Article

UPolySeg: A U-Net-Based Polyp Segmentation Network Using Colonoscopy Images

Subhashree Mohapatra ¹, Girish Kumar Pati ² [ID], Manohar Mishra ³,* [ID] and Tripti Swarnkar ⁴

1   Department of Computer Science & Engineering, Siksha O Anusandhan Deemed to Be University, Bhubaneswar 751030, India
2   Department of Gastroenterology, Institute of Medical Science and SUM Hospital, Bhubaneswar 751003, India
3   Department of Electrical and Electronics Engineering, Siksha O Anusandhan Deemed to Be University, Bhubaneswar 751030, India
4   Department of Computer Application, Siksha O Anusandhan Deemed to Be University, Bhubaneswar 751030, India
*   Correspondence: manoharmishra@soa.ac.in

Abstract: Colonoscopy is a gold standard procedure for tracking the lower gastrointestinal region. A colorectal polyp is one such condition that is detected through colonoscopy. Even though technical advancements have improved the early detection of colorectal polyps, there is still a high percentage of misses due to various factors. Polyp segmentation can play a significant role in the detection of polyps at the early stage and can thus help reduce the severity of the disease. In this work, the authors implemented several image pre-processing techniques such as coherence transport and contrast limited adaptive histogram equalization (CLAHE) to handle different challenges in colonoscopy images. The processed image was then segmented into a polyp and normal pixel using a U-Net-based deep learning segmentation model named UPolySeg. The main framework of UPolySeg has an encoder–decoder section with feature concatenation in the same layer as the encoder–decoder along with the use of dilated convolution. The model was experimentally verified using the publicly available Kvasir-SEG dataset, which gives a global accuracy of 96.77%, a dice coefficient of 96.86%, an IoU of 87.91%, a recall of 95.57%, and a precision of 92.29%. The new framework for the polyp segmentation implementing UPolySeg improved the performance by 1.93% compared with prior work.

Keywords: segmentation; polyp; U-Net; colonoscopy; deep learning

1. Introduction

1.1. Motivation and Incitement

Colorectal cancer (CC) is a major concern in the modern era, where it ranks second in worldwide mortality [1]. It is also the third most common cancer in both genders [2]. Figure 1 shows the cases of colorectal cancer worldwide considering both genders [2]. Polyps are considered an initial sign of colorectal cancer and need to be detected at the early stage. Colorectal polyps can be categorized into various types such as adenoma, serrated adenoma polyp, hyperplastic polyp, and inflammatory polyp sessile [3]. Each category has a different level of risk of developing into CC [4]. Inflammatory and hyperplastic polyps have the lowest risk of developing into CC, while adenoma and serrated adenoma polyps have a high risk of developing into CC. Detection at the early stage is a very crucial task that is carried out by experienced gastroenterologists through colonoscopy [5]. Even though colonoscopy is very effective, it comes with its own limitations. In many cases, polyps can be missed by professionals due to technical or professional errors. This can be due to several factors such as quick scanning through the affected area, polyps not appearing within the visual field, or the polyp size and texture not being very specific [1]. Sometimes, neoplastic polyps can be very hard to detect, even by experts. Another limitation is that it is

a time-consuming process for gastroenterologists and is also a labor-intensive procedure [3]. For these reasons, the cost of the examination is high in high-population countries [3]. In this regard, a smart system can help practitioners reduce the polyp miss rate and thus can further reduce the severity of colorectal cancer. To be more specific, implementing an intelligent segmentation system that can segment the specific polyp region within an image will definitely increase the effectiveness of detecting polyps.
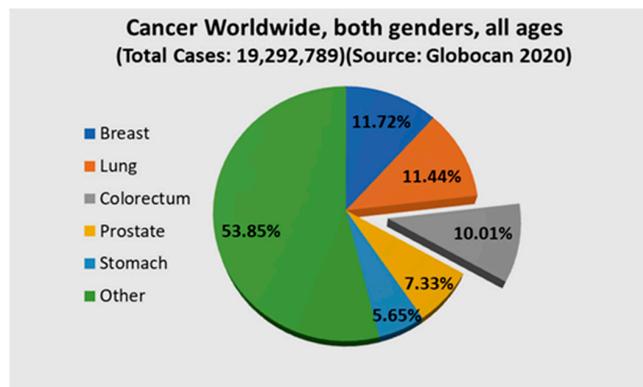

Figure 1. Chart showing the incidences of colorectum cancer worldwide.

