

# An autonomous colonoscopy vine robot for colorectal cancer surveillance and polyp removal

---

Matthew Teichman, Shawn Khan, Min Woo (David) Kong, Manthan Patel

## Abstract

There is a growing utility of artificially intelligent robotic systems in medicine to improve safety and efficiency, reduce costs, and increase access to care. This paper proposes a novel application of an autonomous colonoscopy robot. The process of automating conventional colonoscopy was deconstructed into seven individual tasks, including: (1) path planning and endoscope actuation, (2) localization and orientation of the endoscope, (3) maintenance of sufficient insufflation pressures, (4) assessment of bowel preparation adequacy, (5) tissue identification and localization, (6) polyp characterization, and (7) polyp manipulation and biopsy. The overall system implements a variety of AI frameworks, including Mask R-CNN, stereo SLAM, image processing and fuzzy logic, CNNs and RL techniques. The simulation results are reviewed and methods of improving performance are described. This report identifies the potential of a robotic colonoscope to successfully automate colorectal cancer surveillance.

## I. Project Description

The following report describes the application of artificial intelligence in the autonomation of colonoscopy for routine colorectal cancer surveillance.

## II. Project Statement

### Colorectal Cancer & Colonoscopy Screening

Colorectal cancer is the third most common cancer worldwide. It develops from precancerous lesions, known as polyps, that grow in the inner lining of the colon. Over the course of years, these lesions become gradually larger and can grow uncontrollably. Colorectal cancer can lead to devastating disease, spreading to the liver, lungs, bone, and brain.

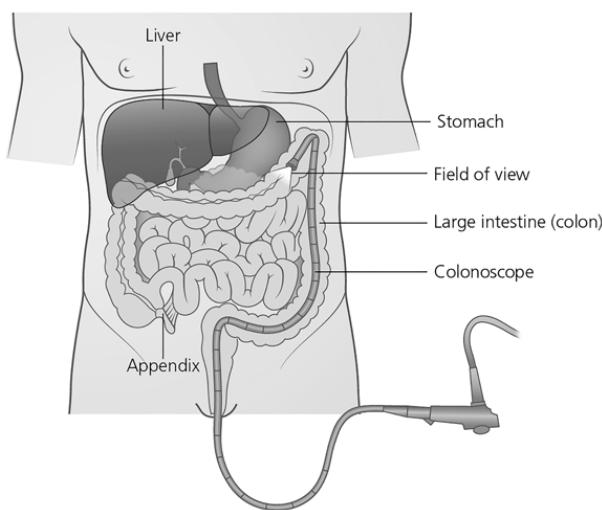
However, if caught early, the prognosis for colorectal cancer can be quite favourable. As such, screening is performed routinely in adults aged 50 - 75. Colonoscopy is the most effective method of identifying lesions and preventing colorectal cancer. It involves an operator inserting a flexible camera to examine the lining of the colorectal tract (Figure 1)[1]. As a result of colonoscopy screening, rates of colorectal cancer have been declining. However, more recently, the prevalence of colorectal cancer has been rising in younger adults[2].

Further, in low resource areas, it can be particularly difficult to undergo colonoscopy due to a shortage of skilled operators—contributing to suboptimal screening rates. Rates of surveillance can be particularly low among underserved and vulnerable populations given lack of access to screening[2].

## Robotic Colonoscopy

The following proposal describes an autonomous colonoscopy robot that identifies suspicious lesions in the colorectum.

The goal of the robot will be to improve accessibility to routine colonoscopy and improve detection and safe removal of lesions suspicious for pre-malignant colorectal polyps and/or masses. The autonomous robotic colonoscopy system is proposed to function in hospitals and/or outpatient clinics.



**Figure 1.** Diagram of colonoscopy showing path through large intestine (colon)

## Steps of Automation

The overall task of colonoscopy requires complex mechanisms, including multiple degrees of motion and real-time adjustments by a human operator. Each fundamental component of the colonoscopy process, and the proposed method of AI automation, is outlined in Table 1.

Of note, the proposed robot does not currently automate the process of maintaining insufflation pressures or currently assess adequacy of bowel preparation. These were not accomplished due to time restraints and limitations in available data. As such, this report does not describe the process of implementing fuzzy logic, as initially proposed. This is further described in the Results section.

## Assumptions

In order to ensure successful robotic colonoscopy, the following assumptions will be made:

- The patient will be positioned in the standard left lateral decubitus position
- The patient will have undergone adequate bowel preparation
- The patient has no past history of colorectal surgery (ie. ileoanal anastomosis) or active inflammation/infection (ie. acute diverticulitis)

**Table 1. Description of tasks the proposed robot will perform autonomously with AI**

Task	Goal of Task	Role of AI	Significance of Task	Completed (Y/N)
Path planning and actuation of endoscope	To ensure safe movement within the colon. It will specifically travel in the following order: <ul style="list-style-type: none"> <li>• Anus</li> <li>• Rectum</li> <li>• Sigmoid colon</li> <li>• Descending colon</li> <li>• Transverse colon</li> <li>• Ascending colon</li> <li>• Cecum</li> </ul>	Mask R-CNN	Mask R-CNN will be used to inform the direction of autonomous movement, such as turning, without collision and/or injury to the bowel wall.	Y
Localization and orientation of endoscope	To understand where the robot is moving in relation to the colon	Stereo SLAM	Allows for accurate localization and path planning, mapping the colon	Y
Maintenance of sufficient insufflation pressures	To maintain steady insufflation pressures in the colon to allow for adequate visualization	Image Processing and Fuzzy Logic	Poor insufflation pressures preclude adequate visualization while high insufflation pressures can result in bowel wall injury. As such, the robot will need to maintain an optimal balance.	N
Assess adequacy of bowel preparation	To ensure patients have undergone bowel preparation to clear the bowel of the stool and ensure sufficient visualization and safety of the procedure.	Image Processing and Fuzzy Logic	If the amount of stool is over a percentage point, the robot will conclude that the bowel is not adequately prepared and exit the colon.	N

Tissue identification and localization	To identify and localize lesions discrepant from normal anatomy (polyps, lumen, stool, cecum, appendiceal orifice, ileocecal valve)	Mask R-CNN	The most important indicator of a good quality colonoscopy is the polyp detection rate, which is defined as the proportion of colonoscopies where a polyp was detected in relation to all the screening colonoscopies the operator performed.	Y
Characterization of polyp	To indicate whether a lesion is likely to be cancerous, precancerous, or benign (non-malignant)	CNN	Cancerous polyps tend to have a unique appearance--such as larger size, irregular borders, ulceration, and a flat appearance. Once a polyp is identified, the ability to characterize the polyp as having a low or high risk for cancer will inform which polyps are prioritized for biopsy.	Y
Polyp manipulation and biopsy	To allow for biopsy and extraction of polyp	Mask R-CNN	Biopsy of tissue allows further clarity regarding tissue histology and risk of cancer, informing need for surgical management	Y

### III. Background

#### 1) Path planning and actuation of endoscope

In the table below, different technologies for colonoscopy are compared based on their respective mechanisms, benefits, and drawbacks. Of note, autonomous magnetic colonoscopy is not

performed commercially due to its relative novelty, its resource intensiveness, and its corresponding lack of existing infrastructure. This proposal outlines the development of a novel AI-driven autonomous colonoscopy system that uses pressure-driven eversion (“vine robotics”).

