Comp448 Medical Image Analysis Term Project Report

Barış Kaplan **(0069054)**
bkaplan18

Canberk Eker **(0060233)**
ceker16

1. Explanation of Dataset Details, Motivation, and Defined Problem

In the term project, the medical imaging dataset called "Kvasir-SEG: A Segmented Polyp Dataset" is used. This dataset of gastroscopic images of polyps was produced by Simula Research Laboratory. The purpose of choosing this dataset is to do research on general gastroenterology. The gastrointestinal system of a human contains several essential parts including the large bowel. Various diseases and anomalies can disrupt the structure of the large bowel, one of them being colorectal cancer. Colorectal cancer is commonly observed among both women and men. Polyps constitute one of the main triggers of colorectal cancer and they can be frequently found in the large bowels after performing necessary colonoscopic examinations. Colonoscopy is a highly essential medical treatment method that can detect polyps and assess their severity, and this treatment can be followed by the removal of polyps if needed. However, various studies indicated that during the colonoscopic examinations, the polyps are missed at a percentage between 5% and 30%, which alters based on the types and sizes of the polyps [1], [2]. Missing them even at these rates can be costly and lead to the deaths of millions. This makes early polyp detection and diagnosis highly important and by doing so, colorectal cancer can be prevented and death rates due to this cancer can be reduced to a certain extent. This constitutes our primary motivation of working with the “Kvasir-SEG” dataset.

The dataset is an important resource for creating image processing, image segmentation, and machine learning algorithms for polyp detection and diagnosis. In the machine learning context, the Kvasir-SEG dataset can serve as a validation dataset or as a training dataset. A subset of the Kvasir dataset, Kvasir-SEG, has 1000 cases of polyps. In the Kvasir-SEG dataset, each image is represented with an 8 by 8 matrix. Moreover, the pixel values of an image of this dataset varies from 0 to 255; where 0 is the background and 255 is the foreground. In the Kvasir-SEG dataset, during gastroscopic examinations, the original polyp images were captured in the colored format, while the respective masks were captured in the grayscale format. Different kinds of polyps and normal zones devoid of polyps are depicted in the photographs. The dataset also contains the respective ground-truth mask annotations of the polyp images for segmenting polyps at the pixel level. The limits and locations of polyps are identified using these annotations. This dataset can be used by researchers to develop and test polyp detection and classification methods.

Numerous scholarly works have used and cited the Kvasir-SEG dataset. This shows that the dataset has received widespread recognition as a trustworthy source. The dataset is useful for research since it provides a rich collection of annotated examples that are both many and diverse. Segmenting polyps is the specified issue for this dataset. The objective is to create an algorithm that precisely recognizes and delineates the pixel-level borders of the polyp regions in the photos. Automating or semi-automating this process can help clinicians make precise diagnoses because it is a critical stage in the analysis and diagnosis of polyps in endoscopic pictures.

Kvasir-SEG is an open-access dataset with a focus on gastrointestinal polyp images and was created primarily for pixel-wise image segmentation in medical image analysis. It offers fully confirmed segmentation masks and bounding boxes of the polyp regions that have been manually annotated by experts in medicine [3]. The dataset makes it easier for researchers to compare and reproduce results using both conventional and deep-learning-based methodologies. It promotes contributions from multimedia and computer vision researchers in the field of polyp segmentation and automated analysis of colonoscopy pictures by extending the Kvasir dataset's frame-wise annotations with segmentation masks.

In conclusion, the Kvasir-SEG dataset is a useful tool for developing image processing, image segmentation, region of interest (ROI) localization, and machine learning techniques utilized in polyp identification in the field of gastroenterology. It is now a well-known dataset for research purposes and is frequently cited in the literature.

2. *Explanation About The Used Platform, Applied Methods and Parameter Selection Reasons*

Since the size of the Kvasir-SEG dataset is high, it required high CPU and GPU resources. Due to higher level of computational resource requirements, Google Colab is used as the development environment. As the programming language, Python is chosen due to its advanced support of built-in computer vision libraries. As the libraries, for performing the image processing and segmentation tasks, PIL and cv2 were commonly utilized in the implementation. The data was stored in Google Drive due to the large size of the dataset and high storage capabilities of Google Drive.