The main challenge in the case of region-of-interest extraction in polyps is that the number, size, and shape of polyps vary widely [6]. This can be well visualized in Figure 2 [7]. To handle this, an efficient segmentation network is required to segment each type of polyp in an image. Another point is that, in the colonoscopy polyp images, there can be some artifacts such as a green or black patch signifying the placement of the colonoscope inside the body, specularity (white patch or spots due to reflection of light) [8], as well as image contrast [9]. This fact motivated the authors to take up the challenge of efficiently segmenting the polyp region in colonoscopy images by implementing various image-processing techniques to remove these artifacts and by designing an intelligent deep learning model based on U-Net [10].

Figure 2. Sample of polyp images along with their masks showing the difference in number, shape, and size of polyps. The first and third images show the original image, while the second and fourth images show the corresponding ground truths provided in the Kvasir-SEG [7] dataset.

1.2. Prior Work

There are several studies available in the literature that show the automatic segmentation of polyp. These methodologies can be broadly divided into two categories. The first approach is based on implementing machine learning techniques that use the hand-crafted feature for segmentation. The second approach is based on using deep learning techniques for polyp segmentation. Here, a summarized review is presented for previous work on polyp detection and segmentation. Yao and Summers [11] proposed a fuzzy c-mean clustering technique for segregating the polyp region using computed tomography colonography images. The fuzzy c-mean was followed by adaptive deformable models for polyp segmentation. Sánchez-González et al. [12] proposed a system to segment the polyp by considering features such as the shape, color, region, and curvature of the edges of the polyps. Yuan et al. [13] proposed automatic detection of polyps using colonoscopy videos. They considered using sparse autoencoders to extract super-pixel features along with various saliency techniques to segment polyp areas.

The authors in [14] designed and implemented two variants of a fully convolution neural network for Gastrointestinal Image ANAlysis (GIANA) polyp segmentation. Baldeon-Calisto and Lai-Yuen [15] proposed a multi-objective adaptive residual U-Net for the segmentation of medical images that can adapt to any new dataset while reducing the network size. Tomar et al. [16] designed a novel network dual decoder attention network experimented on the Kvasir-SEG [7] dataset and validated on an unseen dataset. The network achieved a dice coefficient of 0.7874, a mean intersection of union (mIoU) of 0.7010, a recall of 0.7987, and a precision of 0.8577. Zhang et al. [17] proposed a fused network for segmentation combining transformers and convolutional neural networks (CNNs) in a parallel way known as TransFuse as well as a BiFusion module that competently combines multi-level features from both branches. Zhang et al. [18] proposed an adaptive context selection-based encoder–decoder model for polyp segmentation using the Kvasir-seg and EndoScene datasets. The network comprises different modules such as local context attention, a global context module, and an adaptive selection module.

The authors in [19] designed a CNN named HarDNet-MSEG for the segmentation of polyps using five different datasets, one of which was the Kvasir-SEG polyp dataset, and for this data, the model delivered a mean dice score of 0.904 for 86.7 fps. An encoder–decoder module-dependent deep neural network framework was proposed by Mahmud et al. [20] for polyp segmentation using four different datasets. The network was named PolypSegNet and aims to handle various issues of traditional models. The authors in [6] benchmarked various state-of-the-art techniques using the Kvasir-SEG dataset on ColonSegNet for polyp detection, localization, and segmentation. The model achieved a competitive dice coefficient of 0.8206 and a best average speed of 182.38 frames per second for the segmentation task using 512×512 images. The network was implemented using PyTorch, and the model was trained using NVIDIA Quadro RTX 6000 hardware. The authors in [6] did not consider any image pre-processing techniques to enhance the data. As stated earlier, colonoscopy images can have some issues such as specularity [8], saturation, contrast, and a few others for which some pre-processing steps can be incorporated. Additionally, the hyperparameters of the deep network were not tuned in [6].

1.3. Major Contribution

The above-mentioned literature review provides a detailed analysis of related recent approaches proposed for polyp segmentation. The major outcomes, challenges, and research gaps are clearly highlighted and discussed. In order to deal with the above-mentioned challenges, this study proposed a new model comprising an image pre-processing unit for handling image issues and a deep learning-based network for polyp segmentation.

The main contribution of this work can be stated as follows.