**Table 2. Literature comparison of robotic endoscope technologies [7, 8]**

Technology	Description	Advantages	Disadvantages
Mechanical Endoscope (Conventional)	Insertion of a flexible tube with camera attached on its tip for getting vision inside the colon. If required, special tools are inserted via a flexible tube for polyps removal	<ul style="list-style-type: none"> <li>Entire colon view is possible compared to methods other than colonoscopy</li> <li>Biopsy and polyps removal is achievable</li> <li>Already in use commercially</li> </ul>	<ul style="list-style-type: none"> <li>Risk of injury to bowel (ie. perforation, bleeding)</li> <li>One to one movement tracking is difficult</li> <li>Accuracy of polyp detection dependent on skill of human operator</li> </ul>
Magnetic Endoscope	Enabling conventional endoscopy with magnetic control, the endoscope's magnetic tip can be manipulated by movement of a large magnet outside.	<ul style="list-style-type: none"> <li>Autonomous manipulation prevents chances of human error</li> <li>Biopsy and polyps removal is achievable</li> <li>Highly trained personnel are not required</li> </ul>	<ul style="list-style-type: none"> <li>Not as well studied in comparison to conventional endoscopy</li> <li>Magnetic manipulation may be considered complex and unintuitive</li> <li>Requires the adoption of significant infrastructure</li> </ul>
Vine-Robot Endoscope	Proposed design for autonomous colonoscopy robot, using pressure-driven eversion. This design grows from its center and can be tuned to follow a trajectory. As a soft robot, there may be less risk of bowel perforation.	<ul style="list-style-type: none"> <li>Autonomous operation along with material flexibility</li> <li>Biopsy and polyps removal is achievable</li> <li>Unlike magnetic endoscopes, no extra cost of magnetic devices is required</li> </ul>	<ul style="list-style-type: none"> <li>Has not been studied in field of endoscopy</li> <li>Modeling of the mechanics and feedback control is difficult given the novelty of application</li> </ul>

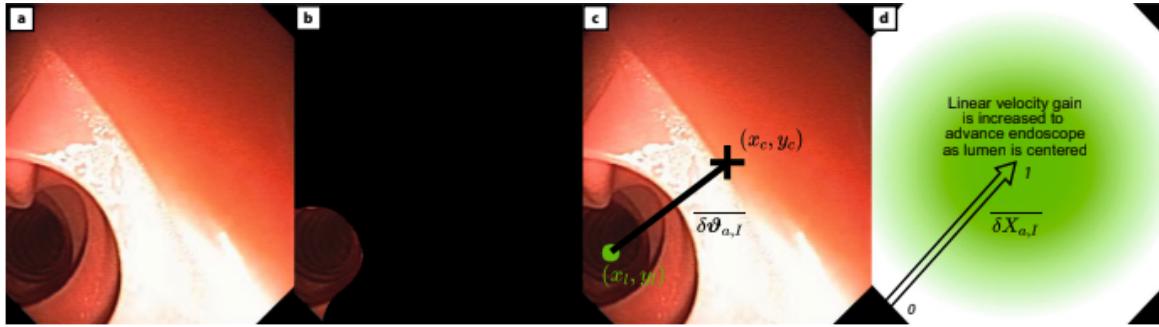
In comparison to the conventional devices, a colonoscopy that functions through pressure-driven eversion (“a vine robot”) may offer several unique advantages. In particular, vine robots are made of soft and flexible material and they allow for adjustable autonomous

movement. Unlike other endoscopes, vine robots increase in length from the centre of the device. As such, with advancement into the colon, vine robots would apply pressure uniformly on the colonic walls rather than applying focal pressure--decreasing patient discomfort and potentially reducing risk of bowel perforation.

Further, no extra pressure is required to inflate the bowel as it can be implemented by inflating the vine robot itself.

Vine robot localisation algorithm will process images captured from a camera as shown in Fig 1a. Using segmented Mask from the Mask R-CNN it will locate the lumen(Fig 2b.)(The lumen is the inside space of the colon. Stool passes through the colonic lumen during bowel

movements; polyps grow inwards into the colonic lumen.) and will give us a positional difference between lumen and current center of endoscope focus (Fig 2c). Then it will use this positional difference to generate a vector which endoscope will follow to recentre with lumen as shown in Fig 2d. Once the lumen is centered, the Vine robot endoscope will increase its velocity to grow ahead.



**Figure 2.** Example of Lumen Detection and Path Planning [7]

## 2) Localization and orientation of endoscope

The proposed robot needs localization in order to understand its orientation and position within the colon. Localization allows for safe navigation as well as the ability to accurately manipulate objects in the environment.

Simultaneous Localization and Mapping (SLAM) technology has been selected to determine the position and the path of the endoscope in an uncertain environment. Many different types of SLAM are available today; in particular feasibility studies of Mono, Stereo, Sonar, and Lidar SLAM had been conducted to conclude with the Stereo. Please refer to Table 4. In summary, Mono is quite

flexible but is associated with significant noise which makes it difficult to control with AI. Stereo SLAM could be used in any kind of environment condition whereas Lidar is largely optimal under dry conditions. Stereo SLAM also allows for a higher range of structures that can be visualized while Sonar and Lidar are more limited to specific distances.

In order to implement Stereo SLAM, the use of stereo video datasets for the colon will be required to generate the 3D images of the colon. Additionally, the calibration process of two cameras needs to be carefully performed in order to improve robot movement and detection of objects in the colorectum.

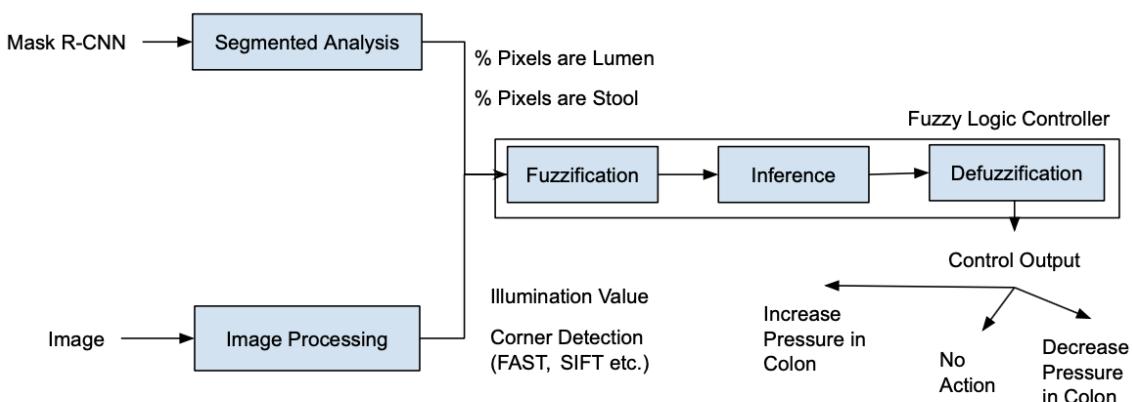
**Table 3. Comparison of SLAM technology identified in the literature**

Types of SLAM	Environment Condition	Range	Selection
Mono [10]	Any	0.5 - 5.0m	N
Stereo [11,12]	Any	Dependent on baseline distance between 2 cameras (mm to cm)	Y
Sonar [11]	Any	Long	N
Lidar [11-13]	Dry	Short	N

### 3) Maintenance of sufficient insufflation pressures

During colonoscopy, physicians use air to inflate the colon. This opens the colon cavity to allow the endoscope to move freely through the body. The physician uses judgement from visual inspection from the endoscope video feed and tactile feeling of the abdomen to determine if the colon is sufficiently inflated. Papers discussing autonomous endoscopes fail to understand the importance and difficulty of the task. As a result, many of the autonomous endoscope robots proposed in academia are impractical as they assume the colon is already inflated.

As a result, our group proposed a novel solution for determining if the colon is sufficiently inflated using a classical robotics AI approach. Data from the segmented mask and the original image is collected and fused together to generate a decision on whether pressure in the colon is required to be modified. The data fusion and decision-process is carried out on the fuzzy logic controller, as shown below in Figure 5. The following parameters such as illumination value, number of corners, percentage of lumen pixels and stool pixels were considered.



**Figure 3.** Fuzzy logic controller to determine colon pressurization

The approach does not require training data, but rather modification of the fuzzy logic inference rules. This can be done by trial and error using medical judgement to modify the rule set, until a

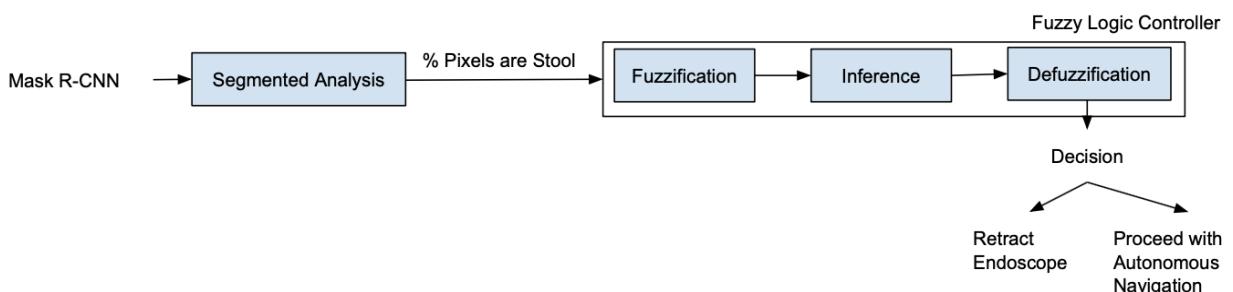
desired behavior is achieved. The fuzzy logic controller must be able to handle the following scenarios;

- 1) Colon has insufficient pressure. As a result, the robot is unable to move forward. The correct action of the robot is to inflate the colon with air.
- 2) Colon has too much air, and the lumen is too large. The corresponding action of the robot would be to decrease the pressure.
- 3) Colon has the correct amount of pressure. The correct action of the robot is to take no action.
- 4) Endoscope is facing the colon wall and not the lumen. The robot cannot assess whether or not the pressure is inadequate and as a result takes no action.

