XIII Seminário de Extensão e Inovação
XXVIII Seminário de Iniciação Científica e Tecnológica da UTFPR
Ciência e Tecnologia na era da Inteligência Artificial: Desdobramentos no Ensino Pesquisa e Extensão
20 a 23 de novembro de 2023 - *Campus Ponta Grossa, PR*
SEI-SICITE 2023

Segmentaçao de Polipos em imagens de colonoscopia para detecção e diagnostico de câncer colorretal

Polyp segmentation in colonoscopy images for detection and diagnosis of Colorectal cancer

Monique Borges Seixas, João Pedro Busnardo de Souza, Murilo Moreira Mello, Jader Tavares Lima, Vinicius Sylvestre Simm, Mateus Dalla Costa, Wesley Pacheco, Paulo Victor dos Santos [2]

RESUMO

Você deve estruturar o resumo de modo a descrever o seu trabalho com simplicidade e objetividade. Tenha sempre em mente que o resumo é a apresentação do seu trabalho a leitores que provavelmente não estão familiarizados com alguns termos e expressões que podem ser comuns em sua área de estudo. O texto do resumo deve ser apresentado em um único parágrafo, sem recuo na primeira linha. A fonte usada deve ser Arial, com 10 pontos de tamanho e alinhamento justificado. O resumo deve ter entre 150 e 200 palavras: por exemplo, este texto possui exatamente 161 palavras. Na medida do possível, evite usar fórmulas, símbolos matemáticos e referências bibliográficas em seu resumo. Se ambos os resumos ultrapassarem a página inicial do documento, não há problema. Preferencialmente, procure apresentar o objetivo, o método de pesquisa, os principais resultados e as conclusões possíveis a partir do seu trabalho. Um bom resumo informa de maneira clara e objetiva o que será apresentado no texto.
PALAVRAS-CHAVE: palavra um; palavra dois; palavra três. (em ordem alfabética).

ABSTRACT

Colorectal cancer (CRC) is a leading cause of cancer-related death worldwide, with a strong correlation to the abnormal growth of tissue known as polyps. These polyps exhibit various shapes and sizes, such as flat, elevated, or pedunculated. This paper presents a deep learning approach for polyp segmentation in colonoscopy images, based on the U-net architecture and incorporating pre-processing and post- processing techniques. The approach aims to assist healthcare professionals in the accurate and early detection and diag- nosis of CRC. The experiments achieved values of 0.962 for Precision and 0.948 for Recall, demonstrating the potential of the proposed model as a valuable tool for enhancing patient outcomes, reducing public health expenses, and improving overall colonoscopy quality.
KEYWORDS: colorectal cancer, polyp, image segmentation, colonoscopy images, U-net, deep learning, pre-processing, post- processing).

INTRODUCTION

Colorectal cancer (CRC) is currently one of the leading causes of cancer-related death worldwide, ranking as the third most common cancer in men and second in women [1]. In Brazil, the number of cases is increasing, particularly in the Southern and Southeastern regions, where the incidence is comparable to that in developed countries [2].

1 UTFPR, Ponta Grossa, Paraná, Brasil E-mail: [moniqueseixas, joaopedrosouza, murilomello, jaderlima, mariafigueiredo, heronlima, marcella]@utfpr.edu.br, smpremebida@gmail.com.

2 [UEM, Maringá, UNIDEP,Pato Branco;UFG, Goiania], Brasil E-mail: viniciussimm, mateus dc1998@hotmail.com, wesley.pacheco@ufg.br,paulo.analise@live.com.

CRC is associated with the abnormal growth of tissue called polyps, which can appear in various shapes and sizes, such as flat, elevated, or pedunculated [3]. These polyps can be classified based on their characteristics, and those identified as adenomas carry a higher risk of developing the disease. Since all CRC cases originate from polyps, accurate diagnosis and treatment are crucial [4].