- An image pre-processing module that pre-processes the input image in three general steps was designed. In the first step, the image was resized and then a coherent transport module was used to remove specularity in the image. The final step included the contrast enhancement module, which used the contrast limited adaptive histogram equalization (CLAHE) technique to enhance the image.
- A U-Net model (UPolySeg) was designed from scratch by implementing some advanced modules within the architecture for segmentation of the polyp using the Kvasir-SEG dataset.
- The hyperparameters for the UPolySeg were selected after extensive experimental work.
- To justify the effectiveness of the UPolySeg, it was compared with a similar model, ColonSegNet, which was designed for the segmentation of polyps.

The authors hypothesize that handling the specularity and contrast issue of colonoscopy images through coherence transport and CLAHE, respectively, will help the segmentation network segment out each polyp more effectively. Again, improving the U-Net structure by applying advanced processing parameters and tuning the hyperparameters can help the network accomplish necessary tasks more accurately.

2. Materials and Methods

The authors designed a U–Net [10]-based polyp segmentation (UPolySeg) framework using the publicly available Kvasir-SEG dataset. Details about the dataset and the techniques used in this work are described in this section. The complete framework of the proposed model is illustrated in Figure 3.
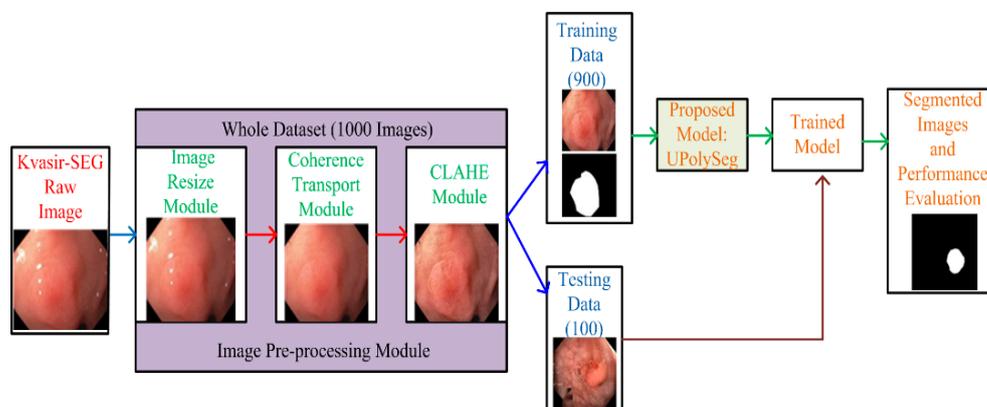

Figure 3. A complete framework of the proposed model.

2.1. Dataset

There are various datasets available [20], such as CVC-ClinicDB [21], ETIS-Larib [22], Endoscene [23], CVC-ColonDB [24], and Kvasir-SEG, that consist of different polyp images along with their ground truths. This study was carried out using the Kvasir-SEG dataset. The dataset consists of two different folders, one with the polyp images and another with the corresponding ground truth of the original polyp images. Each of these folders have 1000 images of the polyps and the ground truth mask images. The dataset also has a JavaScript Object Notation (JSON) file that consists of bounding boxes for the corresponding polyp images. The mask of the polyp region of the original polyp images was created using the Labelbox tool. The margin of polyp regions was manually constructed under the supervision of an engineer and a medical professional. The final annotation of the polyp mask was later confirmed by experienced gastroenterologists. Out of these 1000 images, 900 images were used for training and 100 were used for testing.

2.2. Image Pre-Processing

In this work, the dataset used was the publicly available Kvasir-SEG dataset consisting of colorectal polyp images captured through colonoscopy. In general, the images can be of different sizes and can have some artifacts or noise, affecting the performance of any classification or segmentation model. Then, image pre-processing was performed. In the image pre-processing process, the first step was resizing the image. As each image in the dataset was of varying sizes, the images were resized to 416×416 pixels. This size was chosen after an experimental analysis was performed; higher dimension images increased the execution time as well as the memory required. The next two steps are explained in detail in this section.