#### 4) Assess adequacy of bowel preparation

Throughout the operation of autonomous colonoscopy the robot must continuously assess whether or not the bowel is adequately prepared. An inadequately prepared colon caused much stool in the large intestine. Too much stool in the colon can impact the robots ability to move through the colon and possibly generate friction between device and colon wall, risking the possibility of perforating the wall.

To prevent the robot from perforating the colon wall, a fuzzy logic controller will be designed to determine if it is safe to continue operation of the robot as shown below. The fuzzy logic controller uses a segmented mask from the Mask R-CNN, to determine if the percentage of stool in the colon is above a predefined rule-set programmed in the fuzzy logic controller.



**Figure 4.** Fuzzy logic controller to determine bowel preparation

#### 5) Tissue identification and localization

Tissue identification and localization is central to the functioning of the robot; many of the following sequential tasks of the project will depend on performance of image recognition AI. The following table below summarises the various techniques used in academia to detect and localize polyps in the colon (Table 2).

Based on the literature search, Mask R-CNN was identified as the most optimal framework to accomplish the target goal, as it provides accurate

instant segmentation masks for localized colonic tissue. The Mask R-CNN networks are also the fastest of the R-CNN type architectures, which can provide the best real-time performance while allowing for mask segmentation. This is needed as the project utilizes transfer-learning to accelerate learning and reduce dataset sizes.

It was initially designed to consist of six labelled categories: polyps, lumen, stool, cecum, appendiceal orifice, and ileocecal valve (see Terminology). The cecum, appendiceal orifice, and ileocecal valve are used as anatomical features

to mark the end of the large intestine during traditional colonoscopy. However, there was a paucity of colonoscopy images with these anatomical features. As such, the Mask R-CNN was only trained to identify polyps, stool, and the cecum. Identification of lumen was used for path planning as described in section I.

When training this AI system, we sought to balance data generalizability with optimal fit. Specifically, to reduce overfitting the network, it was proposed that an equal proportion of images will be used to represent the dataset.

**Table 4. Literature of current techniques used in detection and localization**

Paper	AI Architecture	Datasets	Results (Polyp Detection and Localization)	Disadvantages
Artificial Intelligence and Polyp Detection in Colonoscopy [3] (2021)	YOLOv3	Training set of 6,038 images with 2,571 polyps  Tested on 10 unseen videos	Sensitivity: 74.1%  Specificity: 85.1%  F-Score: 83.1%	Does not provide accurate detection and localization compared to R-CNN architectures.
Detection of elusive polys using a large-scale artificial intelligence system [4] (2021)	Using both RetinaNet and LSTM-SSD Networks in parallel  Classification and Box Localization	Training set of 3,611 hours of video or 80 million frames  Testing set of 1393 hours of video	Sensitivity: 97.1%	Boundary boxes are not sufficient to provide accurate information for robot path planning, which is important for navigation and grasping in later tasks.
Computer-aided detection of colorectal polyps using a newly generated deep convolutional neural network [5] (2021)	DCNN uses VGG16 for classification and single-shot multibox detector (SSD)	Training set of 116,529 images  Testing set of 15,534 video frames	Sensitivity: 90%  Specificity: 80%	DCNN uses a multi consecutive frame to extract features and fuse results together. These are generally difficult to develop due to specialized datasets and software.
Automatic detection and	Mask R-CNN	Training set of 3,045 images with	Average precision of 89.5%	The Mask R-CNN does not provide good real-time

segmentation of adenomatous colonoscopy using Mask R-CNN [6] (2021)		3,045 images of polyps.  Testing set of 330 images with 228 images of polyps		performance as most networks are limited to 5 FPS.
---	--	--	--	--

## 6) Characterization of polyp

On appearance alone, operators can appreciate features that are suggestive of malignancy, such as polyp size, irregular borders, ulceration, and the presence of a short and immobile stalk. As such, the proposed autonomous colonoscopy will aim to characterize whether a polyp likely represents harmless tissue or a high suspicion for malignancy.

In particular, we will seek to identify existing datasets to train our AI using CNN on the features concerning potential malignancy to generate confidence intervals of the likelihood of malignancy as well as size measurements, such as depth.

In doing so, polyps with a higher risk of malignancy will be prioritized for excision and biopsy. If the robot identifies very low to no risk of malignancy, a biopsy of the polyp will not be performed. However, if there is a high risk of malignancy, the robot will record the location of the polyp in the colorectal tract and attempt to excise the polyp for biopsy.

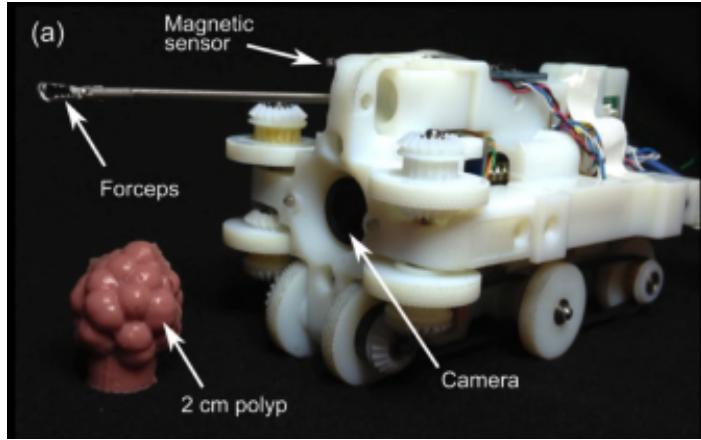
A study by Abdelrahim et al. used image processing to group polyps as category A ( $\leq 5\text{mm}$ ) or category B ( $> 5\text{mm}$ ). They found that real-time automated polyp sizing when combined with

AI-assisted polyp characterisation could improve polyp management strategies.

## 7) Polyp manipulation and biopsy

Performing polyp manipulation and biopsy is not a trivial task as polyps come in a variety of colors, shapes and anatomical features. As a result, there is no uniform orientation to grasp and manipulate polyps. Additionally, there are no commercial products available to perform autonomous polyp removal. There is only one published academic approach, to automating the task of colonoscopy polyp removal which was done by Q. Zhang et al at Southeast University [9].

The research paper presents an autonomous polyp biopsy using robotic endoscope platform (REP). The robotic platform can be seen in the image below. The platform uses a set of wheels and tracks to navigate autonomously through the colon to the location of the polyp. The REP system uses forceps to take Robot uses a single pinhole camera model to locate and track the polyp and series of magnetics sensors to localize the robot and provide a frame of reference. A magnetic sensor is placed on the robotic platform to track the position of the robot using an external magnetic field. Additionally magnetic sensors are biopsy forceps which are used to provide pose measurements of the forceps [9].



**Figure 5.** Autonomous robotics endoscope platform (REP) by Zhang et al [7].

The REP had a number of flaws including the removal success rate of polyps. For polyps of 1 cm the success rate was only 43% and for 2 cm polyps it was only 67%. The robot made use of a single pinhole camera, as a result in order to get depth measurement, the robot image took samples from multiple views and used multiple view geometry mathematics to estimate depth measurement of polyps. These depth measurements were often inaccurate and resulted in forceps under or overshooting the polyp. The REP had a number of issues driving through the colon folds. The robot utilized a wheel and track hybrid system which resulted in its driving issues [9]. Lastly, the autonomous system proposed was hard-coded and modeled to remove polyps that matched the reference target polyp shown in Figure 7. In a practical clinical environment polyps are different anatomical structures including a variety of colors and sizes [9].

