University of Texas Rio Grande Valley
ScholarWorks @ UTRGV

Theses and Dissertations

8-1-2024

White Light Specular Reflection Data Augmentation for Polyp Detection

Jose Angel Nunez
The University of Texas Rio Grande Valley

Follow this and additional works at: https://scholarworks.utrgv.edu/etd

 Part of the Computer Sciences Commons

Recommended Citation

Nunez, Jose Angel, "White Light Specular Reflection Data Augmentation for Polyp Detection" (2024). *Theses and Dissertations*. 1557.
https://scholarworks.utrgv.edu/etd/1557

This Thesis is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact justin.white@utrgv.edu, william.flores01@utrgv.edu.

WHITE LIGHT SPECULAR REFLECTION DATA AUGMENTATION FOR POLYP DETECTION

A Thesis

by

JOSE ANGEL NUÑEZ

Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: Computer Science

The University of Texas Rio Grande Valley
August 2024

WHITE LIGHT SPECULAR REFLECTION DATA AUGMENTATION

FOR POLYP DETECTION

A Thesis
by
JOSE ANGEL NUÑEZ

COMMITTEE MEMBERS

Dr. Bin Fu
Chair of Committee

Dr. Emmett Tomai
Committee Member

Dr. Tamer Oraby
Committee Member

August 2024

Copyright 2024 Jose Angel Nuñez
All Rights Reserved

ABSTRACT

Nunez, Jose A., White Light Specular Reflection Data Augmentation for Polyp Detection. Master of Science (MS), August 2024, 48 pp., 18 tables, 17 figures, 25 references.

Colorectal cancer is among the deadliest cancers, but fortunately, this type of cancer can be prevented. The best current method of prevention is via detecting the bad polyps in the colon in time. Furthermore, the best method we have available to detect these bad polyps is through colonoscopies. Even though a lot of lives have been saved via these methods, it is still not perfect because of human error. Integrating artificial intelligence into colonoscopy procedures is our next evolution in increasing our prevention of colorectal cancer. Polyp detectors are one of the tools brought by advancements in technology that may aid doctors in colonoscopy procedures. It is vital that we continually improve the quality of polyp detectors. A common problem in polyp detectors is that they sometimes confuse the white light specular reflections produced by the endoscope with polyps. This study proposes a new data augmentation technique that artificially augments the images with more white light specular reflections. The hypothesis is that by providing the model more opportunities to make mistakes, it also gives it more chances to learn from those mistakes, thus improving the quality of the model.

DEDICATION

I dedicate this work to my family and friends. It would not be possible without your love, support, and patience. Thank you!

ACKNOWLEDGMENTS

Accomplishing this master's thesis would have been 10 times harder without the following people.

Fabian Vazquez, thank you for being a great research partner. Our teamwork was impeccable, and we accomplished a lot in a short period of time.

Dr. Tamer Oraby, thank you for providing me with a strong foundation in statistics. The honors project I did with you during my undergraduate, “Markov Chains and Random Walks,” helped me tremendously in my journey to learning artificial intelligence.

Dr. Dong-Chul Kim, thank you for being a great teacher in any area related to artificial intelligence. I learned so much in your reinforcement learning and deep learning classes.

Dr. Emmett Tomai, you were my first computer science teacher when I was pursuing an undergraduate degree in business. Thank you for being an amazing teacher; you inspired me to pursue a master's in computer science.

Dr. Bin Fu, thank you for all your guidance and mentorship. I will always remember your wise phrases, especially, “Focus on something small but important.”

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

1.1 A World Health Problem

Colorectal cancer (CRC) is a significant health problem in our world. According to the World Health Organization (World Health Organization, 2021), CRC ranks as the third most common cancer worldwide and is the second leading cause of cancer-related deaths. Furthermore, CRC's 5-year survival rate ranges from 48.6% to 59.4% (Morris et al., 2007). Factors such as gender, age, smoking, body mass index (BMI), and family history increase CRC risk (Joseph et al., 2012). Initially, medical guidelines recommended colorectal cancer screenings starting at age 50. However, the American Cancer Society now advises beginning screenings at age 45 due to a significant increase in incidence rates among younger adults. Notably, incidence rates for colorectal cancer have been rising by 1% to 2% annually, particularly among individuals under 55 years old (Siegel et al., 2024). This adjustment of earlier screening aims to address this younger demographic's growing number of cases and improve early detection and treatment outcomes.

There seems to be some sort of connection between CRC incidence rates and the economic power of a country. It has been observed that higher-income countries experience higher incidence rates of CRC but lower mortality rates than lower-income countries (Favoriti et al., 2016). Some speculate that it could be the dietary changes that evolve as a result of economic improvement such as access to a wider variety of foods, such as sugary and junk foods.

However, it is important to note that although incidence rates are higher in better off countries, the mortality rates due to CRC are lower and is likely due to better treatment options and medical access in these countries. There was a study done on 37 countries to find patterns in CRC incidence rates and it was found that incidence and mortality rate correlate with present human development levels (Arnold et al., 2017). Furthermore, CRC is expected to increase significantly by 2040, with an estimated 3.2 million new cases and 1.6 million deaths in high Human Development Index countries (Eileen et al., 2022).

This type of cancer also poses a substantial economic burden. It causes a strain on healthcare resources, it increases health insurance premiums, affects the workforce, and treatment costs are massive. A study done in Canada, found that increased screening participation to 60% would lead to better clinical outcomes across all income groups, which would translate to lower treatment costs, saving on average 95 million Canadian Dollars per year over the period of 2024 to 2073 (Adegbulugbe et al., 2024). The economic burden of different types of cancers was studied in the European Union (EU). It revealed that CRC incurred an economic cost of 13.1 billion euros in the EU in 2009, which amounted to 10% of the total cancer costs (Ramon et al., 2013).

CRC has a plethora of negative side effects, ranging from affecting our loved ones to affecting our economy. Fortunately, there is hope; CRC is largely preventable. This type of cancer often develops from polyps, which are little bumps that may form in the colon. Most polyps are small, less than 10mm in size. Some of the polyps are benign and pose little to no threat; these are called hyperplastic polyps. While other polyps have a high risk of becoming cancerous, these polyps are known as adenomatous polyps. Thus, the best way to prevent this type of cancer is to detect adenomatous polyps in time and remove them.

Colonoscopies are considered the gold standard for detecting polyps. A colonoscopy is a medical procedure in which an endoscopist examines the colon for polyps using an endoscope, a flexible tube apparatus containing a camera and light. Once an adenomatous polyp is detected, it is removed. It is important to note that even though colonoscopies are great for detecting polyps, this procedure is dependent on the experience of the endoscopist. Furthermore, human error still exists in all professions and at all levels of experience. About 12% of large adenoma polyps (at least 1cm) are missed even by experienced endoscopists (Bernal et al., 2021). According to another study, the rate of missed polyps during a colonoscopy can vary between 15% to 35% depending on the size of the polyp (Lalinia et al., 2023).

Artificial intelligence is helping evolve many industries. Regarding colonoscopies, there are many different AI models that aid in detecting polyps. Providing these AI tools to endoscopists will reduce the chances of missed polyps. In order for a polyp detector to successfully integrate well with a colonoscopy procedure, it must work in real time and be as precise as possible.

1.2 Contribution

Striving towards the goal of increasing the prevention rate of CRC requires decreasing the missed polyp rate. To accomplish this using AI, we must aim to make improvements to polyp detector models in precision while ensuring that they work in real time.

After reading the literature on polyp detectors, I noticed that most mention a common problem with these detectors: they confuse the white light specular reflections produced by the endoscope for polyps. This problem results in more false positives, which as a result reduces the precision of the model. To my knowledge, I have not found any research papers that address solutions to fix the white light specular problem in polyp detection.

Therefore, this thesis attempts to use the knowledge of the white light specular reflection problem and use it to make polyp detection stronger. This is accomplished via creating a new data augmentation. Data augmentations are widely used to help object detection models become more robust and better generalized. Popular data augmentations include mosaic, random horizontal flip, random crop, Gaussian blur, mix-up, and a whole bunch more.

The data augmentation method created for this paper is called White Light Specular Reflection, or WLSR for short. This data augmentation generates more white light specular reflections, let's call them artificial lights for short, and adds them to the training images. The purpose is to use the knowledge that white light specular reflections are troublesome for polyp detectors and, therefore, make the training images harder by adding artificial lights. My hypothesis is that by proving harder training via WLSR data augmentation, it will make the model stronger, and thus make gains in mean average precision (mAP).

CHAPTER II

LITERATURE REVIEW

2.1 Object Detector Basics

An object detector comprises two tasks: classifying and locating the object of interest. Classifying means correctly identifying the object. For example, if given an image of a polyp, it tells us that there is a polyp in the image. Localization tells us where the polyp is in the image, utilizing a bounding box that encapsulates it. Therefore, object detectors would be advantageous in colonoscopy procedures because they would help endoscopists in finding polyps.

Object detectors output a bounding box to signify where they think the object of interest is located. Thus, to measure how well an object detector is at locating objects, we must compare the bounding box produced by the model with the ground truth box. The ground truth box is the real location of the object. This comparison is done by calculating the intersection over union (IoU). IoU ranges from 0 to 1, where 0 indicates no overlap, meaning the model did not locate the object. An IoU of 1, indicates that the model located the object exactly. The closer the IoU is to 1, the better the model is at accurately locating objects of interest. In object detection, an IoU of 0.5 is considered acceptable. Figure 1, shows a visual representation of IoU.