2.2.1. Specular Reflection and Patch

To remove the image artifacts, the authors used the inpainting method. One inpainting technique is the partial differential equation (PDE)-based coherence transport (CT) technique, which is a pixel-based method. CT is an efficient PDE technique that uses a first-order PDE, so iterates through each pixel only once [25]. The first step is to find a binary mask of the image using an image segmenter, where the nonzero pixels in the mask represent the region to be filled up. CT inpaints the pixels by sequentially traversing the particular pixels beginning from the boundary and moving towards the interior. The pixels

are ordered by evaluating the Euclidean distance of the pixel to the image boundary. Each ordered pixel is inpainted according to Equation (1).

$$w_i(p_j) = \frac{\sum_{x \in N_{\varepsilon, i}^<(p_j)} u(p_j, x) w_i(x)}{\sum_{x \in N_{\varepsilon, i}^<(p_j)} u(p_j, x)}, \ j = 1, 2, \ldots, m \tag{1}$$

where w_i gives the weight value of any given pixel; p_j represents ordered pixels to be inpainted; m signifies the total number of pixels to be inpainted; $u(p_j, x)$ is a non-negative weight function; and $N_{\varepsilon, i}^<(p_j)$ is a space that contains original or previously inpainted pixels, which is given by Equation (2).

$$N_{\varepsilon, i}^<(p_j) := \{p \in N_{\varepsilon, i}(p_j) : K_i(p) < K_i(p_j) \ if \ p \in \alpha_i\} \tag{2}$$

where α_i contains the pixels outside the inpainting area.

2.2.2. Contrast

To enhance the quality of the colonoscopy images, the contrast limited adaptive histogram equalization (CLAHE) technique was implemented as it is the most popular technique used for medical image enhancement [26]. The whole process of CLAHE is performed in two broad steps. In the first step, the original image is divided into multiple non-overlapping areas of almost the same size. The individual region is evaluated. After the histogram is evaluated, each region is redistributed such that the height of the histogram does not exceed the clip limit. The clip limit is set by the value α and is given by Equation (3).

$$\alpha = \frac{RC}{N}\left(1 + \frac{\beta}{100}(Sl_{max} - 1)\right) \tag{3}$$

where α represents the clip limit; RC represents the pixel value in each area; N is the number of grayscales; β represents the clip factor, which ranges from 0 to 100; and Sl_{max} is the maximum allowable slope. In this work, the R and C values are taken to be 8, and the clip limit is set to 0.002. Figure 4 illustrates a sample image after each pre-processing step.

Figure 4. The first image is the original image from the Kvasir-SEG dataset, the second image is the image after artifact removal, and the third is the image after contrast enhancement.

2.3. Deep Learning for Image Segmentation

In this work, the UPolySeg model was designed based on the U-Net [10] architecture. U-Net is a very popular deep learning network specially designed for medical image segmentation and performs better than other architectures [27]. The detailed deep learning architecture of UPolySeg is illustrated in Figure 5. The main module is the encoder–decoder connected in a U-shaped structure. Each encoder–decoder in the same layer is linked for feature concatenation. The proposed UPolySeg model has three levels of encoder–decoders. Each of the encoders in the contracting path consists of a 3×3 convolution followed by a leaky rectified linear unit (LReLU). Each contracting module is followed by a 2×2 max pooling layer for downsampling. Again, each module of the decoder in the expanding path starts with 2×2 transposed convolutions for upsampling. Then, 3×3 convolutions are followed by LReLU with 2×2 max pooling. The output of the last decoder module is sent to 1×1 convolution layers. A softmax activation unit is used for evaluating the probability

of each pixel. Finally, a pixel classification unit with dice loss is used to generate the binary mask of the image.
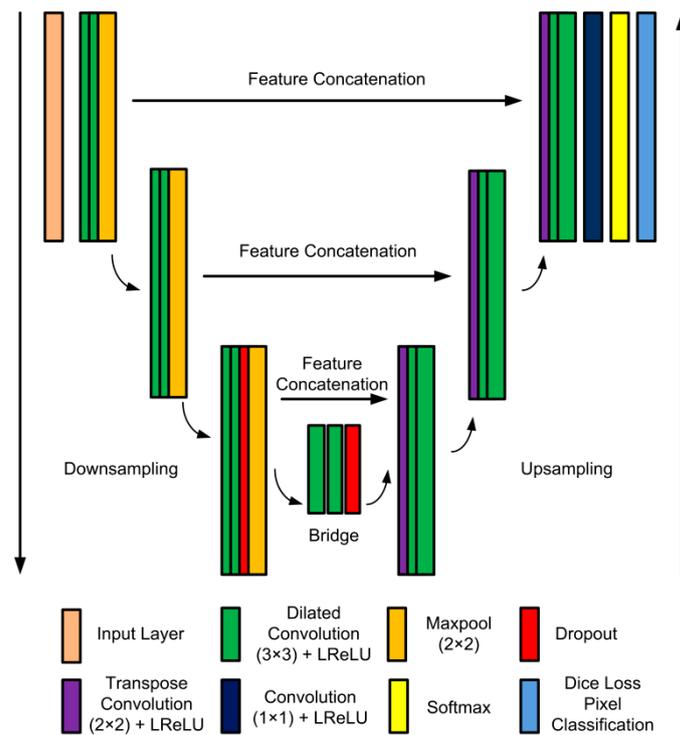