The project will propose the following enhancements to improve and build on the REP solution [9]. The robotic endoscope in our project will be localized using SLAM techniques and will generate better pose estimates of the robot's position in the colon. Stereo cameras will be used to determine more accurate depth measurement; this will correlate to more precise and accurate manipulation and control of the biopsy tools. The mask R-CNN will be integrated into a robotic endoscope to detect polyps with different

anatomical features, colors and sizes. The neural network segmentation mask will also be used to determine centroid of mass, to improve grasping and manipulating the polyp. The robotic platform will be replaced by the vine robot to improve movement through the colon as proposed in task 2. The model of biopsy tools will be utilized for simulation requirements and trajectory equations of motion will be modified to fit the biopsy tool selected. Classical feedback controls will be used to guide the biopsy tool to the polyp.

Lastly, the biopsy tool will be replaced with a suction device to retrieve the specimen. The suction tool both excises and retrieves the polyp specimen and will be incorporated into the model.

#### IV. Proposed Design Methodology

Due to the complexity of conventional colonoscopy, the procedure was itemized into seven individual tasks, as outlined above in the project statement. These include:

- 1) **Path planning and actuation of the endoscope**

Path planning and actuation refers to the process of traversing the loops of the colon from the rectum to the cecum. This process was demonstrated by simulating the movement of the colonoscopy camera (endoscope) through the

colon in the two dimensional space.

Path planning and actuation is important in the development of autonomous colonoscopy to ensure operation safety. In conventional colonoscopy, there remains a risk of perforating the bowel wall. As such, an autonomous colonoscopy robot requires the capacity to both predict its trajectory and make adjustments to the dynamic environment of the colon to minimize risk of injury.

As such, a virtual Python game was developed to train the process of planning a trajectory and adjusting movement to avoid hitting the colon walls.

## Simulation of Path Planning

### (i) Environment

In particular, to simulate the process of autonomous navigation, we developed a Pygame module. We created a sample image of a colon using a 700 x 700-pixel image. In this image, the complex colonoscope was simplified to a cursor (red) that represented the endoscope head. In this environment, the main controls of the endoscope included forward velocity, backward velocity, clockwise angular velocity, and anti-clock wise angular velocity, which are controlled by keyboard arrow keys. These attempted to model the degrees of freedom of a colonoscopy in the two dimensional space. The algorithm was designed such that if the endoscope head touched the wall of the colon, the simulation quit and the robot was reset to its original starting position.

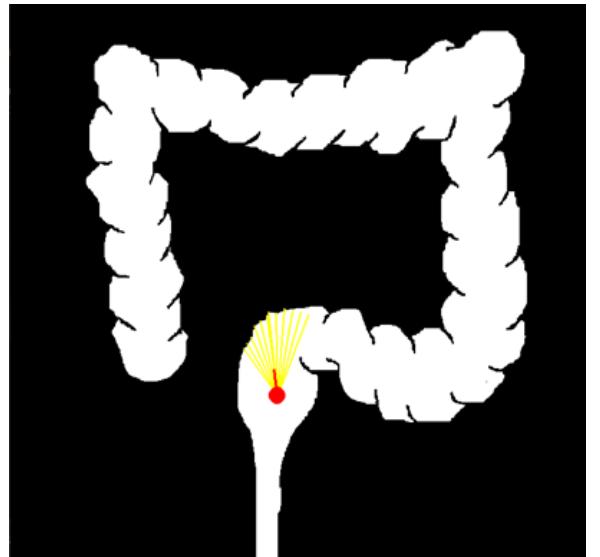
### (ii) Path Planning Algorithm [19]

The colon environment is dynamic as it changes frequently. As such, reinforcement learning was thought to best support the process of autonomous navigation and path planning.

In particular, we used a Deep Q Network for reinforcement learning. For the Q-function, we used a 3-layer neural network with a ReLU activation function between each of the layers.

$$Q(s, a) = r(s, a) + \gamma \max_a Q(s', a)$$

The above equation describes the output of the  $Q$ -value from being at state  $s$  and performing action  $a$ , yielding the immediate reward  $r(s, a)$ , added to the highest  $Q$ -value possible from the next state  $s'$ . The gamma in this equation represents the discount factor which controls the contribution of rewards further in the future.



**Figure 6.** Path planning Pygame

$Q(s', a)$  again depends on  $Q(s'', a)$  which will then have a coefficient of gamma squared. This means that the  $Q$ -value depends on the  $Q$ -values of future states, as shown here:

$$Q(s, a) \rightarrow \gamma Q(s', a) + \gamma^2 Q(s'', a) \dots \dots \dots \gamma^n Q(s'' \dots a)$$

Under these conditions, adjusting the value of gamma shapes the contribution of future rewards.

Since this is a recursive equation, we can start

with making arbitrary assumptions for all  $Q$ -values. Through iterative training cycles, it is expected that the algorithm will converge to the optimal policy. In practical situations, this is implemented through the update:

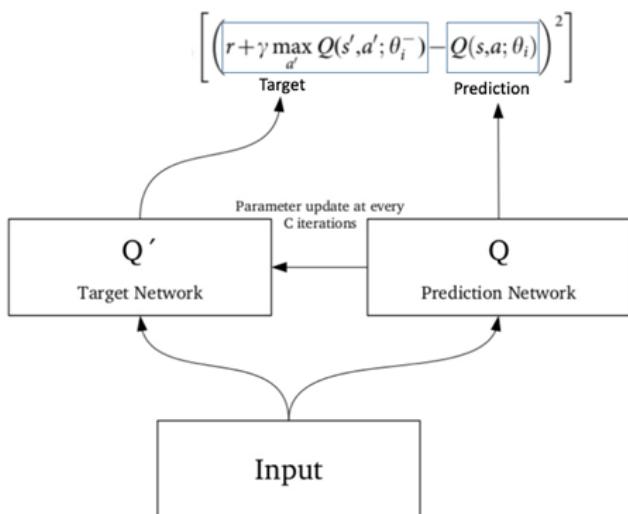
$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$$

where  $\alpha$  represents the learning rate, a hyperparameter. This simply determines the extent to which newly acquired information overrides old information.

### (iii) Agent

The RL agent uses two  $Q$ -learning models: (1) the local model and (2) the target model. The local model receives updates after every experience, while the target model is trained after a certain number of experiences which is a hyperparameter. After a certain number of experiences, the target model undergoes a soft update of its parameters. When the target model is trained, it uses a fixed number of experiences randomly from its memory buffer.

The agent uses an Epsilon-Greedy policy to update the model each time. In the initial stages, an agent will explore randomly. However, over many training cycles, its past experiences will affect its future actions, as shown in Figure 7.



**Figure 7.** Reward Function[17]

### (iv) Reward Function

The application of Reinforcement Learning for path planning in the human colon is novel. As such, no comparable reward functions exist. This required the development of a new reward function, shown below:

**if collision\_with\_colon() == True:**

**reward = -10**

Here, we assess for collisions between the endoscope and the bowel wall. If collisions occur, the path is penalized and given a negative reward.

**if 158<=xpos<=178 and 72<=ypos<=492**

**reward=100**

If the endoscope reaches the end of the colon, the path receives a large positive reward, resulting in the termination of the function.

**Else:**

**reward=2(abs(xpos-320)+abs(ypos-500))**

For intermediate stages, the reward is the function of the travelled distance of the endoscope with its initial position. This attempts to incentivize the process of achieving the maximum reward as the endoscope travels further towards the end of the colon.

## 2) Localization and orientation of endoscope

Localization and orientation of the endoscope refers to the process of situating the camera in relation to the bowel walls. The SLAM algorithm uses a graph bundle optimization approach to generate estimates of probes pose as well as building 3D maps of internal tissues. The stereo

images were taken from unity simulation provided by the open source community [14]. The camera stereo intrinsic parameters were not provided along with images as result values were estimated using an iterative experimentation process.

The SLAM algorithm used in the demonstration was created through open source code and was recreated based on a paper by Taihú et al. [15]. The SLAM algorithm utilizes graph optimization known as g<sup>2</sup>o to perform bundle adjustment removing noise in the estimated pose and map [21-22].

