for polyp segmentation research, offering diverse collections of colonoscopy images with annotated polyp regions. Researchers can choose between these datasets based on their specific research needs and the type of challenges they want to address.

B. EXPERIMENTAL SETTINGS AND EVALUATION CRITERIA

1. EXPERIMENTAL SETTINGS

a) Image Preprocessing:

To standardize the input data, all images are resized to a uniform resolution of 352x352 pixels. To normalize pixel values and scale them to a range between 0 and 1, each pixel's value in the images is divided by 255. The masks are transformed into a binary format, where a value of 1 indicates polyp regions, and a value of 0 represents the background.

b) Loss function:

The Dice Loss function was used to train the model to better segment polyp regions in medical images. By minimizing the Dice Loss, the model learns to produce predicted masks that closely align with the ground truth masks, indicating more accurate segmentation of polyp boundaries. . Dice Loss is defined as 1 minus the Dice Coefficient, follow the formula:

$$\text{Dice Loss} = 1 - \frac{2 \times |Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \#(1)$$

where Y is the actual label and $\hat{Y}$ is the model's predicted label.

2. EVALUATION CRITERIA

The evaluation criteria for this paper adhere to the content presented in the previous chapters. We utilize Dice Similarity Coefficient (DSC), mean Intersection over Union (mIOU) as the primary metrics to measure the model's performance. These metrics evaluate the model's performance by comparing the predicted segmentation map to the ground truth annotation.

The mean Intersection over Union (mIOU) calculates the ratio of the intersection to the union of the predicted and ground truth regions:

$$\text{mIoU} = \frac{1}{N}\sum_{i=1}^{N} \frac{|Y_i \cap \hat{Y}_i|}{|Y_i \cup \hat{Y}_i|} \#(2)$$

where N is the number of classes, Y_i is the predicted segmentation and $\hat{Y}_i$ is the ground truth segmentation for class i.

The Dice Similarity Coefficient, another important metric, is primarily used to assess the similarity between two samples:

$$\text{DSC} = \frac{2 \times |Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \#(3)$$

where $|Y \cap \hat{Y}|$ denotes the number of pixels correctly predicted as foreground (true positives), while $|Y|$ and $|\hat{Y}|$ represent the total number of pixels in the predicted and ground truth foreground, respectively.

C. RESULTS

1. RESULTS FROM THE PREVIOUS EXPERIMENTS)

Table 1. *Result comparison table of the two datasets CVC-ClinicDB and Kvasir-SEG*

Model	**CVC-ClinicDB**		**Kvasir-SEG**	
	DSC	**mIOU**	**DSC**	**mIOU**
UNet [10]	0.710	0.627	0.818	0.746
ResUnet	0.726	0.670	0.787	0.777
ResUnet++	0.763	0.701	0.813	0.793

Table 1 presents the results of the previous study of the ResUnet model for polyp segmentation. The experiment include the basic UNet, ResUnet and ResUnet++. Each model version is evaluated based on two main metrics: the Dice Similarity Coefficient (DSC) and the Mean Intersection over Union (mIOU) across 2 datasets: CVC-ClinicDB and Kvasir-SEG:

Compared to other U-Net variants, model demonstrates strong and competitive performance:

- Compared to U-Net, ResUnet++ on the CVC-ClinicDB dataset increased the Dice Coefficient from 0.710 to 0.763 and mIOU from 0.627 to 0.701, showing significant improvements of 7.46% and 11.79%, respectively.
- On the Kvasir-SEG dataset, ResUnet++ increased the mIOU from 0.746 to 0.793, improving by 6.3% respectively.

- Compared to ResUnet, ResUnet++ on the CVC-ClinicDB dataset increased the Dice Coefficient from 0.726 to 0.763 and mIOU from 0.670 to 0.701, showing significant improvements of 5,1% and 4.63%, respectively.
- On the Kvasir-SEG dataset, ResUnet++ increased the Dice Coefficient from 0.787 to 0.813 and the mIOU from 0.777 to 0.793, improving by 3.3% and 2.06%, respectively.