Figure 5. UPolySeg architecture.

As there is a class imbalance in the segmentation task, dice loss helps improve the condition. In this network, the convolution used is the dilated convolutions with various dilation factors [28]. Dilated convolutions are used as it does not increase the number of parameters but expands the area of the receptive field. The dilation factor controls the area of the receptive field. Again, LReLU, which is an advanced version of ReLU, is used. LReLU is represented by $g(y) = max(0.01 * y, y)$, where y is an input value. The main significance of LReLU is that it always generates an output value for both negative and positive input data. Therefore, it helps eliminate dead neurons in the network.

2.4. Performance Indicators

The network was evaluated using various parameters [6], such as global accuracy (GA), dice coefficient (DC), intersection over union (IoU), recall (R), and precision (P). The UPolySeg model was trained using different hyper parameters. Here, the global accuracy represents the proportion of correct predictions. The global accuracy is calculated using Equation (4). The intersection over union, also known as the Jaccard index, shows the proportion of overlap between the predicted value and the ground truth mask (represented in Equation (5)). The dice coefficient is quite similar to the IoU, but it double counts the intersection, as shown in Equation (6). Precision signifies the purity of a positive detection compared with the ground truth, whereas recall signifies the completeness of a positive detection compared with ground truth. Precision and recall can be evaluated using Equation (7) and Equation (8), respectively. Each of the parameters was evaluated by taking into account the true-positive (T_p), true-negative (T_n), false-positive (F_p), and false-negative (F_n) rates.

$$GA = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{4}$$

$$IoU = \frac{T_p}{T_p + F_p + F_n} \tag{5}$$

$$DC = \frac{2 \cdot T_p}{2 \cdot T_p + F_p + F_n} \tag{6}$$

$$P = \frac{T_p}{T_p + F_p} \tag{7}$$

$$R = \frac{T_p}{T_p + F_n} \tag{8}$$

3. Results

The proposed UPolySeg model was trained on a system with Intel Core 2.60 GHz i7 CPU running Windows 10 with 16 GB RAM, NVIDIA GeForce GTX 1650 GPU. All of the experiments were performed using MATLAB version 2020.

The hyperparameters were set after performing experimental work in exactly five different sets of parameters. Table 1 presents the training accuracy (TA) for different parameter sets. Here, the optimizer used was stochastic gradient descent with momentum (SGDM), the learning rate (LR) was set to 0.0001, L2regularization (L2reg) was 0.005, momentum was 0.9, and the network was trained for 50 epochs. As the network was stable and achieved an accuracy of 97.66% for the training set (Table 1), training was performed for 50 epochs. The evaluated performance measures were compared with the performance value of ColonSegNet [6].

Table 1. Training accuracy for different parameter sets.

LR	L2reg	Momentum	Epoch	TA
0.0001	-	-	20	85.23%
0.0001	-	-	40	87.82%
0.0001	0.001	0.5	40	90.87%
0.0001	0.005	0.5	40	93.02%
0.0001	0.005	0.9	50	97.66%

Figure 6 illustrates an overlay of the final segmented image along with the ground truth image for the best case. Here, the IoU obtained was 0.98, whereas the worst-case IoU achieved 0.8. Figure 7 illustrates an overlay of the final segmented image along with ground truth image for the worst case. Table 2 presents the calculated values of the evaluation parameters for UPolySeg compared with ColonSegNet. It is observed from the evaluation matrix (Table 2) that the UPolySeg model performed better than ColonSegNet. The global accuracy of UPolySeg was 96.77%, DC was 96.86%, IoU was 87.91%, recall was 95.57%, and precision was 92.29%.
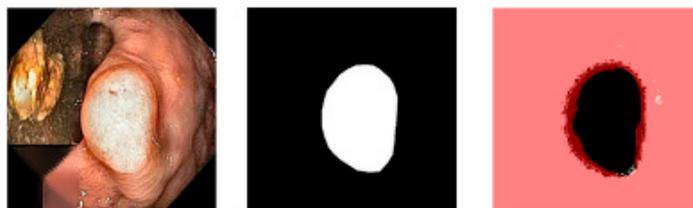