### **3) Maintenance of sufficient insufflation pressures**

To demonstrate the autonomous maintenance of colon pressure, test images of real colonoscopy procedures with varying levels of insufflation pressures will be analyzed. The simulator will analyze the segmented mask from the mask R-CNN and the original image to determine if pressure needs to be increased or decreased in the colon cavity. In particular, with low colon pressures, visualization of the colon will be poor and the AI will be trained to increase pressure. Meanwhile, high pressures within the colon cause the walls to be distended, increasing the risk of perforation. As such, the AI will be trained to lower pressure.

The proposed system will utilize fuzzy logic to fuse associated data to determine the appropriate action of the robot. Successful completion of simulation will be considered when the AI agent is able to recommend the correct course of pressurization to inflate the colon for all provided cases.

### **4) Assessing adequacy of bowel preparation**

To demonstrate the task of assessing bowel

preparation, a simulator will be created to demonstrate AI agent ability to determine if the bowel is adequately prepared. Simulation uses segmented images from mask R-CNN and determines if the current image scenario is too dangerous to proceed with robot movement due inadequate bowel preparation. Successful completion of simulation will be demonstrated when the AI agent can correctly determine if it is safe for the robot to continue with the procedure for a variety of use cases.

### **5) Tissue identification and localization**

To demonstrate tissue identificatification and localization test video data of real colonoscopy procedure will be inputted into the mask R-CNN network. The results from the network predictions including classification prediction, bounding boxes and instance segmentation mask will be overlayed onto the original video data. Additionally, using professional annotated video data, ground truth will be compared to output of the network, generating statistical analytics on performance.

The mask R-CNN used pretrained resnet50 as a feature map. Resnet 50 was chosen due past implementations have successful results in polyp segmentation [6].

In order for mask segmentation, images were retrieved from online published datasets and labelled using open source software called Label Studio (<https://labelstud.io/>). Following, data were exported on COCO and loaded to the pre-trained model. Python scripts were created to sort and generate masking data.

## 6) Characterization of polyp

The proposed robot will examine the physical features of polyps to generate probabilities of the likelihood of the lesion being cancerous. If an identified lesion has no suspicion of it being a cancerous polyp, the robot will not biopsy a polyp. However, if there are lesions with concerning features for malignancy, those lesions will be prioritized for excision.

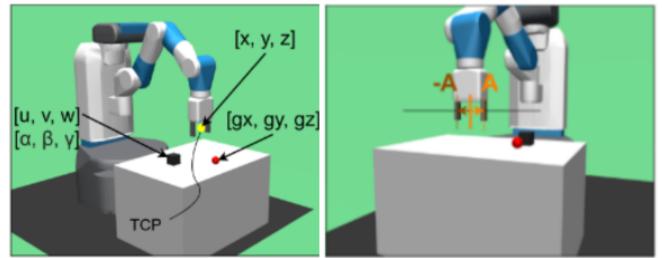
The CNN model implemented for polyp characterization was modified from Google's inception v3 network [23]. It was used based on previous successful implementation on melanoma skin cancer detector proposed by Thurn 2017 [24]. The fully connected layers were modified from 1000 output neurons to 3 output neurons matching the classification labels.

## 7) Polyp manipulation and biopsy

In the absence of an *in vivo* model, polyp manipulation and biopsy was modelled through an open source simulation environment (<https://gym.openai.com/>). In particular, the conceptualized robot was simulated using a robotic arm and the target colorectal polyp was simulated with a three-dimensional box.

Simulation training generated poses of the robot arm to train the trajectory of polyp grasping and the biopsy. Specifically, the robot arm improved its error through Reinforcement Learning (RL) which was chosen to accurately model the complex and variable nature of colorectal polyps.

The degrees of motion of the robotic arm and the simulated polyp, represented by the black box, are shown in Figure 8.



**Figure 8.** Robot arm states. Black box represents an object of interest, modelling the colorectal polyp. Red sphere represents the target location.

**Table 5. Robot arm states [24]**

	States
Goal position	$[gx, gy, gz]$
Achieved goal	$[u, v, w]$
TCP position	$[x, y, z]$
Object position	$[u, v, w]$
Object relative position	$[u - x, v - y, w - z]$
Gripper state	$[A, \bar{A}]$
Object rotation	$[\alpha, \beta, \gamma]$
Object velocity	$[\dot{u}, \dot{v}, \dot{w}]$
Object relative velocity	$[u - x, v - y, w - z]$
TCP velocity	$[\dot{X}, \dot{Y}, \dot{Z}]$
Gripper velocity	$[-\dot{A}, \dot{A}]$

TCP, Tool Center Point

In order to improve performance, the system yielded the reward function with interaction of the simulated polyp.

The overall policy optimization framework can be described with three principles, including: (1) increasing the probability of actions to have a higher return, (2) decreasing the squared residual error to stabilize training, and (3) receiving optimal rewards through the policy gradient.

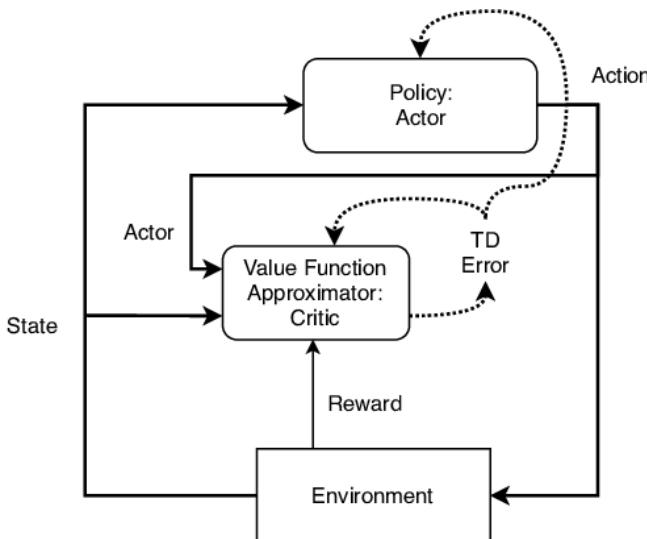
**Table 6. The Soft Actor-Critic algorithm used for policy optimization [17].**

**Algorithm 1** Soft Actor-Critic

```

Initialize parameter vectors  $\psi, \bar{\psi}, \theta, \phi$ .
for each iteration do
    for each environment step do
         $\mathbf{a}_t \sim \pi_\phi(\mathbf{a}_t | \mathbf{s}_t)$ 
         $\mathbf{s}_{t+1} \sim p(\mathbf{s}_{t+1} | \mathbf{s}_t, \mathbf{a}_t)$ 
         $\mathcal{D} \leftarrow \mathcal{D} \cup \{(\mathbf{s}_t, \mathbf{a}_t, r(\mathbf{s}_t, \mathbf{a}_t), \mathbf{s}_{t+1})\}$ 
    end for
    for each gradient step do
         $\psi \leftarrow \psi - \lambda_V \hat{\nabla}_\psi J_V(\psi)$ 
         $\theta_i \leftarrow \theta_i - \lambda_Q \hat{\nabla}_{\theta_i} J_Q(\theta_i)$  for  $i \in \{1, 2\}$ 
         $\phi \leftarrow \phi - \lambda_\pi \hat{\nabla}_\phi J_\pi(\phi)$ 
         $\bar{\psi} \leftarrow \tau\psi + (1 - \tau)\bar{\psi}$ 
    end for
end for

```



**Figure 9.** Policy Optimization Framework

Python with Mujoco and Open AI gym environments were used to train the robot. The return score was the sum of rewards as outlined in Figure 10.

```

def compute_reward(self, achieved_goal, goal, info):
    # Compute distance between goal and the achieved goal.
    d = goal_distance(achieved_goal, goal)
    if self.reward_type == 'sparse':
        return -(d > self.distance_threshold).astype(np.float32)
    else:
        return -d

```

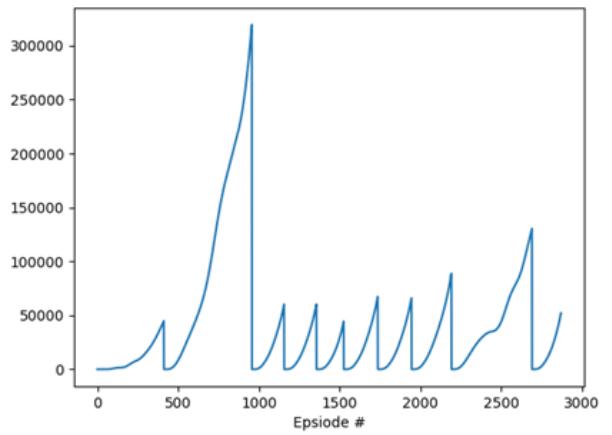
**Figure 10.** Reward function

The sparse reward system was used in this simulation environment as the default state where a ‘-1’ reward was applied for each step and a ‘0’ reward was applied for when the gripper reached the target polyp, represented by the box [18]. Through iterative reward cycles, the robot improved its training performance by optimizing its trajectory to achieve a return score of 0.