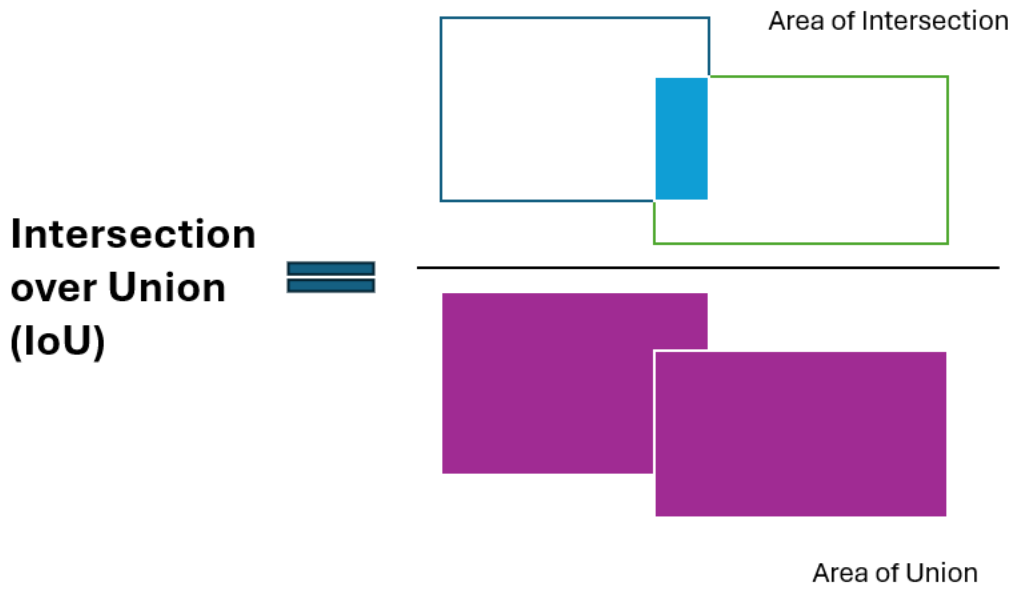

Figure 1: Intersection over Union

The performance metrics that evaluate object detectors utilize IoU in their calculations by specifying an IoU threshold. For example, if we set the IoU threshold to 0.5, then if the model produces a bounding box with an IoU above the threshold, then we would consider that it detected the object successfully, in other words, a true positive. The main performance metrics are precision, recall, F1 score, and mean average precision (mAP). The F1 score and mAP are derived from precision and recall. We will now briefly go over these metrics.

Precision indicates how good the model is at getting its detections correct. A correct detection is known as a true positive. An incorrect detection is a false positive. A false positive would mean that the model indicated that there is a polyp, but there is no polyp. This also means that the model outputs a bounded box but the IoU is below the threshold. In essence, precision is the percentage of predictions that are correct. The equation for precision is shown below:

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall tells us how good the model is at detecting all the objects of interest. If there is a polyp but the model does not detect it, then it is considered a false negative. In medical imaging, we want to reduce false negatives as much as possible because if not, they can result in dire consequences. In this case, not detecting an adenomatous polyp can result in the patient developing colorectal cancer. The equation for the recall metric is shown below:

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

The F1 score combines both precision and recall into one score. The F1 score is known as the harmonic mean of precision and recall. The equation for the F1 score is shown below:

$$F1 = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

Another metric that combines precision and recall is mean average precision (mAP). I think this metric is very intuitive and important. To understand it, we must first understand precision-recall curves. The precision-recall curve captures the trade-off relationship between precision and recall. This trade-off relationship exists because we can modify certain parameters, resulting in higher precision but would cause lower recall, and vice versa. For example, we can make the model more careful when making predictions; this would improve precision since it would make fewer false positive predictions. However, since fewer predictions would be made, we also increase the chances of more false negatives, thus decreasing our recall score. The precision-recall curves show us the possible precision and recall combinations given different careful levels (confidence scores). As a model gets better, the precision-recall curves shift

outwards, providing us with better combinations of precision and recall. Therefore, to determine which model has better precision-recall curves, we calculate the area under that curve,

which gives us mAP. The mAP is usually calculated at IoU of 0.50, giving us mAP@50. Another important mAP calculation is the mAP@50-95, which is the average of mAPs from IoU at 0.5 to IoU at 0.95 in increments of 0.05. Figure 2 shows a precision-recall curve and mAP relationship.

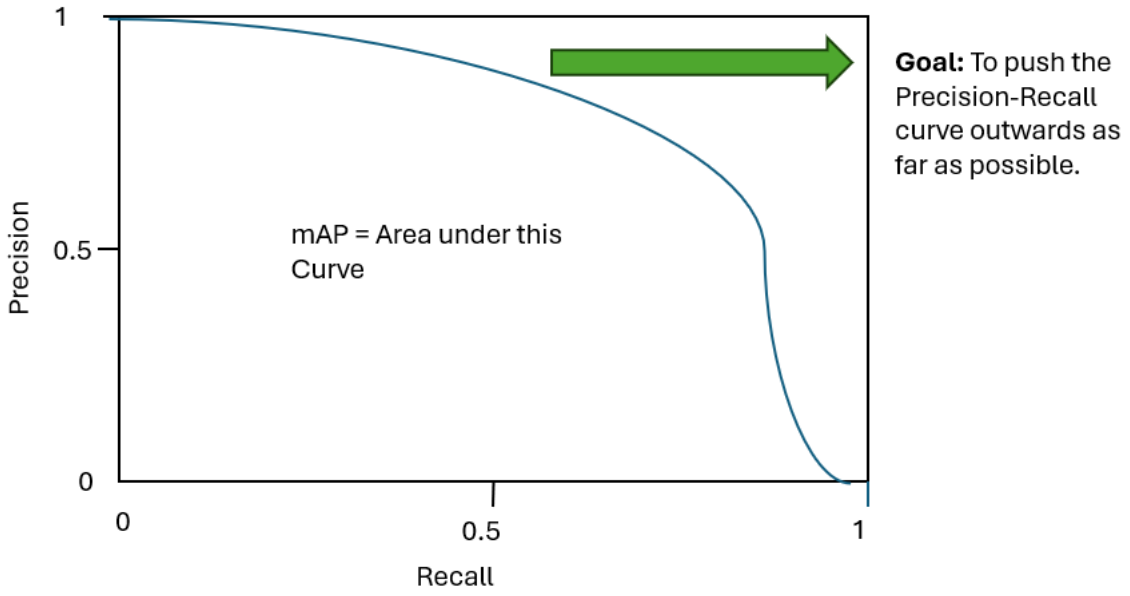

Figure 2: Precision-Recall Curve and mAP relationship

Object detection algorithms can be classified into two main architectures: single-stage and two-stage detectors. Single-stage detectors are the recent advances in object detection. Typically, two-stage detectors have more accuracy than single-stage detectors, but the downside is that they are slower. Therefore, choosing between single-stage or two-stage detectors depends on what is going to be needed. In the next section, we will briefly go over these two different architectures used in polyp detection.

2.2 Two-Stage Polyp Detectors

Two-stage detectors have a region proposal network that indicates areas of interest, and then these regions go to the second stage, which determines if there is an object of interest and the location, represented as a bounded box around the object. The R-CNN architecture is a good starting point if someone is interested in learning more about two-stage detectors. The following two-stage architectures are improvements over the R-CNN mentioned in order from oldest to newest: Fast R-CNN, Faster R-CNN, and Mask R-CNN.

In regards to detecting polyps, an R-CNN was used to detect images of the colon captured by colon capsule endoscopy (Tashk et al., 2020). The researchers mention that it is very time-consuming for medical professionals to analyze all colon images captured by the capsule endoscopy procedure, and there is a chance of missing a polyp. Another problem in polyp detection is that there is a lot of similarity between polyps and their surrounding tissue. Hence, a group of researchers proposed to do contrast enhancement on the images as a preprocessing step before passing it to their faster R-CNN architecture (Chen et al., 2021).

There are a lot of different studies using some form or variation of R-CNN on polyp detection. An important mention to wrap up this section is from a research paper that used a Mask R-CNN for polyp detection and segmentation. The main takeaway was that deeper and more complex networks do not necessarily guarantee better performance (Qadir et., 2019).

2.3 Single-Stage Polyp Detectors

Single-stage detectors achieve object detection in one stage. They accomplish this by using a unified architecture, meaning they perform both object localization and classification in

one pass through the network. The two most popular single-stage detectors are the Single Shot MultiBox Detector (SSD) and You Only Look Once (YOLO).

The speed of single-stage detectors makes them perfect for detecting polyps during colonoscopies, which must happen in real-time. A polyp detector using the SSD architecture and Inception V3 as a feature extractor demonstrated great speed in detection (Liu et al., 2019).

Currently, YOLO is more popular than SSD. One reason for its popularity is that its architecture is straightforward and easy to modify. A testament to its ease of customization is from a research paper in which the researchers customized the YOLOv4-Tiny model and got promising results in detecting polyps (Doniyorjon et al., 2022). Another research team working on polyp detection combined ResYOLO, which is another variant of YOLO, with a tracker named Efficient Convolution Operators (ECO) in order to incorporate temporal data (Zhang et al., 2018). Because of the customization capabilities of YOLO and the big community support due to its popularity, this project uses YOLO.

2.4 Current Top-Performing Polyp Detectors

This section will mention three top-performing polyp detectors with interesting methodologies. They all use the YOLO architecture. Their approaches to improving polyp detectors are very interesting.

The first one I want to mention uses the YOLOv3 architecture and integrates an object tracker into it (Nogueira-Rodríguez et al., 2022). Their goal in using the object tracker was to reduce the false positives by filtering out bounding boxes. Their algorithm runs on video frames, and since between frames, there is very little change if a polyp is found, if suddenly a polyp appears near an area where there was no polyp detected, it is filtered out, thus reducing the false

positives. They created a dataset containing 28,576 polyp images taken from 941 polyp videos. Their testing dataset contained 8,658 images. Their network was trained for 50 epochs but they found they had the best results at the 37th epoch. When incorporating their frame-based model, they achieved an F1 score of 0.88 (precision = 0.89 and recall = 0.87). They also mention that they got lower predictive performance on flat polyps.

The second study used many variants of YOLOv5, but it got its best results with the large variant (Ghose et al., 2024). It took an interesting approach by fine-tuning the anchor boxes in the YOLOv5 model to capture smaller polyps. In addition, they performed the following data augmentations: noise invocation, flipping, rotation, brightness, and contrast adjustments. Their training and testing data came from the Kvasir-SEG dataset, which contains 1,000 polyp images. They split that dataset into 80% for training and 20% for testing images. They then further split the training dataset to make room for a validation set. In total, they had 640 images for training, 160 for validation, and 200 for testing. Their best model resulted from using the largest variant of YOLOv5 and training with early stopping at 1,000 epochs. They reported 0.9840 precision, 0.9861 recall, and 0.964 F1 score.

The third study builds upon YOLOv3 by modifying the architecture to make it anchor-free, they named this modified model, YOLO-OB (Yang et al., 2023). Their motivation was to increase flexibility so polyps could be treated equally at different feature levels regardless of size or shape in the original image. Their total dataset consisted of 102,234 polyp images that came from SUN a public database and a private database named Union. They trained their model for 150 epochs and reported the following results: precision 98.37, recall 98.23, F1 98.32, and mAP 98.19.

2.5 Challenges of Detecting Polyps

Many challenges exist when attempting to detect polyps, such as polyps being too small or too flat, occluded by colon folds, endoscope blind spots, inadequate light conditions during image acquisition, and white light specular reflections. Researchers have attempted various things to combat these challenges. A research team attempted to tackle the inadequate light condition challenge by making their colonoscopy images go through a preprocessing stage in which the images get image enhanced by converting from RGB to HSV color space(Nisha et al., 2022).