In the implementation, firstly, for mounting the Google Drive directory, the mount method from the “google.colab” library was utilized. Then, the dataset directory was defined and joined with the folder of polyp images to construct images directory and with the folder of and ground-truth masks to construct masks directory. Next, all polyp images are iterated with a for loop. Inside the for loop, each polyp image and its respective mask are extracted from their specific folders. Then, by using the “imread” function and grayscale code from the cv2 library, the images and their ground-truth masks are read in the grayscale format. The pixel values of each polyp image vary from 0 to 255; while the pixel values of each ground-truth mask are either 0 or 1. Therefore, to achieve consistency, pixel classification boundary is determined in this step. The pixel values less than or equal to 155 are classified as background, while the pixel values greater than 155 are classified as foreground. Then, the image processing was performed. While processing the images, initially, fast non-local means denoising was used for effectively removing noises that are scattered throughout the image and do not follow a specific pattern. This resulted in a better determination and preservation of the image details. Then, the grayscale image is converted to a PIL image with the usage of “fromArray()” function from the “Image” library. After that, the contrast of the PIL image was enhanced by a factor of 1.65. This value was determined after several trial-and-error-based experimentations and optimal results were obtained with the specified one. Lastly, in preprocessing, the Gaussian Blur was applied with the usage of a 5 by 5 kernel. The Gaussian Blur was selected due to its high effectiveness in reducing the noise. However, it came with the downside of blurred edges. 5, which is relatively small, is chosen as the kernel size for finer-grained noise removal. After preprocessing, Otsu’s thresholding method was performed on the enhanced image and the estimated foreground mask was obtained as a result of thresholding. Subsequently, several morphological operations were performed on the estimated foreground mask. Initially, a binary closing operation with a 13 by 13 kernel was applied on the estimated mask. In binary closing, the kernel size is selected as large to fill larger holes and less disrupt the foreground regions. Then, a binary dilation operation with a 5 by 5 kernel was performed 2 times. The kernel size was chosen as small for dilation to link finer-grained pixel regions that are unconnected and shrink finer-grained holes in the foreground regions. Next, to better preserve the regional connectivity, binary holes in the estimated mask were filled. After that, erosion was applied 2 times with a kernel size of 3 and binary closing was applied with a kernel size of 7 afterward. For erosion, the kernel size is selected as small to better separate the under-segmented regions with more preservation of details. The kernel size for closing is selected as relatively large to better preserve the original image's structure. Finally, a binary opening operation was performed on the estimated mask 2 times with the usage of a 9 by 9 structuring element. The size of the structuring element is chosen as large in opening to better smooth the boundaries. After applying the morphological operations, the estimated segmentation map was plotted. While plotting; the "imshow", "show", "figure", "xlabel", "ylabel", and "title" functions from the "matplotlib" library were utilized. Then; the true positive, true negative, false positive, and false negative counter values were found by comparing the estimated foreground mask and corresponding gold mask pixel by pixel. Then, with these TP, TN, FP, and FN counters, several metrics, namely dice index, specificity, overall accuracy, recall, f-score, and precision were calculated to evaluate the performance of the polyp segmentation.