Overall, ResUnet++ demonstrates a significant improvement in medical image segmentation performance compared to its predecessors. Its ability to capture intricate details and long-range dependencies, combined with its efficient learning capabilities, makes it a powerful tool for a variety of medical image analysis tasks.

2. COMPARISON WITH U-NET BASED MODEL

To further illustrate the advancements in polyp segmentation techniques, we conducted a comprehensive comparison of various UNet models in ***Table 2***. This comparison includes the basic UNet [4], UNet++ [16], Attention UNet [17], Double UNet [18] and SwinUnet [19].

Table 2*. Result comparison table of improved versions of the U-Net model*

Model	**CVC-ClinicDB**		**Kvasir-SEG**	
	DSC	**mIOU**	**DSC**	**mIOU**
UNet [4]	0.710	0.627	0.818	0.746
UNet++ [14]	0.794	0.729	0.821	0.743
Attention U-Net [15]	0.850	0.789	0.814	0.730
Dou UNet [17]	0.878	0.822	0.894	0.833
SwinUnet [18]	0.899	0.836	0.896	0.835
ResUnet++	**0.763**	**0.701**	**0.813**	**0.793**

Compared to other U-Net variants, model demonstrates strong and competitive performance:

- For the CVC-ClinicDB dataset, ResUnet++ performs slightly lower on the CVC-ClinicDB dataset, achieving a DSC of 0.763 and an mIOU of 0.701. This suggests that the model may be more sensitive to variations in image quality and data distribution.
- For the Kvasir-SEG dataset, ResUnet++ achieves a DSC of 0.813 and an mIOU of 0.793 on the Kvasir-SEG dataset. This performance is comparable to more complex models like SwinUnet, indicating that ResUnet++ can effectively handle challenging segmentation tasks.

ResUnet++ generally performs well, especially on the Kvasir-SEG dataset, where it achieves comparable results to more complex models like SwinUnet. However, it exhibits lower performance on the CVC-ClinicDB dataset compared to some other models.

V. CONCLUSION

In this study, we successfully implemented and trained the ResUnet++ architecture for polyp segmentation. ResUnet++ provides multiple benefits for medical image segmentation. Its hierarchical architecture, combined with techniques like residual blocks, squeeze-and-excitation blocks, Atrous Spatial Pyramid Pooling, and attention mechanisms, enables it to efficiently learn features at multiple scales and capture long-range spatial dependencies. This leads to improved accuracy and precise segmentation, especially for intricate and complex medical images.. This highlights the potential of deep learning models for advancing polyp segmentation and ultimately improving early detection and diagnosis of colorectal cancer.

VI. REFERENCES

[1] Maksimova P.E., Golubinskaya E.P., Seferov B.D., Zyablitskaya E.Yu. (2023) Colorectal cancer: epidemiology, carcinogenesis, molecular subtypes and cellular mechanisms of therapy resistance (analytical review). Koloproktologia, 22(2):160-171.

[2] J. H. Bond, "Polyp guideline: diagnosis, treatment, and surveillance for patients with nonfamilial colorectal polyps," Annals of Internal Medicine, vol. 119, no. 8, pp. 836-843, 1993.

[3] Jha, Debesh & Smedsrud, Pia & Riegler, Michael & Johansen, Dag & de Lange, Thomas & Halvorsen, Pål & Dagenborg, Håvard. (2019). ResUNet++: An Advanced Architecture for Medical Image Segmentation.

[4] Zhang, Zhengxin & Liu, Qingjie. (2017). Road Extraction by Deep Residual U-Net. IEEE Geoscience and Remote Sensing Letters. PP.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of IEEE conference on computer vision and pattern recognition (CVPR), 2016, pp. 770–778.

[6] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of IEEE conference on computer vision and pattern recognition (CVPR), 2018, pp. 7132–7141.

[7] K. He., X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.

[8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 4, pp. 834–848, 2018.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in Proc. Med. Image Comput. Comput.-Assist. Intervent. – MICCAI 2015: 18th Int. Conf., Munich, Germany, Oct. 5-9, 2015, pp. 234-241.