This thesis aims to find a possible solution to one of the challenges mentioned above, the white light specular reflection problem. In the next section, the proposed method will be discussed.

CHAPTER III

METHODOLOGY

This chapter will go over in detail the proposed data augmentation and the experiments performed.

3.1 Data

There are not that many publicly available polyp images. The lack of a sufficient amount of polyp images is one of the reasons that polyp detectors struggle.

The data used for this research comes from the Harvard Dataverse. This dataset comprises polyp images sourced from MICCAI 2017, CVC Colon DB, GLRC, KUMC, and the University of Kansas Medical Center, each accompanied by XML annotations. The XML annotations describe important details such as coordinates of where the polyp is located in an image. This data contains 3 main folders: train 2019, validation 2019, and test 2019.

3.2 Object Detector Model: YOLOv5

YOLO is single stage object detector model that is used for object detection. The first version through third version of YOLO were made by Joseph Redmon and his collaborators (Redmon et al., 2016). The YOLO model became very popular because of its speed, and it is used in many areas, including polyp detection. This popularity resulted in more versions of YOLO being produced by different research teams such as YOLOv4, which added mosaic augmentation and other improvements (Bochkovskiy et al., 2020). The same authors of YOLOv4 created YOLOv7 (Wang et al,. 2022). YOLOv6 was made by Meituan, a Chinese technology

company (Li et al,. 2022). YOLOv5 and YOLOv8 were developed by Glenn Jocher and his team at Ultralytics. Currently, more versions of the YOLO model are being created.

This research utilized YOLOv5 because there are currently more resources available for this version. In addition, this version of YOLO is easier to customize and modify its code if needed. Furthermore, YOLOv5 has 4 different models that are pre-trained on the COCO dataset. YOLOv5s, the small model was used for this research because polyp detectors must be fast enough to work in real-time and this model accomplishes that goal.

3.3 Annotation Conversions

When training using a YOLO model, for every image there must be an annotation file in form of a text (.txt) format that contains the bounding box coordinates for each object in the image. The dataset for this project contains the annotation files in XML format.

The coordinates of a bounding box in XML format are given by the top left corner and bottom right corner of the bounding box, in other words, x-min, y-min, x-max, and y-max. The coordinates of a bounding box in YOLO format are given by center of the box, the width of the box, and the height of the box. In addition, these values are normalized by dividing by the width and height of the image, thus each value for a bounding box range from 0 to 1.

Two different codes were made to do the conversion from XML to YOLO format. One was to convert the annotations for the training images folder. The second one was for converting the annotations for the test and validation folder, since these two folders had subfolders inside them. Furthermore, when an annotation did not have a corresponding image then that annotation was not used. In addition, when an annotation did not specify bounding box coordinates for polyps then it was also not used, and neither was its corresponding image. This resulted in trimming our dataset a bit. Further trimming was done a bit more if it didn’t meet a certain

condition during the WLSR data augmentation pipeline, which will be discussed soon. At the end, our dataset consisted of 26,942 training images, 4,214 validation images, and 4,685 test images.

3.4 WLSR Data Augmentation Design

This section will now dive into the high-level view of the WLSR data augmentation.

3.4.1 Bank of Lights

The artificial lights for this data augmentation had to look like real white light specular reflections that occur during colonoscopies. Simply placing white circles or ovals was not going to be beneficial for training since we needed to mimic the real thing. Therefore, the solution was to meticulously pick and choose 300 lights from real polyp images.

First, color segmentation was done on all polyp images so we could target the shades of white light that occur during colonoscopies. Any color outside that shape was turned black so it could be easier to paste our desired lights onto an image later. Figure 3 shows a before and after picture of an image after color segmentation was applied to it with our desired shades of white light.

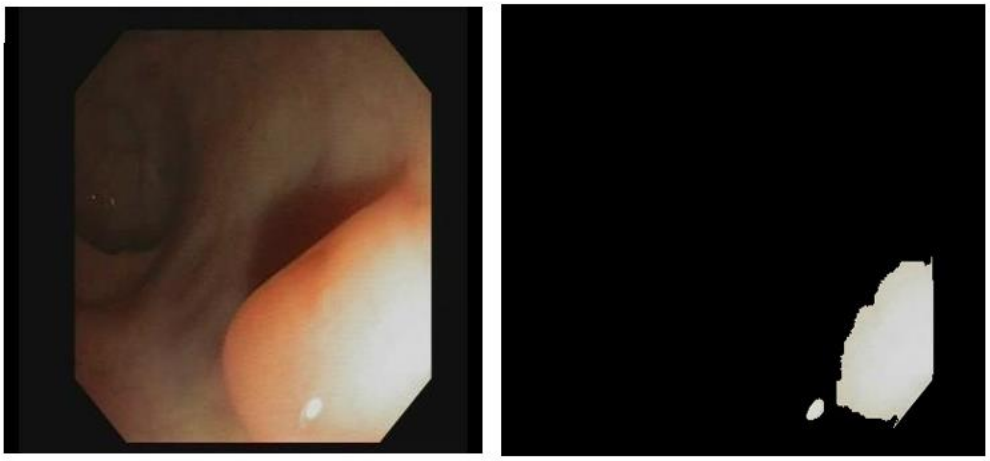

Figure 3: Before and After White Light Shade Color Segmentation

After segmenting our desired regions for all images, then the cropping phase began. Images were carefully analyzed first, then regions that would provide a great variety of white lights were cropped. Overall, 300 regions were cropped. Figure 4 shows some of the cropped images that were selected.

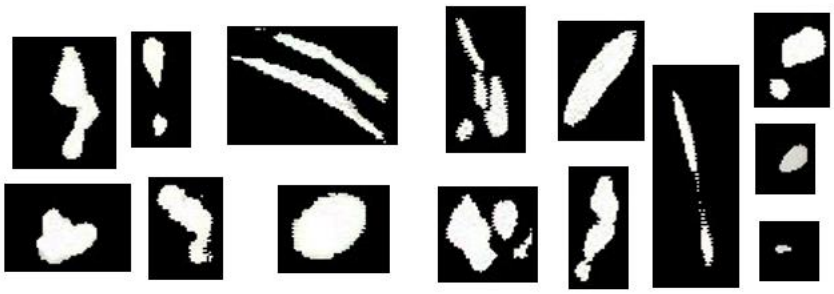

Figure 4: Examples of Selected Cropped Lights

Transformations were done to create more light images from our cropped lights. Eleven different transformations were performed to increase the lights in our Light Bank. Table 1 shows

the transformation performed. Parenthesis indicates the input of a given transformation. Figure 5 shows examples of some of the transformations performed.

Table 1: Transformations to Create the Bank of Lights.

Name	Quantity
Normal	300
Flipping Horizontal(Normal)	300
Flipping Vertical(Normal)	300
Random Scaling (Normal)	300
Random Rotation(Normal)	300
Random Rotation (Flipping Vertical)	300
Random Rotation(Flipping Horizontal)	300
Random Scaling (Flipping Vertical)	300
Random Scaling (Flipping Horizontal)	300
Random Scaling(Random Rotation(Normal))	300
Random Scaling (Random Rotation(Flipping Vertical))	300
Random Scaling (Random Rotation (Flipping Horizontal))	300
Total	**3600**

- *Random rotation range: -30 to 30 degrees*
- *Random scaling between 0.8 to 1.2*
- *Lights less than 1347 pixels can be scaled from 1 to 1.2. Else, the scale ranges from 0.8 to 1.*

Figure 5: Examples of Transformations Performed

All the transformations listed above on our 300 selected cropped lights make up all the possible artificial lights for the WLSR data augmentation. Therefore, in total the WLSR data augmentation has 3,600 different artificial lights that it can use. Figure 6, shows a sample of the artificial lights that make up the Bank of Lights.

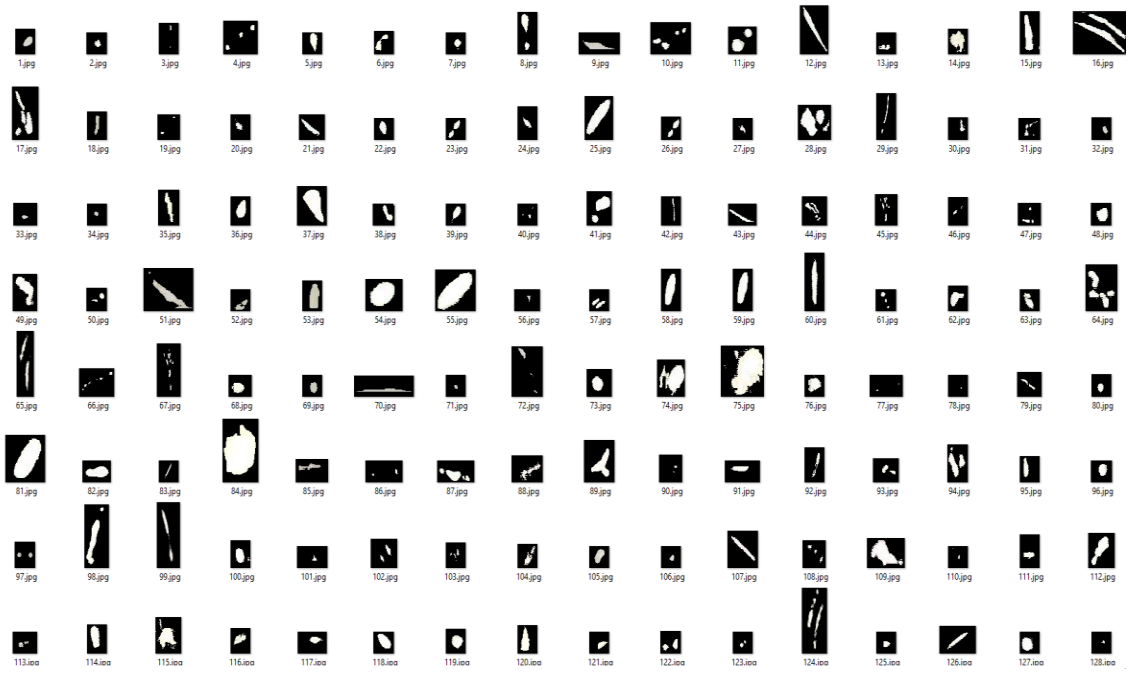

Figure 6: Sample of Bank of Lights

3.4.2 Orange Areas

There are three things to consider when placing our artificial lights in our images.

1) It must not overlap current lights because it could deform them and create the shapes of lights that don't exist during colonoscopies.
2) To avoid possible total occlusions of polyps, we will not let the artificial lights overlap at all with polyps.
3) Since the images in the dataset contain black borders of different sizes and shapes, we will ensure that the artificial lights land on colon areas and not on the black borders.