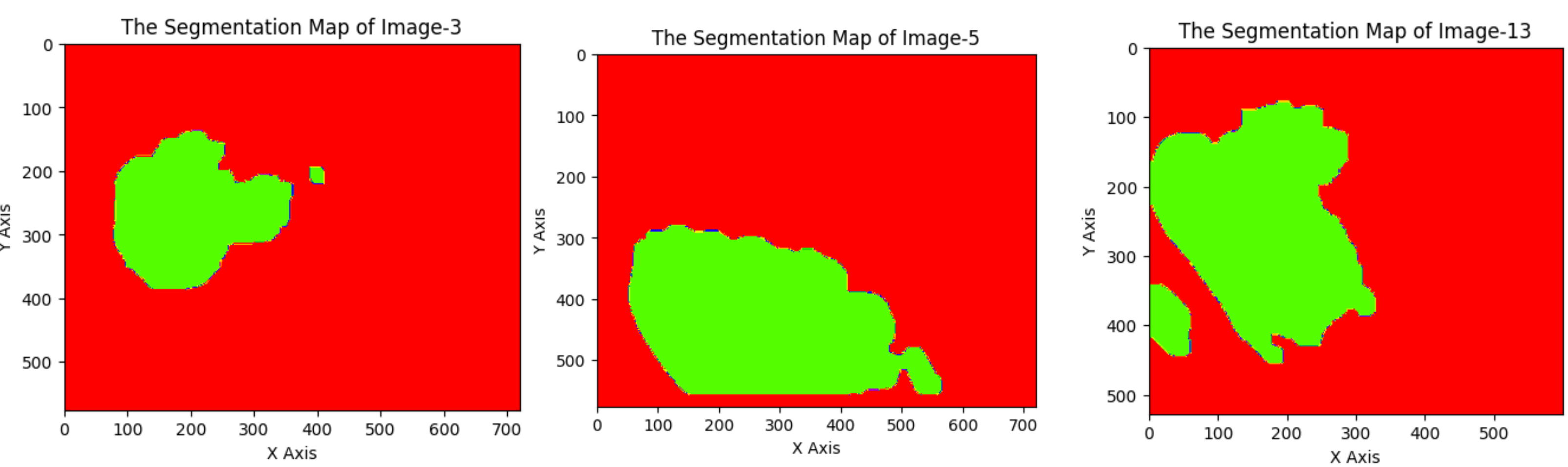
3. Experimental Results

3.1 Values of the Evaluation Metrics

	Overall Accuracy	Precision	Recall	Fscore	Specificity	Dice Index
Image-3	0.89686	0.43004	0.43004	0.43004	0.92907	0.486420
Image-5	0.75354	0.24523	0.24523	0.24523	0.79341	0.32294
Image-6	0.80775	0.71867	0.71867	0.71867	0.89488	0.65981

Note: The experiment results are denoted for the subset of Kvasir-SEG dataset, which consists of 3 images. The indexes of the images represent their original indexes within the Kvasir-SEG dataset.

3.2 The Segmentation Maps



4. Interpretation and Discussion of the Experiment Results (including advantages & limitations)

We performed experimentations on the subset of the Kvasir-SEG dataset, which contains 3 polyp images. We achieved high overall accuracy values for all images we experimented with. Moreover, the specificity values we found can also be considered as high. However, we achieved relatively lower dice index, precision, recall, and f-score values compared to the results obtained for the other metrics (i.e., specificity and overall accuracy). This shows that after applying the segmentation algorithm, the number of true negatives found was slightly higher than the number of true positives found. In other words; it is found that the pixels belonging to the background regions were more accurately estimated than the pixels which belong to the foreground parts. At the end, as the advantage of the applied methods, a higher number of true negatives was achieved overall. This meant a more accurate estimation of the pixels of the background. However, as the limitation of them, less true positives achieved compared to the number of true negatives. This meant a less accurate estimation of the pixels belonging to the foreground. With further analysis and experimentations with different segmentation parameters, these results can be improved and a higher number of true positives can be obtained. Including more polyp images and masks in the Kvasir-SEG dataset may also aid in enhancing the performance of the applied segmentation algorithm and thus the values of the evaluation metrics.

References

[1] Kim NH, Jung YS, Jeong WS, Yang HJ, Park SK, Choi K, Park DI. Miss rate of colorectal neoplastic polyps and risk factors for missed polyps in consecutive colonoscopies. Intest Res. 2017 Jul;15(3):411-418. doi: 10.5217/ir.2017.15.3.411. Epub 2017 Jun 12. PMID: 28670239; PMCID: PMC5478767.

[2] Ahn SB, Han DS, Bae JH, Byun TJ, Kim JP, Eun CS. The Miss Rate for Colorectal Adenoma Determined by Quality-Adjusted, Back-to-Back Colonoscopies. Gut Liver. 2012 Jan;6(1):64-70. doi: 10.5009/gnl.2012.6.1.64. Epub 2012 Jan 12. PMID: 22375173; PMCID: PMC3286741.

[3] Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., & Johansen, H. D. (2020). Kvasir-seg: A segmented polyp dataset. In *MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26* (pp. 451-462). Springer International Publishing.