## V. Results & Discussion

### 1) Path planning and actuation of endoscope

While path planning was performed, there were 3 main limitations identified that prevented optimal performance.



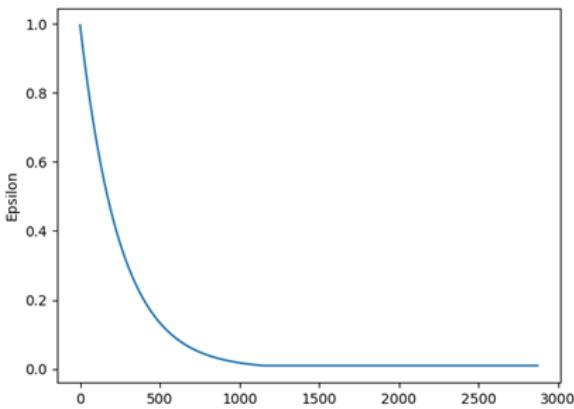
**Figure 11.** Reward Plot

The plot shown in Figure 11 depicts the observed iterative training process, demonstrating a lack of convergence. This can likely be attributed to three main factors.

Firstly, the experienced buffer size was small. As such, the sampling dataset that was used to train the agent did not have an adequate number of experiences to learn from. The model was training on the past 100 experiences, resulting in poor generalizability.

In addition, the number of training cycles was suboptimal. Specifically, the maximum number of iterations was 3000. In comparison, a typical DQN algorithm can use millions of iterations to run. In part, this was due to restraints in computational resources, given a lack of GPU.

Lastly, the Epsilon-Greedy parameter decayed too quickly. As a result, the network was taking actions from the Q network rather than from random exploration, which limited its capacity to improve training.



**Figure 12.** Epsilon Decay Plot

Given the current path planning algorithm performed suboptimally, we suggest the future use of additional computational resources to record more experiences and train more efficiently. In doing so, we would expect an increased number of training iterations and improved performance.

## 2) Localization and orientation of endoscope

Stereo images of colon were found from ‘A Template-based 3D Reconstruction of Colon Structures and Textures from Stereo Colonoscopic Images’ [14].

A total of 260 sequential images were used to generate a SLAM mapping using the Python framework developed by Taihu et al. (2017) [15, 16].

The intrinsic parameters of the SLAM algorithm were approximated through an iterative cycle. Throughout this process, the localization and mapping of the scope and soft tissue was noisy, impairing accuracy. To further enhance the SLAM algorithm, we propose undergoing proper intrinsic calibration through MATLAB to optimize the geometric representation between the stereo frames and improve performance.

**Table 7. SLAM demonstration**

Datasets	SLAM

Further, the training data was taken from a simulation resulting in further limitations in modelling ground truth data. As such, it was unclear how well the simulated algorithm represented the real colon environment.

In order to better characterize the human colon, early modelling studies could be performed in *in vivo* models, such as a pig, with stereovision colonoscopy to gather and tune the scope’s

intrinsic parameters. The anatomy of the pig large intestine would largely reflect the anatomy of the human colon, allowing for more accurate testing while ensuring patient safety.

### **3) Maintenance of sufficient insufflation pressures**

Assessing the sufficiency of insufflation pressures was proposed to occur through fuzzy logic, by characterizing the size of the colonic lumen. Small lumens would prompt an increase in colon pressure in order to distend the colon walls. Similarly, a large lumen would prompt decreasing the pressure of the colon to prevent overdistension and bowel wall injury. The fuzzy logic algorithm is reliant on Mask R-CNN to detect and localise lumen.

However, segmentations of the lumen were limited as a result of a paucity of available training data. As such, the Mask R-CNN was not able to be adequately trained with time constraints and limitations in available training data. However, the use of fuzzy logic appears to be a promising method of ensuring steady insufflation pressures.

### **4) Assessing adequacy of bowel preparation**

Similarly, assessing the adequacy of bowel preparation was proposed to occur through fuzzy logic, by characterizing the quantity of stool in the colon.

During conventional colonoscopy, a large amount of stool burden impairs the visualization of colonic polyps. The human operator can choose to wash away the stool to improve visualization or to quit the procedure and ask the patient to re-attempt bowel preparation at a future time.

The segmentation of stool was attempted; however, there was insufficient data for adequate training of the Mask R-CNN and thus the fuzzy logic algorithm.

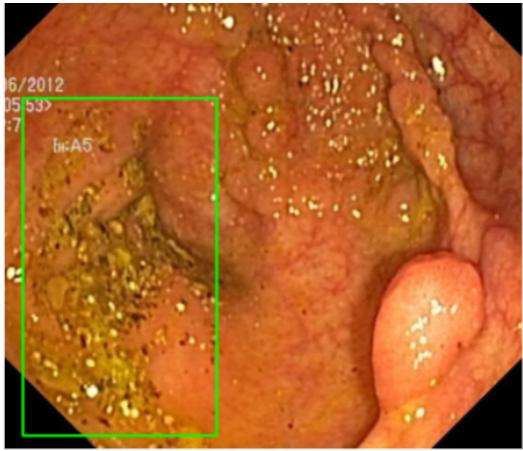
### **5) Tissue identification and localization**

The Kvasir-SEG, KUMC, ETIS-Larib, CVC-Clinic, and Harvard Dataverse datasets were used to hand label (1) stool, (2) lumen, (3) polyps, (4) ileocecal valve, and (5) appendiceal orifice.

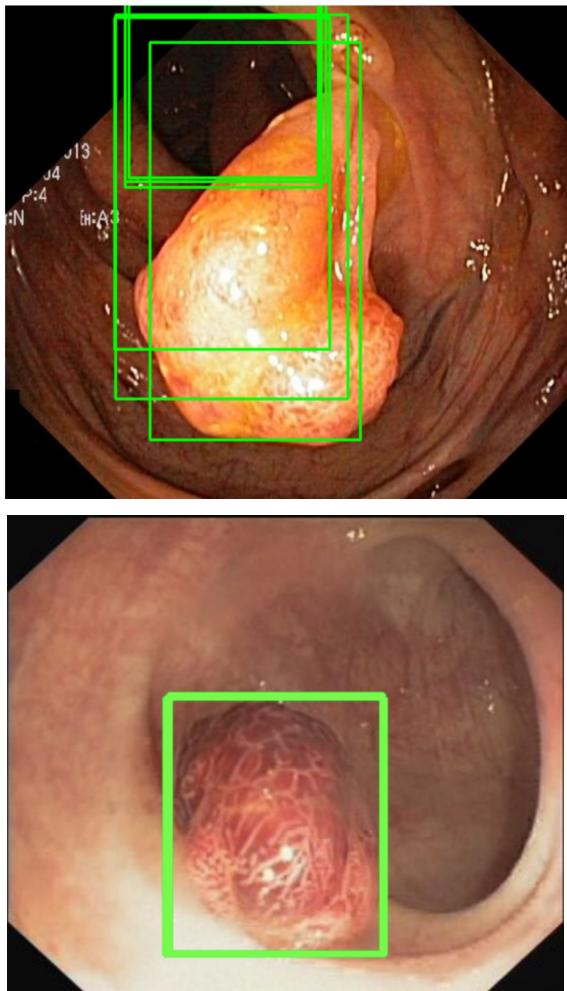
Tissue segmentation occurred using the following number of segmented images: 300 stool, 500 lumen, 2,400 polyp, 100 ileocecal valve, and 100 appendiceal orifice images. Feature identification of stool, lumen, and polyps were trained on 3 separate networks. The ileocecal valve and appendiceal orifice segmentation were removed given an absence of training data.

Descriptive statistics were calculated to characterize the performance of polyp identification and localization. Specifically, out of 198 polyp images, the trained model had an accuracy of 87.8% (174/198). The IoU accuracy, reflecting segmentation inaccuracy, was 0.371. Further, the mean Dice Coefficient score for the test set was 0.982, and the mean f1 score for the test set was 0.0734.