The solution to meet the conditions above is to convert the prohibited areas to a color that does not appear during colonoscopy procedures. In this case, orange is chosen.

First, we converted the borders to orange. Since the images have black borders of different sizes and shapes, a good enough range to encapsulate the borders was chosen: 20 percent for the top, left, down, and right sides of the images. Any pixel that is inside the border range mentioned

above and within a defined range of black gets turned orange. Figure 7 shows the orange border conversion.

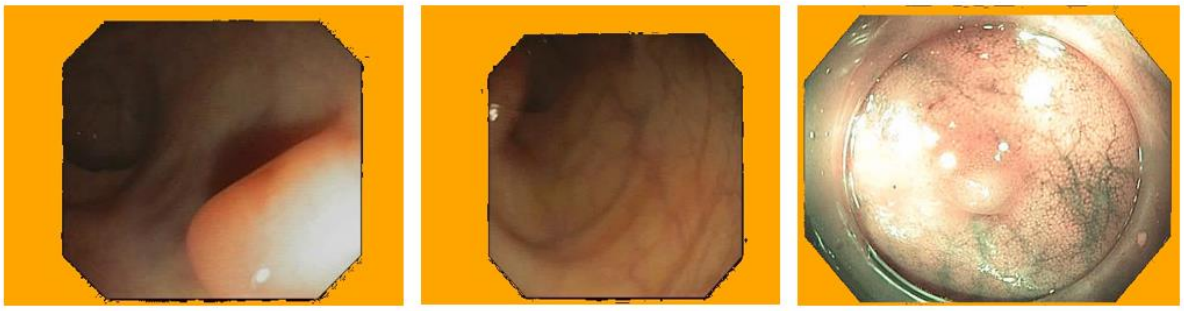

Figure 7: Orange Borders

Next, we must get bounding boxes for the current white lights to avoid overlapping them. We use the color-segmented images that we generated when doing the bank of white lights. We iterate through each segmented image and find the contour of each segmented object in an image in order to get the bounding boxes. Then bounded boxes get filtered with non-max suppression and lastly, they get converted to YOLO annotation format and saved.

To avoid passing two different annotation text files per image (annotation for polyps and annotation for white lights), a code was written to combine them into one annotation text file. Figure 8 shows the bounding boxes and orange boxes of an image.

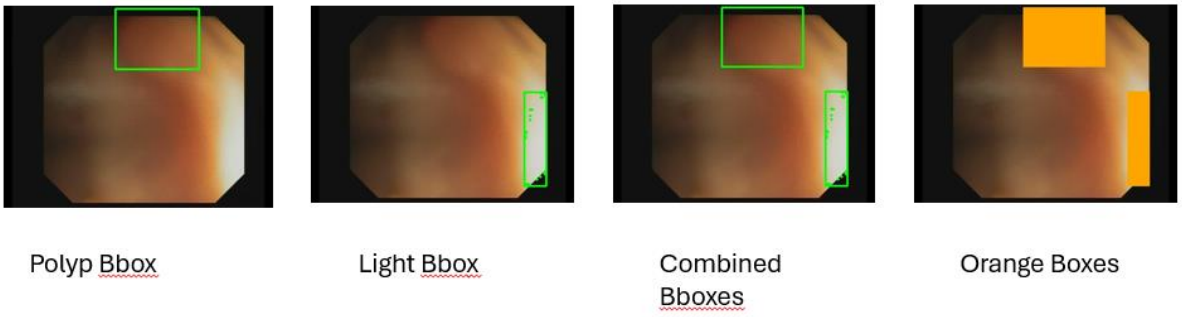

Figure 8: Bounding Boxes to Orange Boxes

Finally, to get all the desired areas we are trying to avoid, we combine both orange borders and orange boxes. To do this, we first do the orange border conversion and then we pass that image to the orange box converter. Figure 9 shows the orange area process.

Figure 9: Orange Area Flow

3.4.3 Areas that Fit

After obtaining the areas where we cannot add an artificial light, we then proceed to find the areas that fit. We accomplish this via a sliding window approach. We first generate a random artificial light from the bank of lights and then starting from the top left corner it moves in sliding window fashion. From the left to right until it reaches the end of the image and then it steps down and follows the same pattern from left to right until all areas of an image have been covered. The step size to move horizontally is the width of the cropped light image plus one. The step size to move vertically is the height of the cropped light image plus one.

If at any point during the sliding window process there is no orange pixel in the whole area that the cropped light image is covering, then the top left coordinate of that location gets stored in an array. The array of top left coordinates provides us with all the areas that fit. Lastly, from that array it randomly picks one top left coordinate where we are going to place the artificial light. I decided to do only one additional artificial light to see if it results in at least a minimal improvement to the model. In the future, I plan to test it with more than one artificial light added to images. Figure 10 illustrates the areas that fit and the random selection of an artificial light.

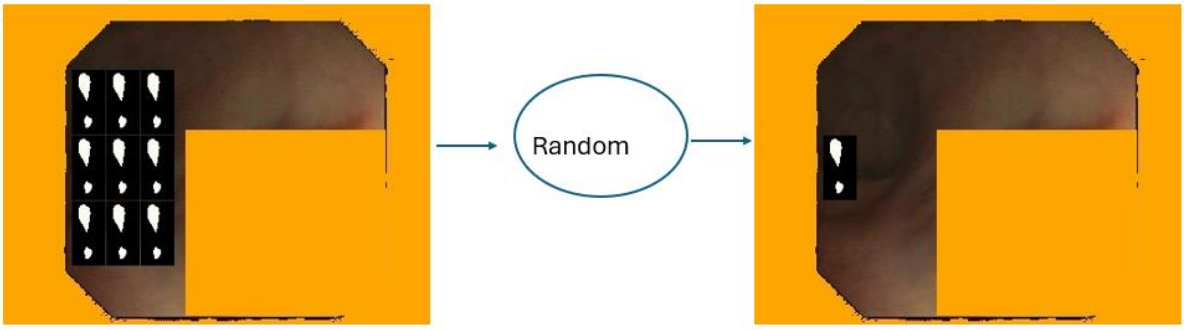

Figure 10: Random Selection from Areas that Fit

If no areas fit, the algorithm tries again n number of times with different cropped light images. After n tries, that image and its respective annotation are not used for training.

Lastly, only the non-black portions are pasted onto the image. Figures 11 and 12 show a few of the results of the WLSR data augmentation, along with an extra image that circles where the new light is located. Finally, Figure 13 shows a high-level view of the WLSR data augmentation algorithm.

