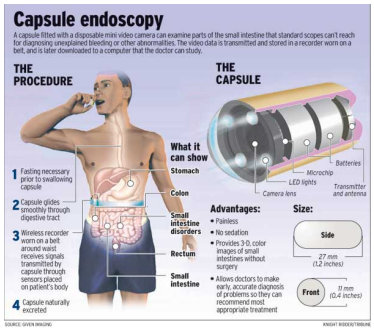
Exam Project

Jürgen Herp[*], Victoria Blanes-Vidal[†], Esmaeil S. Nadimi[‡]

December 4, 2018



Motivation

The aim of this study is to develop a machine learning-based model to classify bowel cleansing quality into unacceptable and acceptable, and further segment bowel wall and polyps.

The human bowl system is a long continuous tube running from the stomach to the anus. The system is roughly comprised of the small intestine (small bowel) and the large intestine (colon or large bowel). The small intestine is about 6 meters long, and on of its prime functions is to absorb most of the nutrients from the semi-fluid mass of partly digested food. The large intestine is approximately 1.5 meters long and up to 9 centimeters in diameter. The colon removes water, salt, and some nutrients forming stool. The colon absorbs water from wastes, creating stool. As stool enters the rectum, nerves create the urge to defecate.

Capsule endoscopy, in contrast to traditional colonoscopy, is a procedure involving a pill shaped capsule, equipped with a camera, that wireless sends information to an receiver outside the human body. Among others, capsule endoscopy can help screening for polyps, and aid in diagnosing cancer.

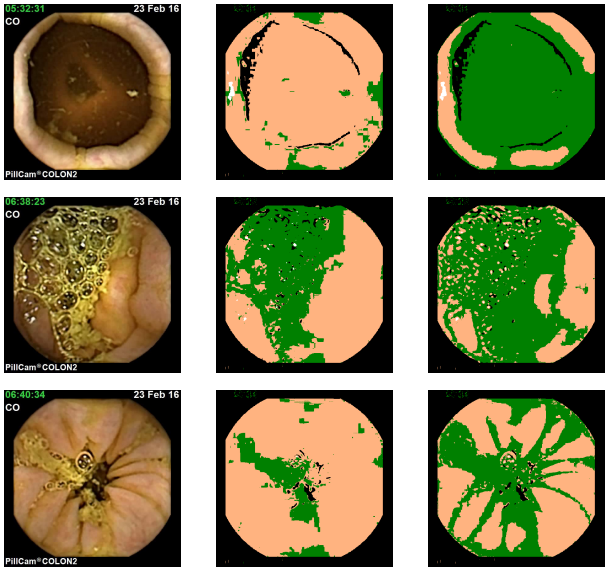
Bowel polyps are small growths on the inner lining of the bowel. Most colon polyps don't cause symptoms. Thus, most people with polyps will not know that they have one or more. They are often detected up during screening. Since capsule endoscopy does not have functions which allow suctioning of fluid or washing the bowel mucosa during the examination, its diagnostic yield can be limited by presence of stool, biliary secretion, bubbles and blood that may hide relevant finding (i.e. polyps).

It is thus of utter most importance to be able to distinguish between the following: (i) Is a captured image clean enough to extract meaningful features? (ii) Can we segment (classify) pixels of the image to belong to specific classes?

1 Bowl Assessment

(a) **Data Collection:** Download the image set from BlackBoard. Crop the image, if neccesarry, to get rid of the meta-date on the edges of the images, as seen in the column (a) of Figure 1. Divide your dataset into a Training-Validation and Test set.

[*]herp@mmmi.sdu.dk
[†]vbv@mmmi.sdu.dk
[‡]esi@mmmi.sdu.dk



(a) Raw images (b) Index Classifier (c) SVM Classifier

Figure 1: Pixel classification using a nonlinear index and a SVM classifier. *Buijs et al., Assessment of bowl cleansing quality in colon capsule endoscopy using machine learning. A pilot study, Endoscopy International Open, 2018*

(b) **Data Labeling:** Label a subset of pixels and/or images according to the classes:

(a) Clean
(b) Dirty
(c) Bowel Wall
(d) Polyp

Save your initial labels in a separated file. These will be your ground-truth.

(c) **Data Classification:** Train one or more algorithm of your choice to predict whether a pixel and an image belongs to either of the chosen classes.

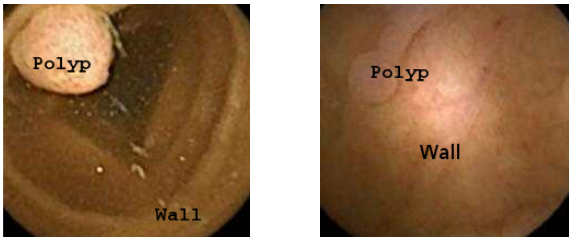


Figure 2: Example for wall and polyp.

(d) **Data Segmentation:** Segment the pixels of each image. For inspiration how that might look like see Figure 1 in case of an index and support vector machine (SVM) classifier.

(e) **Describe your findings:** Include the following questions when elaborating on your finding:

- What are the model(s) errors?
- Which model offers the best prediction?
- What are the relevant observed variables to consider, if any?

2 Project Requirements

- The project has to be submitted in PDF format and should not exceed 10 pages excluding appendix.
- The source code has to be submitted separately with sufficient documentation.
- The ground-truth labels have to be submitted in a separate file with sufficient documentation.