Testing performance of stool and lumen was limited due to a lack of segmented data. The capacity for the system to identify these features is shown below, with bounding boxes representing identification of stool, lumen, and polyps.



**Figure 13A.** Tissue identification with Mask R-CNN bounding boxes overlying stool.



**Figure 13 B, C.** Tissue identification with Mask R-CNN bounding boxes overlying the colonic lumen and a polyp

## 6) Characterization of polyps

The Harvard Dataverse ‘Colonoscopy Polyp Detection and Classification’ dataset, reported approximately 40 000 labelled polyps for training of the inception network. Of note, this data could not be used for tissue localization and identification given an absence of image segmentation. The data contained in this dataset included categorization of polyp images as “adenoma”, “hyperplastic”, or “unspecified”[20].

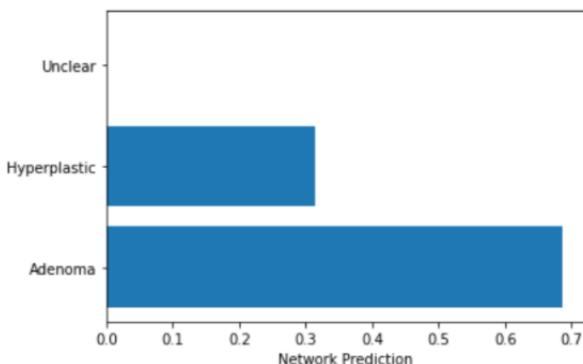
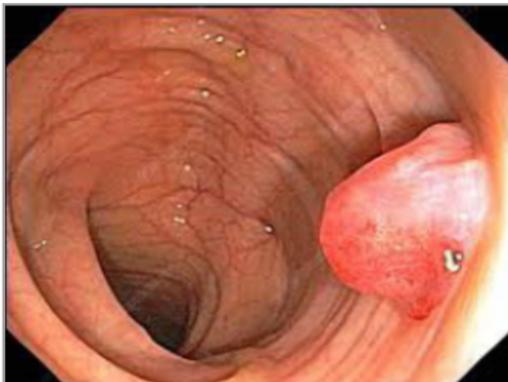
An example of the network performance is described below for adenomatous polyp, hyperplastic polyp, and unspecified respectively.

When tested with an image of an adenomatous polyp, the model accurately identified it as an adenomatous polyp with a network prediction of 70% for adenoma and 30% for hyperplastic.

When tested with an image of a hyperplastic polyp, the model misidentified it as an adenomatous polyp with a network prediction of 52% for adenoma and 48% for hyperplastic.

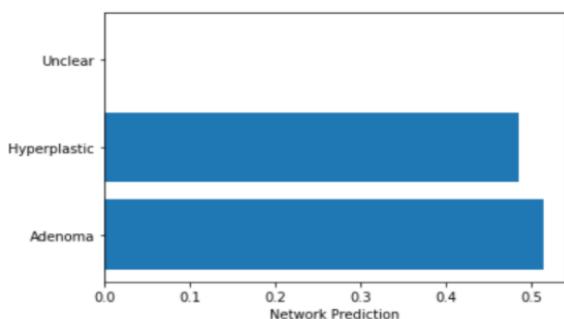
As demonstrated by the confusion matrix below, the overall model had an accuracy of 87%, 66%, and 46% when identifying adenomatous polyps, hyperplastic polyps, and unspecified images respectively (see Figure 14).

**Adenomatous Polyp**

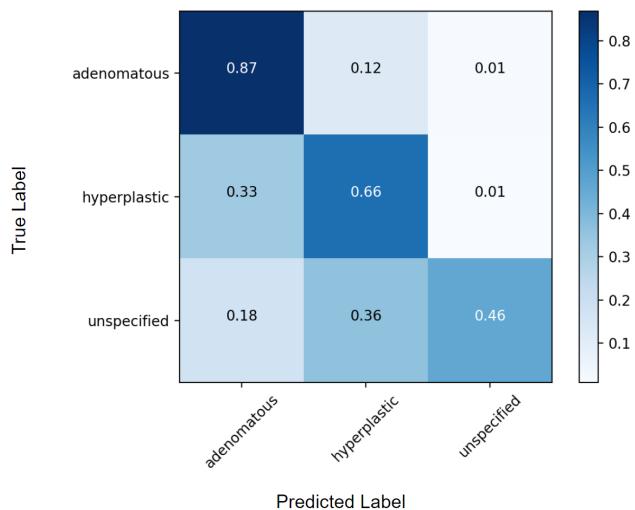


**Figure 14.** Network prediction for adenomatous polyp detection

**Hyperplastic Polyp**



**Figure 15.** Network prediction for hyperplastic polyp detection



**Figure 16.** Confusion matrix

Overall, the confusion matrix for polyp characterization demonstrates poor accuracy in identifying the third category of “unspecified” images. This category was added to allow for uncertainty when training with blurry and/or uncategorized images. This reflected the labelling structure of the Harvard dataset. However, in order to improve the accuracy of the model, there may be utility in removing this third category to improve accuracy. The inclusion of only two categories reflects the two options of adenoma and hyperplastic polyps. The capacity for the model to label images as “other/unspecified” increases the inherent uncertainty of the model. By removing this third category, it is expected that the network would be able to better categorize the two broad categories of polyps and improve accuracy.

## 7) Polyp manipulation and biopsy

The use of reinforcement learning appropriately allowed for the training of polyp manipulation, as represented by a box. Specifically, the simulated robot arm was consistently able to touch the box; however, it failed to appropriately pick up the box and transfer it to the set target. The system iteratively improved its error of polyp

manipulation based on the provided reward function, as outlined in the Methods section. However, training of the RL model did not successfully converge due to inadequate training iterations.

This can be further enhanced through the use of a GPU with larger training batches over a higher number of epochs until the network converges to optimal performance.

## VI. Future Directions

Overall, the conceptualized AI-driven robot demonstrates the potential utility of autonomous colonoscopy in routine colorectal cancer surveillance. To the authors' knowledge, this is the first report that outlines a design for autonomous colonoscopy navigation.

Specifically, it showed moderate-to-high accuracy in detecting adenomatous polyps, those with the highest malignancy potential. As such, with further refinement, the capacity of autonomous colonoscopy to improve both access to routine surveillance as well as increase the frequency of colorectal polyp detection in comparison to conventional methods is highlighted. These have large implications in reducing the significant financial burden that colorectal cancer places on our healthcare system [1,2].

In particular, every individual aged 50 requires a routine colonoscopy. With a rise in younger populations developing colorectal cancer, the need for colonoscopy may further expand to include earlier screening. Even with unremarkable findings, colonoscopies need to be repeated routinely, often every 10 years to identify potentially new lesions. In those with suspicious features, patients may require even more frequent follow up assessments, such as every 1 to 5 years [1,2].

Within the current healthcare system infrastructure, conventional colonoscopy is resource and labour intensive. They often require an endoscopy suite or operating room with multiple trained health care professionals. The associated costs and need for skilled operators can be unfeasible for many low resource communities [1,2]. As such, we envision the potential to improve access to care with the adoption of autonomous colonoscopy.

While the robot was not physically constructed, the use of a growing vine robot design through pressure-driven eversion is novel and appears to be a promising method of reducing bowel injury. With pressure-drive eversion, the robot lengthens from the centre, and the robot places even pressure on the walls of the bowel. This is important as an additional safety measure to reduce the risk of bowel perforation.

Although the proposed framework appears promising, the authors acknowledge potential limitations with the described AI system. Specifically, the performance of Mask R-CNN can be enhanced with additional training data and higher computational infrastructure. This may occur through the prospective collection of patient colonoscopy images from hospitals and clinics, and the use of GPU for processing.

Furthermore, we recognize the limitations of SLAM. The SLAM algorithm often performs optimally under static environments. However, this represents an idealized environment and does not accurately reflect the dynamic nature of the human colon. Rather, the colon is deformable and has variable motion, which can impair the SLAM estimation processes involved with mapping soft tissues.

In addition, it is expected that both path planning and polyp manipulation could be improved

through RL transfer learning, such as by applying the developed 3D simulation to an *in vivo* model. This would allow for the incorporation of additional complexities within the model that better represent the characteristics of real human tissues. The dynamic parameters involved with fluid dynamics or the nature of soft tissues is complex. The deformability of soft tissues—which can vary based on patient factors, such as age and sex—would likely best benefit from transfer learning.