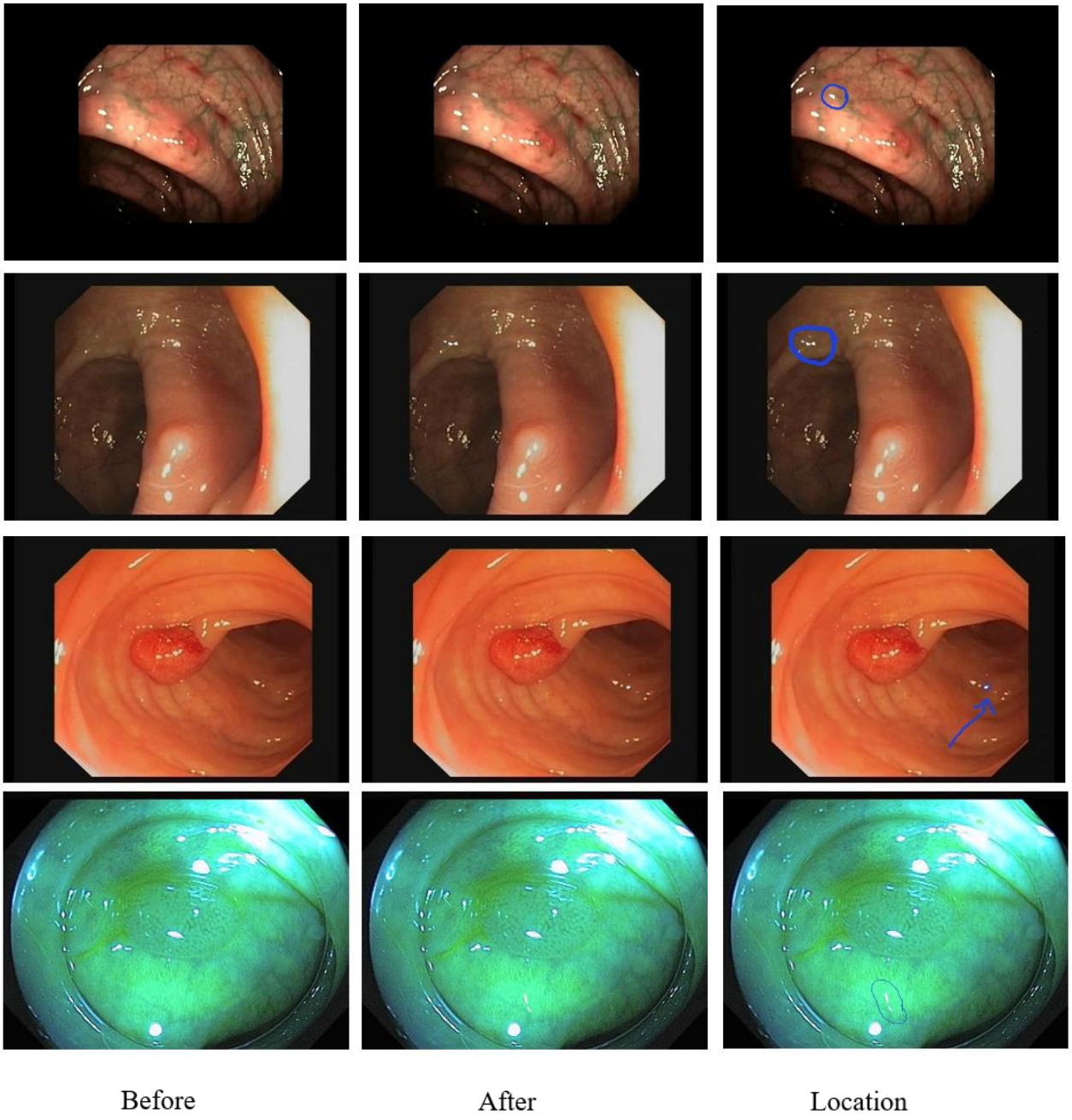

Figure 11: WLSR Data Augmentation Results Part 1

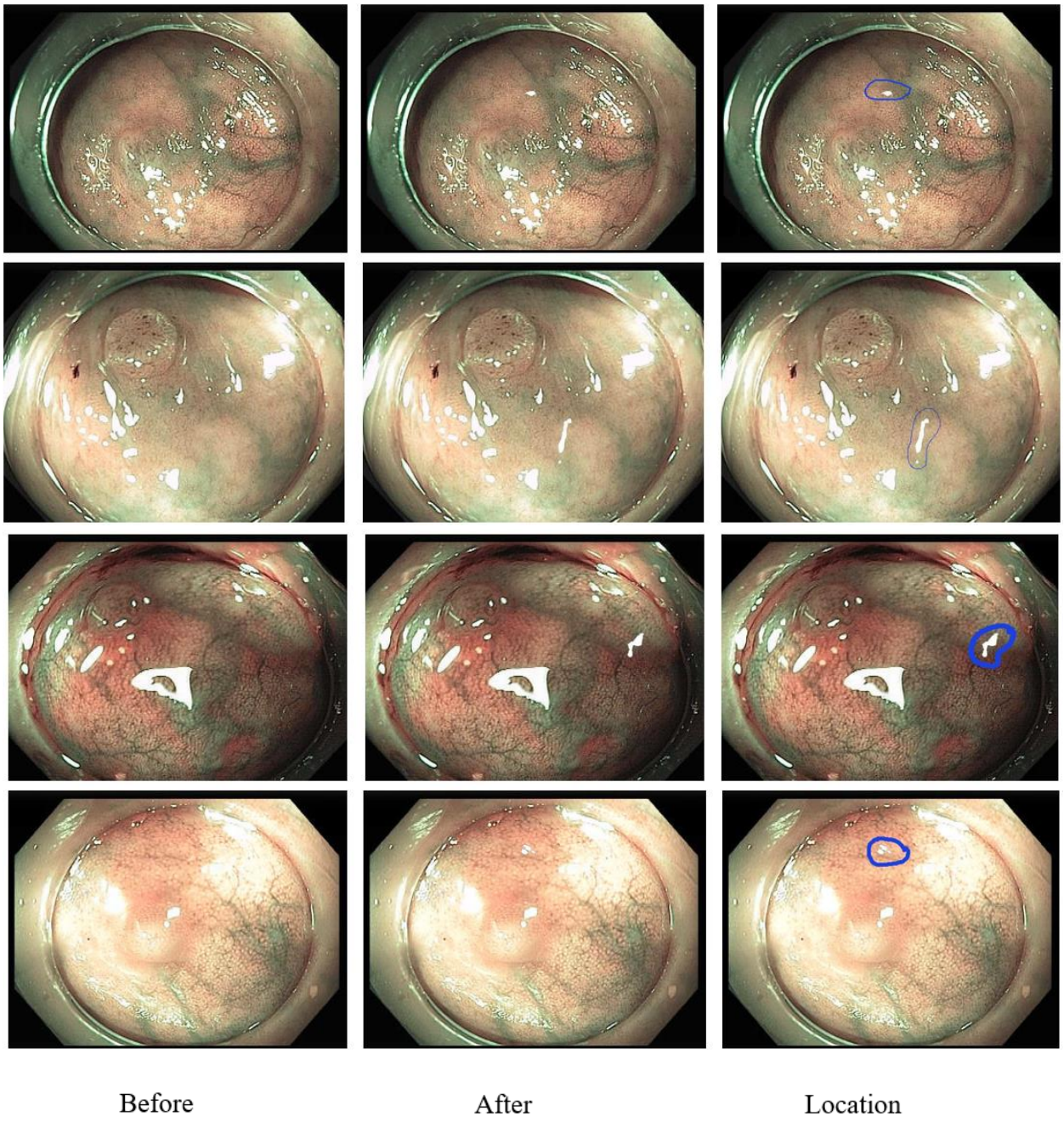

Figure 12: WLSR Data Augmentation Results Part 2

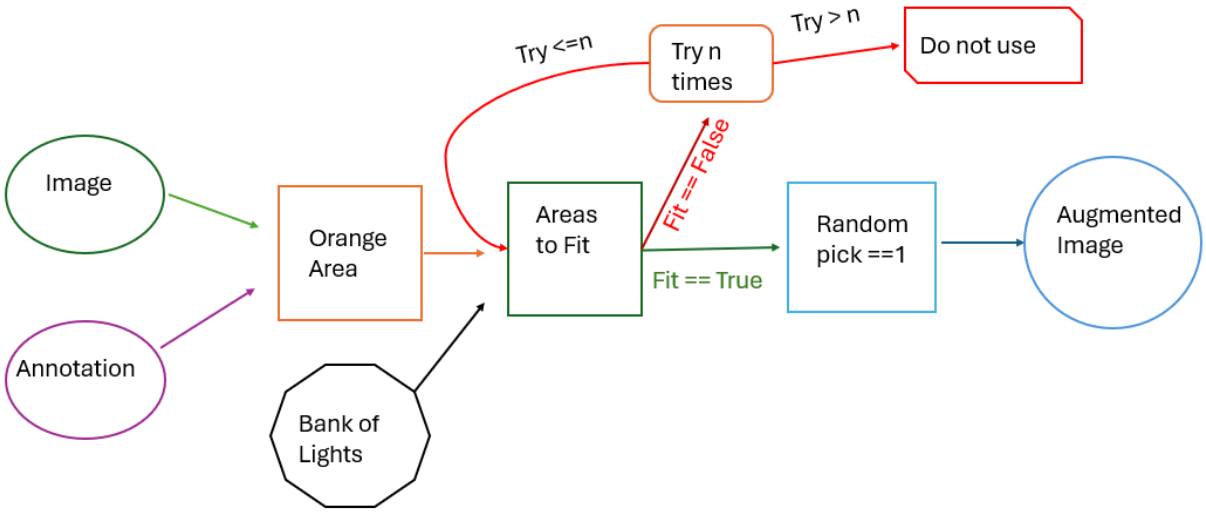

Figure 13: High-Level View of the WLSR Data Augmentation Algorithm

3.5 Experiments

The following experiments aim to test the WLSR data augmentation under various conditions. All the models that were trained utilized the pre-trained COCO model as a starting point. This was done to speed up training, and because the data size for polyps is small, it is highly recommended to start training on a pre-trained model. Default training parameters, test parameters, and hyperparameters were used in these experiments. We kept them consistent across all experiments to clearly see if models with WLSR data augmentation included offered an increase in performance metrics. The default data augmentations were also disabled to make it clearer if WLSR is benefiting the model. Stochastic gradient descent was used as the default optimizer. Consistent parameters and hyperparameters are shown in the tables below.

Table 2: Consistent Training Parameters Across Experiments

Parameter	Value
Image Size	640
Batch Size	16

Table 3: Consistent Test Parameters Across Experiments

Parameter	Value
Image Size	640
Batch Size	16
Confidence Threshold	0.25
IoU Threshold	0.45

Table 4: Consistent Hyperparameters Across Experiments

Hyperparameter	Value
Initial learning rate	0.01
Final learning rate	0.2
Momentum	0.937
Weight decay (L2 regularization)	0.0005
Warm up epochs	3
Warm up momentum	0.8
Warm up bias learning rate	0.1
Box loss gain	0.05
Class loss gain	0.5
Class losss positive weight	1
Objectness loss gain	1
Objectness loss positive weight	1
IoU treshold for matching	0.2
Anchor matching threshold	4

We will now describe the experiments done to determine whether WLSR benefits polyp detector models.

3.5.1 Replacing 100% of Data

First, I wanted to understand how WLSR performs by just replacing the whole training dataset with its augmentation to see how it performed. This would give me a good starting point to understand how this augmentation fundamentally affects the images and the training.

In addition, to better understand a specific data augmentation, we must test it under various conditions, such as the number of training images and the number of epochs.

The effects of data augmentations may vary with the quantity of epochs. An epoch represents one full cycle through the whole training dataset. When creating an object detection model, more isn't always better. Too many epochs may cause the model to overfit, which means that it becomes too good at predicting patterns in its own training dataset, in a way becoming too rigid and losing flexibility when it comes to predicting patterns on unseen data. Some data augmentations may help in making the model more flexible and thus mitigating overfitting effects. Thus, it is vital to understand how data augmentations affect the performance of a model with different quantities of epochs. Table 5 shows different combinations of data size and epoch numbers used for the 100% replacement test.

Table 5: Different Combinations of Data Size and Epoch Numbers

Subset Size	Number of Images	Epochs
20%	5,388	10, 30, 50
50%	13,471	10, 30, 50
100%	26,942	10, 30, 50

Most of the time, a number is useful when we can compare it. Therefore, to better gauge the results of the 100% replacement test on WLSR, the same tests and training were performed on the unaltered training images (Baseline) and on five other popular data augmentations: Random Flip, Color Jitter, Cut Out, Gaussian Blur, and Brightness Adjustment. Figures 14 and 15 show before-and-after pictures of these data augmentations.