It is evident that accurate and early diagnosis significantly impacts the survival rate. Patients with advanced stages of CRC have a lower life expectancy compared to those who receive an early diagnosis[5,6]. In addition, CRC poses a considerable public health burden. It is estimated that by 2030, the costs associated with CRC will reach a billion Brazilian reais. Beyond the direct healthcare expenses, there are significant losses in productivity due to early deaths and morbidity among patients, which have a direct impact not only on the healthcare system but on the country's overall economy [2]. Moreover, Kuga et al. (2023) demonstrate that despite the high performance of endoscopists, there is still room for improvement. Colonoscopy is subject to human errors, poor bowel preparation, and challenges in visualizing polyps. Leufkens et al. (2012) confirmed that 22-28% of polyps are missed in patients [6]. Therefore, mechanisms that enhance colonoscopy quality, assisting endoscopists in the diagnosis and visualization of polyps, can increase early diagnosis rates, reduce public health expenses, and ultimately improve patient life expectation [4]. In this scenario, Deep Learning (DL) algorithms can be a useful tool to assist endoscopists in the diagnosis and visualization of polyps due to its higher performance in image segmentation and classification.

A possible approach for segmentation of medical images, specifically for colonoscopy, are the convolutional neural networks, among them, we can highlight the U-net. Introduced in a 2015 paper by Ronneberger et al. [7], U-Net features a distinctive design that incorporates skip connections between the encoding and decoding portions of the network. This design enables U-Net to capture fine details and shapes in biomedical images more effectively while also handling small training datasets more efficiently than other architectures. Besides, several studies showcase the potential for further modifications and improvements to the U-Net architecture to better suit the needs of specific applications, such as colorectal polyp detection.[8,9,10].

In this paper we analyze the performance of a U-net architecture to segment polyps in colonoscopy images in order to assist health professionals in the detection and diagnosis of CRC.

METHODOLOGY

We employ a U-net architecture, implemented in Python, for image segmentation and classification. The model's efficiency is enhanced using pre- and post-processing techniques which will be further explained in this section.

For model training, validation, and testing, we use two public colonoscopy image databases: CVC-ClinicDB and Kvasir-seg. CVC-ClinicDB contains 612 images and masks, and was used in the MICCAI 2015 Sub-Challenge on Automatic Polyp Detection [11]. Kvasir-seg comprises 1,000 images and masks, designed for advancing segmentation and classification methods [12].

The segmentation model follows the U-net architecture patterns proposed in [7]. The input image has a resolution of 256x256 pixels with 3 color channels.

The model requires batches of image pairs to train: the original image and its respective mask, where the polyp region is represented by white pixels. The pair must also have the same resolution.

The cost function used to train the model is "Dice Loss" with Adam as optimizer. The Dice Loss measures the similarity between the mask and the predicted mask, evaluating the overlap of images. This measure achieves its best value when both masks have the exactly same region activated. Furthermore, the Dice Loss reflects both size and localization agreement [5].

The Adam optimizer is a strategy for stochastic optimization that makes use of only first-order gradients and has low memory requirements. The technique uses estimates of the first and second moments of the gradients to calculate customized learning rates for each parameter [5].

In addition, callback options are inserted to reduce the training time and select the model that best performed on the validation set. The first callback from TensorFlow, named CSVLogger, is used to store information about each epoch into a CSV file. Later, the history data is used to predict overfit and convergence of both models. The second callback from TensorFlow, ModelCheckpoint, saves the best model using a specific criteria: the model associated with the epoch with lowest validation loss.

The colonoscope, an instrument used in colonoscopy procedures, is a long flexible tube with a diameter of around 1.3 cm, equipped with a tiny camera and a light source at its tip. The inner linings of the intestines, known as the mucosa, can be somewhat glossy due to the presence of mucus, fluids, and other secretions. This glossy surface can cause the light from the colonoscope to scatter and create specular highlights. These specular highlights can distort or hide information in the image that is essential for the segmentation process. As such, it is essential to preprocess the images, removing specular highlights and other artifacts. To do so, the following steps are employed:

- Convert the image to grayscale and calculate a threshold based on the median value of the image.
- Create a binary mask by setting pixels greater than the threshold to 1 and the rest to 0.
- Dilate the binary mask to cover the edges of the highlights.
- Inpaint the image using information from the neighboring region.