Figure 6. The first image is the original image, the second image is the ground truth, and the third is an overlay of the ground truth and the segmented image for the best case.

Table 2. Performance evaluation of UPolySeg with ColonSegNet.

Model	GA	DC	IoU	R	P
ColonSegNet	0.9493	0.8206	0.7239	0.8597	0.8435
UPolySeg (Proposed)	0.9677	0.9686	0.8791	0.9557	0.9229

Figure 7. The first image is the original image, the second image is the ground truth, and the third is an overlay of the ground truth and the segmented image for the worst case.

4. Discussion

In this work, the authors designed a framework for colonoscopy polyp image preprocessing along with polyp segmentation. Various challenges from ColonSegNet were addressed in this work. Handling the artifacts present in medical images along with incorporating various advanced options in UPolySeg helped improve the performance of the segmentation task. In the pre-processing stage, various techniques such as coherence transport and CLAHE were implemented. UPolySeg, a segmentation network that gives a better performance than a prior work such as ColonSegNet, was proposed. Here, dilated convolution was used to increase the area of the receptive field and LReLU helped remove dead neurons in the network to enhance the efficiency of the network. The proposed model obtained a global accuracy of 96.77%, a DC of 96.86%, an IoU of 87.91%, a recall of 95.57%, and a precision of 92.29%, whereas the values given by ColonSegNet were 94.93% overall accuracy, 82.06% DC, 72.39% IoU, 85.97% recall, and 84.35% precision. These results show an improvement of 1.93% in accuracy obtained by UPolySeg. The authors conclude that enhancing the input image by applying coherence transport and CLAHE before training, implementing various advanced parameters in the deep network, and tuning the hyperparameters of the network helped UPolySeg obtain a better performance. It is to be noted that the hardware environment and the hyperparameters used in ColonSegNet are different compared with that of this work of art. Therefore, this comparison could be conducted because the dataset is the same and the approach is quite similar. The deep learning model could be deployed to assist gastroenterologists and can help reduce the adenoma miss rate and detect the disease at the early stage to reduce the death rate due to colorectal cancer.

Even though the network has shown a good performance, there is still room for further exploration. In this work, the authors implemented image pre-processing techniques and tried to improve the U-Net model by applying various advanced units in the training network. Another aspect of the research that can be explored is an ablation study to determine the improvement in efficiency at each step of the study. Here, the authors have used image processing for handling the specularity and contrast of the colonoscopy images. In medical images, there can be other issues (such as noise and distortion) that need to be resolved for better classification performance. Another limitation is that the authors have focused on enhancing only the U-Net architecture for the segmentation of polyps. It would be more effective if several other deep networks could be designed and trained for such a purpose along with a comparative analysis. This will help researchers understand the strengths and weaknesses of various networks used for polyp segmentation. Some advanced optimization techniques can be implemented to tune the hyperparameters of the network. This work is carried out by taking the polyp images from a single source: the Kvasir-SEG database. In this regard, more sophisticated retrospective studies can be carried out by combining different datasets available for public research. The study can even be converted into a prospective study by taking into account a case study of patients directly from a hospital. The research scope remains open for researchers working in this domain to develop a more efficient system for the segmentation of polyps.

5. Conclusions

Even though colonoscopy can help obtain a detailed visual of an internal portion of the colon and is better at determining the presence of a polyp, the adenoma miss rate is still high. This can be reduced by considering deep learning and finding polyps by segmenting colonoscopy images. This could help professionals even determine the severity of the disease by observing the size of the polyp that is segmented out. In the literature, various state-of-the-art work has been carried out on the segmentation of polyps but few challenges have yet to be handled. The proposed framework was designed by keeping in mind unresolved challenges in ColonSegNet. UPolySeg has a pre-processing module for enhancing the image contrast and for removing specularity in coloscopy images. Additionally, some advanced options are selected and designed in the network based on the U-Net architecture. The pre-processing unit along with UPolySeg increased the performance by 1.93% compared with other work, but there is still room for improvement. Detecting different categories of polyps using deep learning techniques can be very helpful for experts to determine the level of risk for colorectal cancer. Other segmentation networks can also be implemented to evaluate the segmentation task on the Kvasir-SEG dataset or a different dataset. Fine-tuning of the network can be performed using various optimization techniques, which gives a scope for future research.