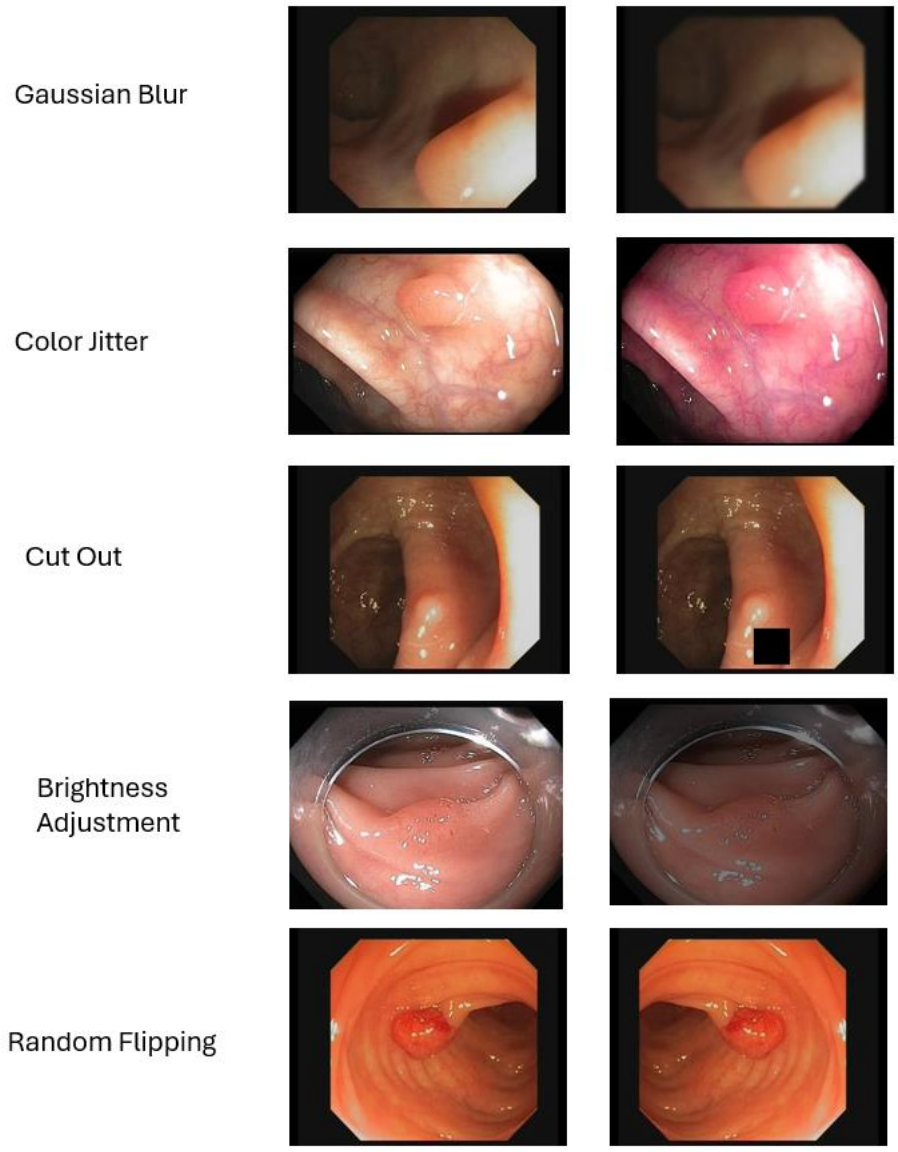

Figure 14: Other Data Augmentations: Before and After

Lastly, all these data augmentations, including the WLSR, were performed before the training. This was done to better compare with the WLSR since it cannot yet perform on-the-fly data augmentation.

3.5.2 Adding 30% Augmentation to Data

Data augmentations are often done to increase the training dataset size, not to replace it completely. Therefore, we must see how a model with WLSR performs as a data augmentation addition to the normal dataset. We chose a 30% data augmentation addition to the dataset and to train the model for 30 epochs. We selected 30 epochs because of the results from the last experiment, which we will show in the next chapter. Lastly, this test was also done on the other data augmentations to better compare and interpret.

3.5.3 False Positives from Specular Reflections

I wondered if the WLSR data augmentation would make the model more robust against specular reflections. Therefore, for this experiment, I first used the baseline model to detect all images in the training folder. This gave an output of images with bounding box detections made by the model. The images were then analyzed and compared to the images that have ground truth boxes. If there is a false positive as a result of a specular reflection, then we see if the model that includes 30% of WLSR from the last experiment made the same detection mistake.

3.5.4 Ensemble Augmentation

A model usually benefits more by providing it with a mix of data augmentations rather than just one. In this next experiment, we test if adding WLSR to a mix of data augmentations results in greater metric performance gains than not adding it to the mix. Since different data sizes and different sizes of data augmentations may mislead our interpretations, I compared two different ensemble models, one with the same data size as the ensemble model with WLSR and the other with the same data augmentation sizes. This relationship is better described in the tables below.

Table 6: Data Augmentation Ensemble Model with WLSR

Data Augmentation	Percent	Num of Images
Gaussian Blur	16%	4310
Flip	16%	4310
Brightness Adj.	16%	4310
Cut Out	18%	4849
Color Jitter	18%	4849
WLSR	16%	4314
Total	100%	26942

Table 7: Model with the Same Data Size as Model with WLSR

Data Augmentation	Percent	Num of Images
Gaussian Blur	19%	5118
Flip	18%	4849
Brightness Adj.	19%	5118
Cut Out	22%	5927
Color Jitter	22%	5930
Total	100%	26942

Table 8: Model with the Same Data Augmentation Size as Model with WLSR

Data Augmentation	Percent	Num of Images
Gaussian Blur	16%	4310
Flip	16%	4310
Brightness Adj.	16%	4310
Cut Out	18%	4849
Color Jitter	18%	4849
Total	84%	22628

The ensemble data augmentations were added to the normal dataset to increase the data size. Thus, the ‘Num of Images’ you see in the last column of tables 6 to 7 represents the additional images for the training dataset. So, the models represented by tables 6 and 7 had their data size multiplied by 2. Lastly, these models were each trained for 30 epochs.

3.5.5 Ensemble On-the-Fly Augmentation

So far, all data augmentations have been done before training. However, data augmentations tend to be more powerful when they can be created during training. This is called "On-the-Fly" augmentation. One reason why on-the-fly augmentation is better is that it can generate completely new augmentations on images per epoch, thus greatly reducing the chances of overfitting.

The WLSR data augmentation cannot do on-the-fly augmentation yet. But I wanted to test how it performs with the same concept of an ensemble of data augmentations, but now those data augmentations are performed during training as opposed to before.

Two models were created for this experiment, one without WLSR and one with WLSR, they were both trained for 30 epochs each. Both had the same ensemble of data augmentations with the same values. Table 9 shows the data augmentations used for this experiment.

Table 9: Data Augmentations Used for On-the-Fly Ensemble

Data Augmentation	Value	Description
hsv_s	0.1	Saturation Augmentation (+or - 10%)
hsv_v	0.1	Value Augmentation (+ or - 10%)
degrees	5	Rotation Augmentation (+ or - 5 degrees)
translate	0.1	Translate (shift any direction up to 10%)
scale	0.1	Scale (scale range from -10% to 10%)
fliplr	0.5	Flip left-right (50% chance)

CHAPTER IV

EXPERIMENTAL RESULTS

This chapter will discuss the experiment results. I mainly focus on mean average precision since I think it is an excellent metric for encapsulating the model's performance at different levels. However, for the sake of a more complete analysis, I will also add the F1 metric when needed.

An important calculation that I use to compare metrics of different models is percentage gain/ loss, also known as percentage change. If the value is positive, then it is a percentage gain. Otherwise, it is a percentage loss.

$$percentage\ change = \frac{(New - Old)}{Old}\ x\ 100$$

In the tables, I abbreviate percentage gain/loss as %G/L.

4.1 Replacing 100% of Data Results

In the following percentage gain/loss tables, I also add the mean average percentage gain/loss for every dataset size. That number will be used to analyze the performance trend across dataset sizes.

Table 10: WLSR Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	WLSR mAP@50	%G/L	mean %G/L
20%	10	0.819	0.797	-2.69%	
20%	30	0.78	0.789	1.15%	
20%	50	0.794	0.775	-2.39%	-1.308%
50%	10	0.846	0.813	-3.90%	
50%	30	0.782	0.779	-0.38%	
50%	50	0.756	0.823	8.86%	1.526%
100%	10	0.779	0.774	-0.64%	
100%	30	0.736	0.785	6.66%	
100%	50	0.786	0.809	2.93%	2.981%

Table 11: Flip Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	Flip mAP@50	%G/L	mean %G/L
20%	10	0.819	0.813	-0.73%	
20%	30	0.78	0.696	-10.77%	
20%	50	0.794	0.712	-10.33%	-7.28%
50%	10	0.846	0.795	-6.03%	
50%	30	0.782	0.733	-6.27%	
50%	50	0.756	0.719	-4.89%	-5.73%
100%	10	0.779	0.794	1.93%	
100%	30	0.736	0.772	4.89%	
100%	50	0.786	0.779	-0.89%	1.98%

Table 12: Color Jitter Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	Color Jitter mAP@50	%G/L	mean %G/L
20%	10	0.819	0.827	0.98%	
20%	30	0.78	0.765	-1.92%	
20%	50	0.794	0.79	-0.50%	-0.48%
50%	10	0.846	0.812	-4.02%	
50%	30	0.782	0.793	1.41%	
50%	50	0.756	0.798	5.56%	0.98%
100%	10	0.779	0.845	8.47%	
100%	30	0.736	0.809	9.92%	
100%	50	0.786	0.766	-2.54%	5.28%

Table 13: Cut Out Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	Cut Out mAP@50	%G/L	mean %G/L
20%	10	0.819	0.828	1.10%	
20%	30	0.78	0.803	2.95%	
20%	50	0.794	0.799	0.63%	1.56%
50%	10	0.846	0.811	-4.14%	
50%	30	0.782	0.814	4.09%	
50%	50	0.756	0.797	5.42%	1.79%
100%	10	0.779	0.826	6.03%	
100%	30	0.736	0.808	9.78%	
100%	50	0.786	0.794	1.02%	5.61%

Table 14: Gaussian Blur Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	Gaussian Blur mAP@50	%G/L	mean %G/L
20%	10	0.819	0.665	-18.80%	
20%	30	0.78	0.639	-18.08%	
20%	50	0.794	0.621	-21.79%	-19.56%
50%	10	0.846	0.639	-24.47%	
50%	30	0.782	0.667	-14.71%	
50%	50	0.756	0.648	-14.29%	-17.82%
100%	10	0.779	0.667	-14.38%	
100%	30	0.736	0.717	-2.58%	
100%	50	0.786	0.695	-11.58%	-9.51%

Table 15: Brightness Adjustment Percentage Gain/Loss for 100% Replacement

Dataset Size	Epochs	Baseline mAP@50	Bright Adj. mAP@50	%G/L	mean%G/L
20%	10	0.819	0.692	-15.51%	
20%	30	0.78	0.731	-6.28%	
20%	50	0.794	0.696	-12.34%	-11.38%
50%	10	0.846	0.757	-10.52%	
50%	30	0.782	0.718	-8.18%	
50%	50	0.756	0.663	-12.30%	-10.34%
100%	10	0.779	0.717	-7.96%	
100%	30	0.736	0.708	-3.80%	
100%	50	0.786	0.713	-9.29%	-7.02%

First, I want to point out that I know some data augmentations got percentage losses. This does not mean that those data augmentations are not useful. What this test indicates is how much the data augmentation changes the integrity of the polyp. For example, gaussian blur makes all polyps blurred in this experiment, then when it gets tested with the test dataset is having trouble making correct predictions because now it thinks all polyps are supposed to be blurry.

On the other hand, the WLSR is getting percentage gains even though it replaces the whole data. This is because this data augmentation does not change the integrity of the existing polyps. It adds new artificial lights, but it ensures that the new artificial lights do not overlap the polyps.

Now, analyzing these tables of the 100% replacement test, I noticed two different trends. These trends are visualized in the next two figures.