This technique also removes letterings from the colonoscopy images present in the Kvasir dataset. By doing so, we improve the quality of the colonoscopy images and enhance the polyp segmentation algorithm.

Moreover, in the pre-processing stage, the proposed U-net model takes 256x256 images as input, so the images were cropped and resized to a 256x256 format, ensuring that the segmentation model receives inputs of a consistent size without losing information regarding polyps. Initially, images and their corresponding masks are loaded, and the square bounding box coordinates for the binary mask are found. Both images are cropped using these coordinates.

In the resizing step, all the images are resized to the input size of the segmentation model using OpenCV and then converted to JPEG format. Through cropping and resizing, the model receives consistent input shapes, allowing it to focus on regions containing polyps and achieve accurate segmentation results.

The post-processing step utilized the morphologyEx method, which is publicly available in the OpenCV library. This method involves performing two morphological operations in sequence: erosion followed by dilation. This sequence is commonly used to reduce noise in images, particularly in the context of binary masks.

In this case, the post-processing operation was applied to reduce noise in the generated masks. The model sometimes produced loose positive pixels scattered randomly across the masks, which did not carry any meaningful information and could be removed without affecting the overall accuracy of the segmentation. By applying the erosion-dilation sequence, these isolated positive pixels were effectively eliminated, resulting in cleaner, more accurate segmentation masks. This, in turn, led to improved polyp detection and better overall performance of the segmentation pipeline algorithm.

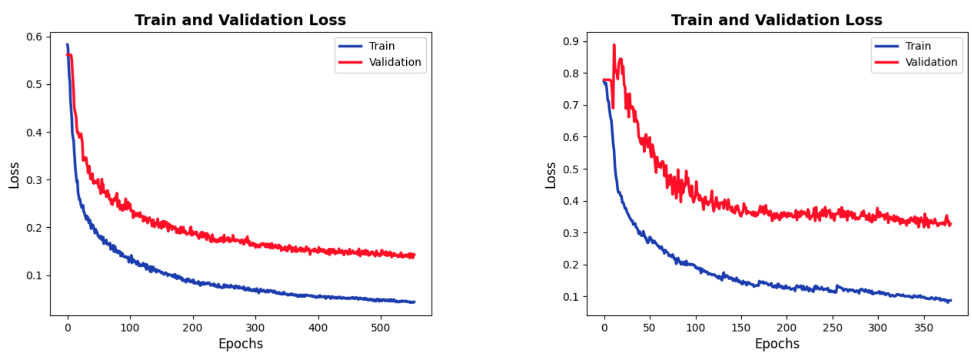
Two different models are trained using images subjected to this process: the first model is trained with the full dataset without any pre-processing, while the second model is trained with all images, excluding the cropped data, with pre-processing applied.

Several metrics are used to evaluate the model: Structural Similarity Index Method (SSIM) [13], Jaccard Index [14], Precision and Recall. The Intersection Over Union (IoU) was not used as an evaluation metric once it was applied as the cost function.

RESULTS

We trained two U-net models on CVC-ClinicDB and Kvasir-seg for colonoscopy image segmentation. The first model used un-preprocessed, cropped images, and the second used preprocessed, uncropped images. Both models were evaluated using metrics such as SSIM, Jaccard Index, Precision, and Recall (Table 1).

Both models showed overfitting, with the second model being more susceptible, possibly due to noise introduced by preprocessing. The training losses for both models decreased consistently across epochs until no further improvement was observed in the validation set (Figs. 1(a) and 1(b)).