Author Contributions: S.M.: methodology, software, validation, data curation, investigation, and formal analysis; G.K.P.: conceptualization, methodology, writing—review and editing, and supervision; M.M.: writing—original draft preparation, validation, and formal analysis; T.S.: conceptualization, methodology, writing—review and editing, and supervision. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are openly available in [Kvasir-SEG] at [https://datasets.simula.no/kvasir-seg/, accessed on 17 July 2022] [7].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in the manuscript.

CC	Colorectal Cancer
CLAHE	Contrast Limited Adaptive Histogram Equalization
CT	Coherence Transport
IoU	Intersection over Union
CNN	Convolutional Neural Network
JSON	JavaScript Object Notation
PDE	Partial Differential Equation
LReLU	Leaky Rectified Linear Unit
GA	Global Accuracy
DC	Dice Coefficient
R	Recall
P	Precision
TA	Training Accuracy
SGDM	Stochastic Gradient Descent with Momentum
LR	Learning Rate
L2reg	L2regularization

References

1. Jha, D.; Smedsrud, P.H.; Johansen, D.; de Lange, T.; Johansen, H.D.; Halvorsen, P.; Riegler, M.A. A Comprehensive Study on Colorectal Polyp Segmentation with ResUNet++, Conditional Random Field and Test-Time Augmentation. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 2029–2040. [CrossRef] [PubMed]
2. Xi, Y.; Xu, P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl. Oncol.* **2021**, *14*, 101174. [CrossRef] [PubMed]
3. Xu, Y.; Ding, W.; Wang, Y.; Tan, Y.; Xi, C.; Ye, N.; Wu, D.; Xu, X. Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis. *PLoS ONE* **2021**, *16*, e0246892. [CrossRef] [PubMed]
4. Mahmud, T.; Paul, B.; Fattah, S.A. PolypSegNet: A modified encoder-decoder architecture for automated polyp segmentation from colonoscopy images. *Comput. Biol. Med.* **2020**, *128*, 104119. [CrossRef] [PubMed]
5. Jha, D.; Smedsrud, P.H.; Riegler, M.A.; Johansen, D.; De Lange, T.; Halvorsen, P.; Johansen, H.D. ResUNet++: An Advanced Architecture for Medical Image Segmentation. In Proceeding of the IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 9–11 December 2019; pp. 225–2255. [CrossRef]
6. Jha, D.; Ali, S.; Tomar, N.K.; Johansen, H.D.; Johansen, D.; Rittscher, J.; Riegler, M.A.; Halvorsen, P. Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy Using Deep Learning. *IEEE Access* **2021**, *9*, 40496–40510. [CrossRef] [PubMed]
7. Jha, D.; Smedsrud, P.H.; Riegler, M.A.; Halvorsen, P.; Lange, T.D.; Johansen, D.; Johansen, H.D. Kvasir-seg: A segmented polyp dataset. In *International Conference on Multimedia Modeling*; Springer: Cham, Switzerland, 2020; pp. 451–462.
8. Kayser, M.; Soberanis-Mukul, R.D.; Zvereva, A.M.; Klare, P.; Navab, N.; Albarqouni, S. Understanding the effects of artifacts on automated polyp detection and incorporating that knowledge via learning without forgetting. *arXiv* **2020**, arXiv:2002.02883.
9. Mohapatra, S.; Nayak, J.; Mishra, M.; Pati, G.K.; Naik, B.; Swarnkar, T. Wavelet Transform and Deep Convolutional Neural Network-Based Smart Healthcare System for Gastrointestinal Disease Detection. *Interdiscip. Sci. Comput. Life Sci.* **2021**, *13*, 212–228. [CrossRef] [PubMed]
10. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2015; pp. 234–241.
11. Yao, J.; Summers, R.M. Adaptive deformable model for colonic polyp segmentation and measurement on CT colonography. *Med. Phys.* **2007**, *34*, 1655–1664. [CrossRef] [PubMed]
