A polyp is an abnormal tissue found in the large intestine that is usually noncancerous but can develop into cancer. Colonoscopy allows endoscopists to identify polyps and determine if they require removal. However, this whole procedure is highly dependent on human skill and thus poses challenges to its effectiveness. This project aims to utilize computer vision to detect polyps in single video frames from colonoscopies. Specifically, it consists of polyp detection and image segmentation. Detection is defined as finding a bounding box for each polyp in an image. At the same time, segmentation means assigning pixels into two categories (Polyp or normal).

Our deep-learning method will be built on two datasets including **PolypGen** and **Real-Colon**. **PolypGen** contains still images and video sequences from colonoscopies, including both positive and negative frames, from 6 medical hospitals in Europe and Africa. It includes 3,762 polyp positive video frames and 4,275 polyp negative video frames, with positive images being annotated with both bounding boxes and pixel-level annotations. Further information can be found on [Synapse](https://www.synapse.org/Synapse:syn26376615/wiki/613312). **REAL-Colon** contains 2,757,723 total video frames from colonoscopies from 4 medical hospitals in North America, Europe, and Asia. Positive frames are annotated with bounding boxes, but not pixel-level annotations. For details, please refer to [Figshare](https://plus.figshare.com/articles/media/REAL-colon_dataset/22202866).

Project stakeholders for this project include:

* **Endoscopists and Pathologists:** They are primary users of the system. Endoscopists care about accuracy, speed, and clinical workflow integration. Pathologists could use this tool for validating findings in imaging and support diagnoses.
* **Hospital and Healthcare:** Hospital management can use this tool to fit in with their desired budget while not interfering with accuracy. This tool could reduce procedure times, lower costs and improve patient outcomes.
* **Medtech Developers:** This tool also serves as a bridge between research and industry for pharma or medtech companies. It would provide commercialization and integration with existing endoscopy hardware and practices.

KPIs: Polyp detection will be our primary goal and segmentation is our secondary goal. We aim for measurable yet realistic improvements over a select baseline model (e.g. YOLOv3 for detection) trained on PolypGen. Success will be defined by achieving a 10% absolute precision gain at higher recall levels (e.g. ≥75 %) compared to the baseline model. Secondary KPIs include (a) stable training and inference within the planned compute budget and (b) reproducible end-to-end training and evaluation pipelines.

For detection, precision is calculated based on an Intersection over Union (IoU) threshold. IoU measures how well a predicted bounding box overlaps with a ground-truth box. A true positive (TP) means that IoU with the best-matching ground-truth box is greater than a chosen threshold.

For segmentation, TP is a pixel accurately predicted as belonging to a polyp. There are many existing metrics for segmentation including precision, recall, overall accuracy (Acc), Hausdorff distance (Hd), etc. We will first focus on precision and recall. If time permits, we will try advanced metrics like Hd, Jaccard Coefficient (JC), or F2-score.