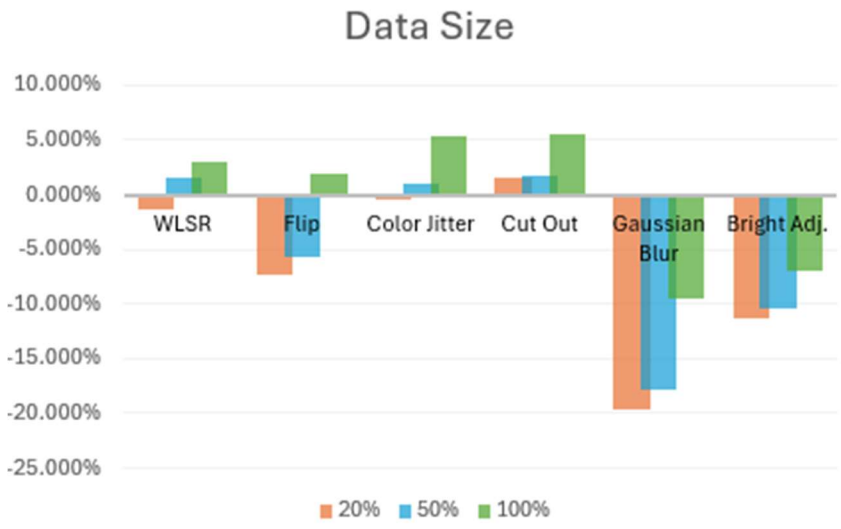

Figure 15: Mean Percentage Gain/Loss across Data Sizes

The first thing I noticed was that no matter the data augmentation, the mean percentage change was improving as the data size grew. This further emphasizes the importance of having more data to improve polyp detector metrics.

The next trend observation is visible across all different data sizes, but since we have witnessed that 100% data size gave the best results, the next trend will be visualized when the data size is 100%.

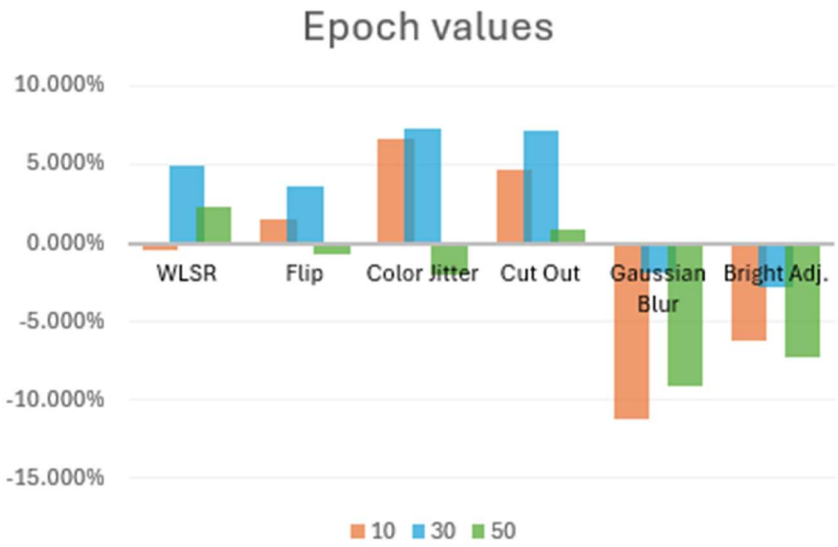

Figure 16: Mean Percentage Gain/Loss Across Various Epoch Values

We can observe from Figure 16 that the best percentage change results when epoch number is 30. This indicates that the model is overfitting when it is at epoch 50. Because of this reason, the next of the experiments are performed at 30 epochs.

4.2 Adding 30% Augmentation Results

Table 16: Metric Results of Adding 30% Data Augmentation

Data Augment.	Precision	Recall	F1	%G/L F1	mAP50	%G/L @mAP50	mAP50-95	%G/L @mAP50-95
Baseline	*0.887*	*0.552*	*0.681*		*0.736*		*0.511*	
G. Blur	0.836	0.702	***0.763***	**12.15%**	**0.801**	**8.83%**	0.551	7.83%
WLSR	0.821	0.693	*0.752*	10.45%	0.797	8.29%	**0.558**	**9.20%**
Flip	0.824	0.698	*0.756*	11.06%	0.784	6.52%	0.55	7.63%
Bright Adj.	0.815	0.688	*0.746*	9.64%	0.783	6.39%	0.534	4.50%
Cut Out	0.846	0.62	*0.716*	5.15%	0.758	2.99%	0.532	4.11%
Color Jitter	0.891	0.567	*0.693*	1.84%	0.744	1.09%	0.523	2.35%

The table above ranks the augmentations after the baseline based on their mAP50. All these models were trained for 30 epochs, including the baseline model. We can observe that the model that included 30% of Gaussian Blur resulted in the highest mAP50 and also the highest F1 score. However, it is important to observe that the model that added 30% WLSR resulted in the highest mAP50-95. This indicates that this model is better at detecting the hardest examples.

4.3 False Positives from Specular Reflections Results

The models used for this experiment were the baseline model and the WLSR model from the test above, which are trained for 30 epochs each.

I expected the baseline model to make more false positives from specular reflections than it did. Maybe the test dataset did not contain as many hard specular reflections. Nonetheless, when a false positive did occur due to the specular reflections, the WLSR model did not make that mistake. Figure 17 shows some of the results of this experiment.

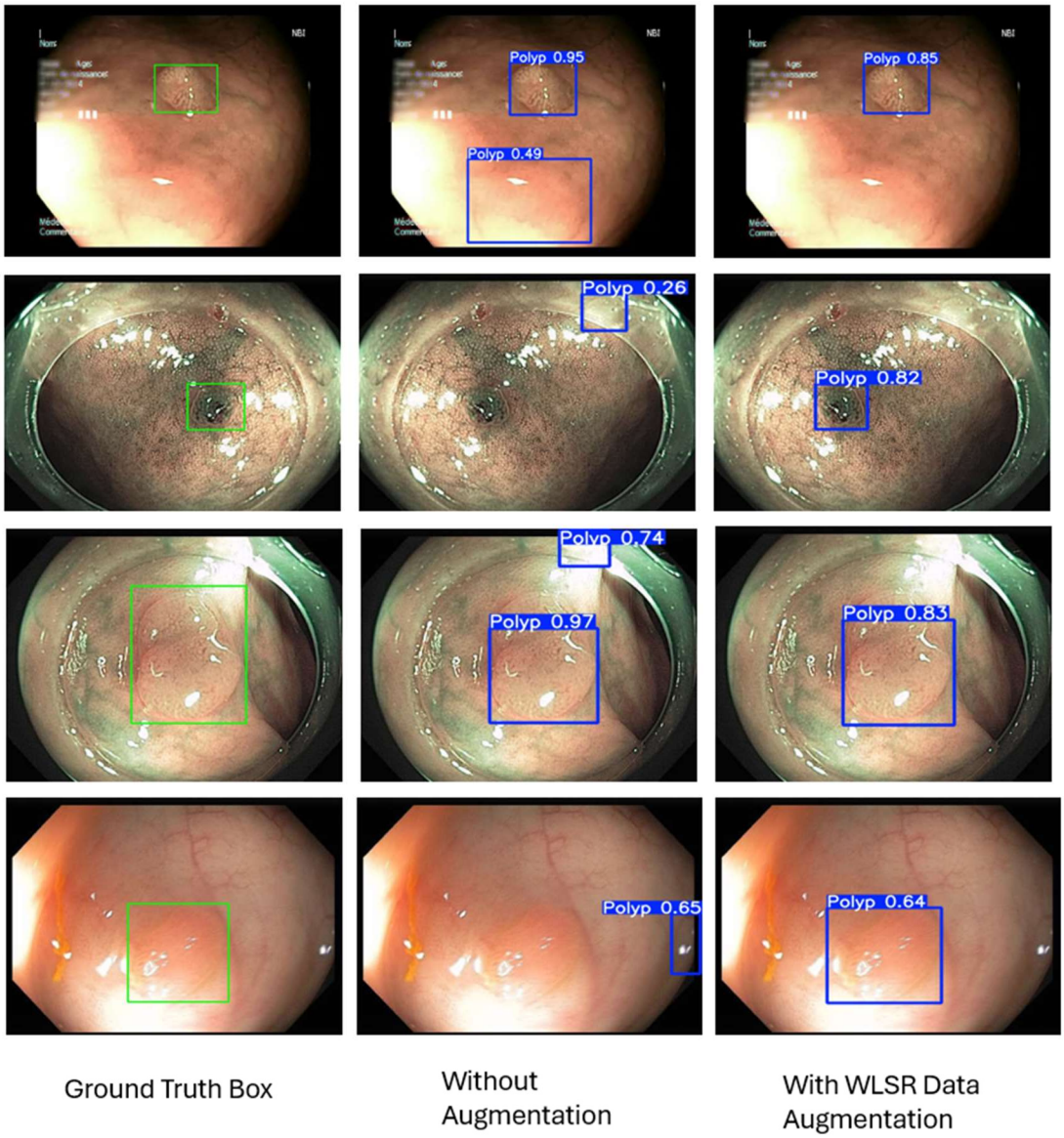

Figure 17: False Positives from Specular Reflections Comparison

This experiment indicates that training with WLSR data augmentation will make the model more robust against specular reflections.

4.4 Ensemble Augmentation Results

Here, we can see that combining data augmentations is better than just using a single one. Table 17 shows the performance metric results of this experiment. The models were abbreviated as follows:

- WLRS_E = Data Augmentation Ensemble method with WLSR included
- DA_E = Data Augmentation Ensemble method with the same Data Augmentation Sizes as WLRS_E
- DS_E = Data Augmentation Ensemble method with the same Data Size as WLRS_E

Table 17: Metric Results of Ensemble Data Augmentation Models

Data Augment.	Precision	Recall	F1	%G/L F1	mAP50	%G/L @mAP50	mAP50-95	%G/L @mAP50-
Baseline	*0.887*	*0.552*	*0.681*		*0.736*		*0.511*	
WLSR_E	0.838	0.723	***0.776***	14.07%	**0.818**	11.14%	0.566	10.76%
DA_E	0.809	0.737	*0.771*	13.35%	0.813	10.46%	**0.569**	11.35%
DS _E	0.859	0.667	*0.751*	10.35%	0.792	7.61%	0.558	9.20%

Something I noticed in this section is that model DA_E added the least of data augmented images and got the second best score in mAP50 and the best score in mAP50-95. I can speculate that maybe there is data augmentation or many data augmentations in the ensemble that may be hindering the growth or causing some sort a diminishing marginal returns.

Lastly, since these data augmentations are done before training, there is still a risk that the model may overfit the augmented training data. Furthermore, that is where on-the-fly data augmentation comes in. On-the-fly means that data augmentations are performed during

training. This means that the any data augmentation we have chosen will be able to generate new training images at every epoch, thus reducing the chances of overfitting greatly. In the next section, we will show the superior results accomplished using on-the-fly data augmentation.