12. Sánchez-González, A.; García-Zapirain, B.; Sierra-Sosa, D.; Elmaghraby, A. Automatized colon polyp segmentation via contour region analysis. *Comput. Biol. Med.* **2018**, *100*, 152–164. [CrossRef] [PubMed]
13. Yuan, Y.; Li, D.; Meng, M.Q.H. Automatic polyp detection via a novel unified bottom-up and top-down saliency approach. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1250–1260. [CrossRef] [PubMed]
14. Guo, Y.B.; Matuszewski, B. Giana polyp segmentation with fully convolutional dilation neural networks. In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Prague, Czech Republic, 25–27 February 2019; SciTePress-Science and Technology Publications: Setubal, Portugal, 2019; pp. 632–641.
15. Baldeon-Calisto, M.; Lai-Yuen, S.K. AdaResU-Net: Multiobjective adaptive convolutional neural network for medical image segmentation. *Neurocomputing* **2020**, *392*, 325–340. [CrossRef]
16. Tomar, N.K.; Jha, D.; Ali, S.; Johansen, H.D.; Johansen, D.; Riegler, M.A.; Halvorsen, P. DDANet: Dual Decoder Attention Network for Automatic Polyp Segmentation. In *International Conference on Pattern Recognition*; Springer: Cham, Switzerland, 2021; pp. 307–314. [CrossRef]
17. Zhang, Y.; Liu, H.; Hu, Q. Transfuse: Fusing transformers and CNNs for medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2021; pp. 14–24.
18. Zhang, R.; Li, G.; Li, Z.; Cui, S.; Qian, D.; Yu, Y. Adaptive context selection for polyp segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2020; pp. 253–262.
19. Huang, C.H.; Wu, H.Y.; Lin, Y.L. Hardnet-mseg: A simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps. *arXiv* **2021**, arXiv:2101.07172.
20. Patel, K.; Bur, A.M.; Wang, G. Enhanced u-net: A feature enhancement network for polyp segmentation. In Proceedings of the 2021 18th Conference on Robots and Vision (CRV), Burnaby, BC, Canada, 26–28 May 2021; pp. 181–188.
21. 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. *Comput. Med. Imaging Graph.* **2015**, *43*, 99–111. [CrossRef] [PubMed]
22. Silva, J.; Histace, A.; Romain, O.; Dray, X.; Granado, B. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *Int. J. Comput. Assist. Radiol. Surg.* **2014**, *9*, 283–293. [CrossRef] [PubMed]
23. Vázquez, D.; Bernal, J.; Sánchez, F.J.; Fernández-Esparrach, M.G.; López, A.M.; Romero, A.; Drozdzal, M.; Courville, A. A Benchmark for Endoluminal Scene Segmentation of Colonoscopy Images. *J. Healthc. Eng.* **2017**, *2017*, 4037190. [CrossRef] [PubMed]
24. Bernal, J.; Sánchez, J.; Vilariño, F. Towards automatic polyp detection with a polyp appearance model. *Pattern Recognit.* **2012**, *45*, 3166–3182. [CrossRef]
25. März, T. A well-posedness framework for inpainting based on coherence transport. *Found. Comput. Math.* **2015**, *15*, 973–1033. [CrossRef]

26. Koonsanit, K.; Thongvigitmanee, S.; Pongnapang, N.; Thajchayapong, P. Image enhancement on digital X-ray images using n-clahe. In Proceedings of the 2017 10th Biomedical Engineering International Conference (BMEiCON), Hokkaido, Japan, 31 August–2 September 2017; pp. 1–4.

27. Rundo, L.; Han, C.; Zhang, J.; Hataya, R.; Nagano, Y.; Militello, C.; Ferretti, C.; Nobile, M.S.; Tangherloni, A.; Gilardi, M.C.; et al. CNN-Based Prostate Zonal Segmentation on T2-Weighted MR Images: A Cross-Dataset Study. In *Neural Approaches to Dynamics of Signal Exchanges*; Springer: Singapore, 2020; pp. 269–280.

28. Zhang, J.; Lu, C.; Wang, J.; Wang, L.; Yue, X.-G. Concrete Cracks Detection Based on FCN with Dilated Convolution. *Appl. Sci.* **2019**, *9*, 2686. [CrossRef]