[10] [5]J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 3431-3440.

[11] [6]Y. Chen, G. He, R. Yin, K. Zheng, and G. Wang, "Comparative study of marine ranching recognition in multi-temporal high-resolution remote sensing images based on DeepLab-v3+ and U-Net," Remote Sensing, vol. 14, no. 22, p. 5654, 2022.

[12] [7]K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 2961-2969.

[13] [8]J. M. J. Valanarasu, P. Oza, I. Hacihaliloglu, and V. M. Patel, "Medical transformer: Gated axial-attention for medical image segmentation," in Proc. Med. Image Comput. Comput.-Assist. Intervent. – MICCAI 2021: 24th Int. Conf., Strasbourg, France, Sep. 27-Oct. 1, 2021, pp. 36-46.Chen, Liang-Chieh & Papandreou, George & Schroff, Florian & Adam, Hartwig. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation.

[14] D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. De Lange, D. Johansen, and H. D. Johansen, "Kvasir-seg: A segmented polyp dataset," in Proc. MultiMedia Modeling: 26th Int. Conf., MMM 2020, Daejeon, South Korea, Jan. 5-8, 2020, pp. 451-462.

[15] C. TS and B. N. Jagadale, "Comparative Analysis of U-Net and DeepLab for Automatic Polyp Segmentation in Colonoscopic Frames Using CVC-ClinicDB Dataset," 2023.

[16] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4 (pp. 3-11). Springer International Publishing.

[17] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, et al., "Attention U-Net: Learning where to look for the pancreas," arXiv preprint arXiv:1804.03999, 2018.

[18] 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 Proc. 2020 IEEE 33rd Int. Symp. Comput.-Based Med. Syst. (CBMS), 2020, pp. 558-564.

[19] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-Unet: Unet-like pure transformer for medical image segmentation," in Proc. Eur. Conf. Comput. Vis., 2022, pp. 205-218.

PHÂN ĐOẠN ẢNH POLYP BẰNG KỸ THUẬT HỌC SÂU VỚI KIẾN TRÚC RESUNET++

Tran Nguyen Quynh Tram

Tóm tắt — Nghiên cứu này trình bày một phương pháp nhận dạng phân đoạn polyp mới sử dụng mô hình mạng ResUnet++. Sử dụng 2 bộ dữ liệu Kvasir-SEG và CVC-ClinicDB với mô hình mạng ResUnet++ cho kết quả vượt trội hơn đáng kể so với UNet và ResUnet truyền thống. Kiến trúc phân cấp của ResUnet++ , kết hợp các kỹ thuật như residual block, queeze and excite blocks, Atrous Spatial Pyramid Pooling và cơ chế chú ý, cho phép ResUnet++ học hiệu quả các tính năng ở nhiều thang đo và nắm bắt các phụ thuộc không gian tầm xa. Điều này dẫn đến độ chính xác và phân đoạn chính xác được cải thiện, đặc biệt là đối với các hình ảnh y tế phức tạp và tinh vi, giúp phát hiện ung thư đại tràng sớm.

Từ khóa — Phân đoạn hình ảnh, nội soi đại tràng, học sâu, thị giác máy tính, tin học y tế.

Tran Nguyen Quynh Tram (**https://scholar.google.com/citations?user=UGueQD0AAAAJ&hl=vi**) received a B.S. degree in information system from the HCM City University of Foreign Language and Information Technology, Vietnam, in 2005, and an M.S. degree in information system management from International University, Vietnam National University at HCMC, Vietnam, in 2017. She graduated with an M.S. degree from the Department of Artificial Intelligence, at Chonnam National University, South Korea in 2019. From 2005 to 2009, she was a Web Programmer Freelancer. Then, she was a Lecturer with the Faculty of Information Technology, Vinatex Economic-Technical College of HCM, Vietnam, for a period of seven years. Next, she moved to the Faculty of Information Technology, HCM University of Foreign Languages and Information Technology, in a Lecturer Role. Her research interests include pattern recognition, deep learning, and computer vision.