4.5 Ensemble On-the-Fly Augmentation Results

This method resulted in the best model. The abbreviations for the following table are as follows:

- OTF = Model done with data augmentations performed on-the-fly but without WLSR.
- WLSR_OTF = Model was done with 30% of WLSR added to the training dataset and other data augmentations that were performed on the fly. Important reminder: WLSR was not performed on the fly; it does not have that capability yet. Only the data augmentations that were mentioned in Table 9 were performed on-the-fly.

Table 18: Metric Results of On-the-Fly Data Augmentation Models

Data Augment.	Precision	Recall	F1	%G/L F1	mAP50	%G/L @mAP50	mAP50-95	%G/L @mAP50-
Baseline	*0.887*	*0.552*	*0.681*		*0.736*		*0.511*	
WLSR_OTF	0.911	0.709	**0.797**	17.18%	**0.831**	12.91%	**0.589**	15.26%
OTF	0.911	0.667	*0.770*	13.17%	0.809	9.92%	0.576	12.72%

The best model acquired in this study is when combining on-the-fly data augmentation with WLSR. This model resulted in better performance metrics in F1score, mAP50, and mAP50-95 than the other models in this study.

CHAPTER V

DISCUSSION

5.1 Interpretation of Results

5.1.1 Intuition Behind WLSR

It never ceases to amaze me the powerful object detectors that humans are naturally born with. I think by understanding how that mechanism works in our brain, we can create better object detectors.

Recently, I had the opportunity to witness how powerful our innate object detector is. I have two black cats, and there are a bunch more black cats in my neighborhood. Initially, I thought it would be hard to recognize my own black cats from the neighborhood cats, but I noticed something surprising. In the beginning, I learned some physical attributes to tell my cats apart. However, as I encountered more black cats, or the environment got more challenging, like turning dark outside, my mind would find new physical traits for me to continue telling the cats apart.

Therefore, when I realized that many polyp detectors confuse white light specular reflections with polyps, I saw an opportunity to increase the performance metrics of polyp detectors. My intuition is that by adding more artificial lights, the model will be pushed to find new attributes (features) that will help it distinguish the polyps from the specular reflections.

I think that is one of the reasons why the data augmentation I proposed, provided great results.

5.1.2 Generative Adversarial Network (GAN) a Possible Approach

Since I obtained good results with my data augmentation, I think using a generative adversarial network (GAN) could provide even more challenging examples for polyp detectors. GANs are composed of two different neural networks. The main goal is to generate new data. This is accomplished by having the two neural networks compete with each other. One tries to generate new data, while the other tries to discriminate the generated data as real or fake. They go back and forth, with the objective of improving. Eventually, both get really good at their roles, and the results are new data that can be used for training purposes. I think using GAN to generate hard examples for polyp detector training can result in the models learning better features and thus improving our main goal of reducing missed polyps as much as possible.

5.2 Future Work

I would like to create version 2 of the WLSR data augmentation to see how much more it can improve a polyp detector.

The first thing I would like to see is if there are more improvements with more than one artificial light added to the images. In this paper, the WLSR only added one artificial light to the image. I am sure there is an optimal value to the number of artificial lights that can be added to produce the best results.

The second thing would be to expand the areas in which the artificial lights can be placed. In this paper, I used bounding boxes around prohibited objects to ensure the artificial lights would not be added there. I think doing segmentation instead of bounding boxes could help in expanding the available area to place the artificial lights. In addition, this version of

WLSR avoided any overlap of artificial lights with polyps. I think controlling to allow some partial overlap but not total overlap can provide further benefits to a polyp detector model.

The third and crucial improvement to WLSR would be to make it work during training instead of before training. The results section shows that on-the-fly data augmentation benefits object detector models greatly. To accomplish this I would need to modify the WLSR architecture to make it more efficient and fast. The orange border converter is what currently takes the longest time in the WLSR pipeline.

Lastly, incorporating GAN into WLSR for generating new and challenging artificial lights would provide great benefits to polyp detectors.

CHAPTER VI

CONCLUSION

Colorectal cancer is a world threat, and its incidence rate keeps rising. This type of cancer can be prevented by successfully detecting and removing bad polyps. Human error exists at any level of expertise, and missing a polyp can mean death. Artificial intelligence can be integrated into the goal of reducing missed polyp rates, which is done via object detectors. Object detectors that focus on detecting polyps are called polyp detectors.

Currently, many researchers are working on polyp detectors and attempting to improve their performance metrics. The important performance metrics are precision, recall, F1 score, and mean average precision (mAP). The goal is to increase those numbers as much as possible. With those metrics, more is better.

There are many challenges that polyp detectors face when trying to detect polyps. These challenges include polyps being too small, too flat, hidden in colon folds, and white light specular reflections caused by the endoscope.

This paper created a new data augmentation technique based on the knowledge that specular reflections are troublesome to polyp detectors. This technique, called WLSR, generates new artificial lights and randomly places them onto the images to provide a challenging environment, with the main goal of improving the polyp detector model.

The WLSR data augmentation method was tested under various conditions to determine its benefit to polyp detectors. With the 100% replacement test, the model with the WLSR data augmentation gave a mean percentage gain of 2.981% over the baseline model when the whole dataset was used with the following epochs: 10, 30, and 50. When only 30% of the training data was augmented by WLSR and the model trained for 30 epochs, it resulted in an 8.29% gain on the mAP50 metric over the baseline model. It ranked number 2 on that test among the other 5 data augmentations tested with the same conditions. In the ensemble comparison test, the model with WLSR resulted in the highest mAP50, reaching an 11.14% gain over the baseline model. In the On-the-Fly augmentation test, the model with WLSR was the best model in this paper. It reached an F1 score of 0.797 and mAP50 of 0.831. Lastly, it resulted in fewer false positives from white light specular reflections. Therefore, we can conclude that the WLSR data augmentation created for this thesis positively benefits the performance metrics of polyp detectors.

REFERENCES

Adegbulugbe, A. A., Farah, E., Ruan, Y., Yong, J. H., Cheung, W. Y., & Brenner, D. R. (2024). The projected health and economic impact of increased colorectal cancer screening participation among Canadians by income quintile. Canadian Journal of Public Health, 115(3), 384-394.

Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2017). Global patterns and trends in colorectal cancer incidence and mortality. Gut, 66(4), 683–691.

Bernal, J., & Histace, A. (Eds.). (2021). Computer-Aided Analysis of Gastrointestinal Videos (pp. 163-169). Springer.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

Chen, B. L., Wan, J. J., Chen, T. Y., Yu, Y. T., & Ji, M. (2021). A self-attention based faster R-CNN for polyp detection from colonoscopy images. Biomedical Signal Processing and Control, 70, 103019.

Doniyorjon, M., Madinakhon, R., Shakhnoza, M., & Cho, Y. I. (2022). An improved method of polyp detection using custom YOLOv4-tiny. Applied Sciences, 12(21), 10856.

Eileen, Morgan., Melina, Arnold., Andrea, Gini., Virgile, Lorenzoni., Citadel, J, Cabasag., Mathieu, Laversanne., Jérôme, Vignat., Jacques, Ferlay., Neil, Murphy., Freddie, Bray. (2022). Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. Gut, 72(2):338-344. doi: 10.1136/gutjnl-2022-327736

Favoriti, P., Carbone, G., Greco, M., Pirozzi, F., Pirozzi, R. E. M., & Corcione, F. (2016). Worldwide burden of colorectal cancer: a review. Updates in surgery, 68, 7-11.

Ghose, P., Ghose, A., Sadhukhan, D., Pal, S., & Mitra, M. (2024). Improved polyp detection from colonoscopy images using finetuned YOLO-v5. Multimedia Tools and Applications, 83(14), 42929-42954.

Joseph, C., Anderson., Douglas, J., Robertson. (2012). Overview of Colorectal Cancer. 1-28. doi: 10.1007/978-1-4614-5943-9_1.

Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., ... & Wei, X. (2022). YOLOv6: A single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976.

Liu, M., Jiang, J., & Wang, Z. (2019). Colonic polyp detection in endoscopic videos with single shot detection based deep convolutional neural network. IEEE Access, 7, 75058-75066

Morris, M., Iacopetta, B., & Platell, C. (2007). Comparing survival outcomes for patients with colorectal cancer treated in public and private hospitals. Medical Journal of Australia, 186(6), 296-300.

Nisha, J. S., Gopi, V. P., & Palanisamy, P. (2022). Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture. Biomedical Signal Processing and Control, 73, 103465.

Nogueira-Rodríguez, A., Domínguez-Carbajales, R., Campos-Tato, F., Herrero, J., Puga, M., Remedios, D., ... & Glez-Pena, D. (2022). Real-time polyp detection model using convolutional neural networks. Neural Computing and Applications, 34(13), 10375-10396.

Qadir, H. A., Shin, Y., Solhusvik, J., Bergsland, J., Aabakken, L., & Balasingham, I. (2019, May). Polyp detection and segmentation using mask R-CNN: Does a deeper feature extractor CNN always perform better?. In 2019 13th International Symposium on Medical Information and Communication Technology (ISMICT) (pp. 1-6). IEEE.

Ramon, Luengo-Fernandez., Jose, Leal., Alastair, Gray., Richard, Sullivan. "Economic burden of cancer across the European Union: a population-based cost analysis." Lancet Oncology, 14 (2013).:1165-1174. doi: 10.1016/S1470-2045(13)70442-X

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Siegel, R. L., Giaquinto, A. N., et al. (2024). Cancer Statistics. CA: A Cancer Journal for Clinicians, 74(1), 12-49. DOI: 10.3322/caac.21820

Tashk, A., & Nadimi, E. (2020, July). An innovative polyp detection method from colon capsule endoscopy images based on a novel combination of RCNN and DRLSE. In 2020 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-6). IEEE.

Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475).

World Health Organization. (n.d.). Colorectal cancer. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer

Yang, X., Song, E., Ma, G., Zhu, Y., Yu, D., Ding, B., & Wang, X. (2023). YOLO-OB: An improved anchor-free real-time multiscale colon polyp detector in colonoscopy. arXiv preprint arXiv:2312.08628.

Zhang, R., Zheng, Y., Poon, C. C., Shen, D., & Lau, J. Y. (2018). Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker. Pattern recognition, 83, 209-219.

VITA

In May 2017, Jose Angel Nuñez earned a Bachelor's Degree in Business Administration with a double major in Finance and Economics and a minor in Applied Mathematics from the University of Texas Rio Grande Valley (UTRGV). A few years later, he returned to UTRGV to earn a Master's Degree in Computer Science in August 2024.

His email is jnunez26.22@gmail.com