Despite these limitations, the development of an autonomous colonoscopy through AI appears promising and, with further refinement, its capacity to transform healthcare appears monumental. The design process outlined for autonomous colonoscopy also may extend to other forms of health care investigations, such as upper endoscopy to identify stomach polyps or bleeding in the upper gastrointestinal tract. These applications of AI-driven robotics have so far gone unrealized and remain a promising area of study and robotics development.

## VII. Conclusion

In conclusion, artificial intelligence has promising applications in the automation of colonoscopy to support navigation in the dynamic environment of the colon. In doing so, the development of an autonomous colonoscopy robot would have important applications in increasing accessibility to screening colorectal cancer. Further training data is required to improve the overall system performance. Lastly, integration of each AI process is needed to better characterize its safety and efficacy in a model colon environment.

## Abbreviations

AI, artificial intelligence  
CAD, computer-aided diagnosis

CNN, convoluted neural network
R-CNN, regional convolutional neural network
FPS, frame per second
FAST, features from accelerated segment test
LR, logistic regression
REP, robotic endoscope platform
RFML, random forest machine learning
SIFT, scale-invariant feature transform

## Terminology

- **Bowel preparation** - A regimen where patients drink laxatives to clear their colon of stool prior to undergoing a colonoscopy
- **Colon** - The terms colon and bowel are used interchangeably to refer to the large intestine
- **Colonoscopy** - The use of a camera to visualize the colon; a type of endoscopy
- **Endoscopy** - The use of a camera to visualize the digestive tract; colonoscopy is one type of endoscopy
- **Polyps** - Colorectal polyps are precancerous growths in the colon
- **Lumen** - Center of colon which the colon wall encapsulates
- **Cecum** - Lower abdominal cavity that receives material from the small intestine.
- **Appendiceal orifice** - Opening of the appendix; an anatomic landmark of the cecum (beginning of the colon)
- **Ileocecal valve** - Transition between the end of the small intestine and the beginning of the colon; an anatomic landmark of the cecum

## References

- [1] Force UP, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson K, Epling JW, Garcia FA, Gillman MW, Harper DM, Kemper AR. Screening for colorectal cancer. *Jama*. 2016;315:2564-75.
- [2] Brenner DR, Ruan Y, Shaw E, De P, Heitman SJ, Hilsden RJ. Increasing colorectal cancer incidence trends among younger adults in Canada. *Preventive medicine*. 2017 Dec 1;105:345-9.
- [3] Li JW, Chia T, Fock KM, Chong KW, Wong YJ, Ang TL. Artificial intelligence and polyp detection in colonoscopy: Use of a single neural network to achieve rapid polyp localization for clinical use. *J Gastroenterol Hepatol*. 2021 Jul 29. doi: 10.1111/jgh.15642. Epub ahead of print. PMID: 34327729.
- [4] Dan M. Livovsky, Danny Veikherman, Tomer Golany, Amit Aides, Valentin Dashinsky, Nadav Rabani, David Ben Shimol, Yochai Blau, Liran Katzir, Ilan Shimshoni, Yun Liu, Ori Segol, Eran Goldin, Greg Corrado, Jesse Lachter, Yossi Matias, Ehud Rivlin, Daniel Freedman:Detection of elusive polyps using a large-scale artificial intelligence system.June 29, 2021. <https://doi.org/10.1016/j.gie.2021.06.021>
- [5] Pfeifer, Lukasa; Neufert, Clemensa; Leppkes, Moritz; Waldner, Maximilian J.a; Häfner, Michaelb; Beyer, Albertc; Hoffman, Arthurd; Siersema, Peter D.e; Neurath, Markus F.a; Rath, Timoa Computer-aided detection of colorectal polyps using a newly generated deep convolutional neural network, *European Journal of Gastroenterology & Hepatology*: May 21, 2021 - Volume - Issue - doi: 10.1097/MEG.0000000000002209
- [6] Meng, Jie, Xue, Linyan, Chang, Ying, Zhang, Jianguang, Chang, Shilong, Liu, Kun, Liu, Shuang, Wang, Bangmao and Yang, Kun. "Automatic detection and segmentation of adenomatous colorectal polyps during colonoscopy using Mask R-CNN" *Open Life Sciences*, vol. 15, no. 1, 2020, pp. 588-596. <https://doi.org/10.1515/biol-2020-0055>
- [7] Martin, J.W., Scaglioni, B., Norton, J.C. et al. Enabling the future of colonoscopy with intelligent and autonomous magnetic manipulation. *Nat Mach Intell* 2, 595–606 (2020). <https://doi.org/10.1038/s42256-020-00231-9>
- [8] M. M. Coad et al., "Vine Robots: Design, Teleoperation, and Deployment for Navigation and Exploration," in *IEEE Robotics & Automation Magazine*, vol. 27, no. 3, pp. 120-132, Sept. 2020, doi: 10.1109/MRA.2019.2947538.
- [9] Zhang Q, Prendergast JM, Formosa GA, Fulton MJ, Rentschler ME. Enabling Autonomous Colonoscopy Intervention Using a Robotic Endoscope Platform. *IEEE Trans Biomed Eng*. 2021 Jun;68(6):1957-1968. doi: 10.1109/TBME.2020.3043388. Epub 2021 May 21. PMID: 33296299.
- [10] A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052-1067, June 2007, doi: 10.1109/TPAMI.2007.1049.
- [11] 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS 2015) Sensor Technologies and Simultaneous Localization and Mapping (SLAM) T.J. Chong, X.J. Tang, C.H. Leng, M. Yogeswaran\*, O.E. Ng, Y.Z. Chong Universiti Tunku Abdul Rahman, UTAR Complex, Jalan Genting Kelang, 53300 Setapak, Kuala Lumpur, Malaysia
- [12] SLAM (Simultaneous Localization and Mapping) <https://www.mathworks.com/discovery/slam.html>

- [13] Accurate and Informative Data: The LiDAR Advantage  
<https://redzone.com/nr/the-lidar-advantage/>
- [14] S. Zhang, L. Zhao, S. Huang, M. Ye and Q. Hao, "A Template-Based 3D Reconstruction of Colon Structures and Textures From Stereo Colonoscopic Images," in *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 1, pp. 85-95, Feb. 2021, doi: 10.1109/TMRB.2020.3044108.
- [15] Taihú Pire, Thomas Fischer, Gastón Castro, Pablo De Cristóforis, Javier Civera and Julio Jacobo Berlles. S-PTAM: Stereo Parallel Tracking and Mapping Robotics and Autonomous Systems, 2017.
- [16] Taihú Pire, Thomas Fischer, Javier Civera, Pablo De Cristóforis and Julio Jacobo Berlles. Stereo Parallel Tracking and Mapping for Robot Localization Proc. of The International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015.
- [17] Salcedo Bosch, Martí, "Data Integration Strategies for Distributed Reinforcement Learning in Robotics"
- [18] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, Pieter Abbeel, Wojciech Zaremba, "Hindsight Experience Replay"
- [19] <https://www.analyticsvidhya.com/blog/2019/04/introduction-deep-q-learning-python/>
- [20] Wang, Guanghui, 2021, "Replication Data for: Colonoscopy Polyp Detection and Classification: Dataset Creation and Comparative Evaluations", <https://doi.org/10.7910/DVN/FCBUOR>, Harvard Dataverse, V1
- [21] uoip, 2018, Stereo Implementation PTAM SLAM, [https://github.com/uoip/stereo\\_ptam](https://github.com/uoip/stereo_ptam)
- [22] R. Kuemmerle, G. Grisetti, H. Strasdat, K. Konolige, W. Burgard, 2011, g2o - General Graph Optimization,  
<https://github.com/RainerKuemmerle/g2o>
- [23] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [23] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Erratum: Corrigendum: Dermatologist-level classification of skin cancer with Deep Neural Networks," *Nature*, vol. 546, no. 7660, pp. 686–686, 2017.
- [24] M. Bosch, "Data Integration Strategies for Distributed Reinforcement Learning in Robotics" MAsc Dissertation, TUM School of Engineering and Design. Dec 12 2019. Accessed on Nov 10 2021. Available:  
<https://upcommons.upc.edu/bitstream/handle/2117/185242/thesis.pdf?sequence=1&isAllowed=y>