(a) (b)

Figura 1 – (a) Training loss without pre-processing and cropped images and (b) Training loss with pre-processing and no cropped images

Post-processing improved the visual quality of the generated masks but had mixed effects on the metrics. Specifically, the first model saw minor improvements in SSIM and Precision but declines in Jaccard and Recall. Conversely, the second model's metrics worsened across the board (Table 1). The visual enhancements were achieved through erosion-dilation using the morphologyEx method, resulting in cleaner segmentation masks (Fig. 2).

Tabela 1 – Results of both models with and without post-processing

pre-processing	Cropped Images	Post-Processing	SSIM	Jaccard	Precision	Recall
Yes	No	No	0.930	0.821	0.959	0.852
Yes	No	Yes	**0.932**	0.819	**0.962**	0.848
No	Yes	No	0.925	**0.837**	0.879	**0.948**
No	Yes	Yes	0.886	0.738	0.850	0.876

Figura 2 – Comparison between masks from each model and the original mask

CONCLUSION

In this study, we introduced a U-net-based model for polyp image segmentation in colorectal cancer detection, supplemented by effective post-processing techniques. The model serves as a valuable tool for colonoscopists, aiding in early and accurate diagnosis while improving adenoma detection rates. This has implications for enhancing patient outcomes and reducing healthcare costs.

However, the model faces limitations like overfitting. Future work should focus on enlarging the training dataset and refining preprocessing methods. Exploring architectural modifications and transfer learning could also enhance generalization.

REFERENCES

1. World Health Organization, **"Cancer statistics"**, *WHO*, 2020.
2. Instituto Nacional de Câncer, **"Estimativa 2023 - Incidência de câncer no Brasil"**, *INCA*, 2023.
3. Universidade Federal do Rio Grande do Sul, **"TelessaúdeRS - Pólipos Colorretais"**, *UFRGS*, 2022.

4. N. C. Thanh, T. Q. Long, “**CRF-EfficientUNet: An improved UNet framework for polyp segmentation in colonoscopy images**”, *IEEE Access*, vol. 9, pp. 156987–157001, 2021.

5. R. A. Rostirolla, J. C. Pereira-Lima, C. R. Teixeira, A. W. Schuch, C. Perazzoli, C. Saul, “**Desenvolvimento de neoplasias/adenomas avançados colorretais**”, *Arq. Gastroenterol.*, vol. 46, no. 3, pp. 167–172, 2009.

6. A. M. Leufkens, M. G. van Oijen, F. P. Vleggaar, P. D. Siersema, “**Factors influencing the miss rate of polyps**”, *Endoscopy*, vol. 44, no. 5, pp. 470–475, 2012.

7. O. Ronneberger, P. Fischer, and T. Brox, “**U-Net: Convolutional Networks for Biomedical Image Segmentation**”, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W.M. Wells, and A.F. Frangi, Eds. Cham: Springer, 2015, pp. 234–241.

8. Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, and J. Liang, “**UNet++: A Nested U-Net Architecture for Medical Image Segmentation**”, in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, D. Stoyanov, A. Taylor, T. Simpson, A. Noble, Z. Chen, Y. Xu, and G. Yang, Eds. Cham: Springer, 2018, pp. 3–11.

9. X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, “**H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes**”, *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.

10. F. Milletari, N. Navab, and S.-A. Ahmadi, “**V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation**”, *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, 2016.

11. Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., Vilariño, F., “**WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians**”, *Computerized Medical Imaging and Graphics*, 43, 99-111, 2015.

12. Debesh Jha, Pia H. Smedsrud, Michael A. Riegler, Pål Halvorsen, Thomas de Lange, Dag Johansen, and Håvard D. Johansen, “**Kvasir-SEG: A Segmented Polyp Dataset**”, in *MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II*. Springer-Verlag, Berlin, Heidelberg, 451–462, 2020.

13. A. Hore & D. Ziou, “**Image Quality Metrics: PSNR vs. SSIM**”, *2010 20th International Conference on Pattern Recognition*, 2010.

14. R. Real & J. M. Vargas, “**The Probabilistic Basis of Jaccard’s Index of Similarity**”, *Systematic Biology*, 45(3), 380-385, 1996.