



UPPSALA
UNIVERSITET

UPTEC X 16015

Examensarbete 30 hp
Augusti 2016

Stool Detection and Classification in Colorectal Cancer

Sabri Jamal



UPPSALA
UNIVERSITET

Degree Project in Bioinformatics

Masters Programme in Molecular Biotechnology Engineering,
Uppsala University School of Engineering

UPTEC X 16 015	Date of issue 2016-08
Author Sabri Jamal	
Title (English) Stool Detection and Classification in Colorectal Cancer	
Title (Swedish)	
Abstract <p>This project has been dedicated to the field of medical image analysis concerning the issue of colorectal cancer. Cancers can be evolved in almost any part of the body and is therefore a disease that impacts the whole world. Colorectal cancer is just one of such cancer types and has been coined as one of the more frequent cancer types encountered. Colonoscopy is the accepted screening method for identifying elements known as polyps. Polyps are perceived as swollen tissue found in the colon. Before the search for the polyps begins, an assessment of how clean the bowel is first made to ensure it is safe to identify the above-mentioned elements. This thesis has been focused on detection and classification in order to calculate the percentage of each stool type present in the colon. To address this, k-means clustering was implemented using features such as texture and color to classify the different stool types. Firstly the images were preprocessed, the preprocessing was followed by color segmentation and finally the images were classified. Once the classification of each pixel had been done the classified pixels were assigned a class label. Each label was connected to color and finally a visual representation of the classified image was presented through repainting the entire image. The results show that in a perfect segmentation of the colon the classifier performs well. While in the case of a partial segmentation the frequency of misclassifications increases.</p>	
Keywords <p>BBPS (Boston Bowel Preparation Score), polyp, visual descriptor, feature, LBP (Local Binary Pattern), dilation/in-paint, segmentation, illumination, illumination invariant, PCA (Principal Component Analysis)</p>	
Supervisors Maria Begoña Garcia Zapirain Soto Universidad de Deusto/University of Deusto	

Scientific reviewer		
Carolina Wählby Uppsala University		
Project name	Sponsors	
Language English	Security	
ISSN 1401-2138	Classification	
Supplementary bibliographical information	Pages 65	
Biology Education Centre Box 592, S-751 24 Uppsala	Biomedical Center Tel +46 (0)18 4710000	Husargatan 3, Uppsala Fax +46 (0)18 471 4687

Stool detection and classification in Colorectal Cancer

Sabri Jamal

Populärvetenskaplig text

Cancer är en grupp av sjukdomar som delar samma karakteristiska drag, dvs. onormal celledelning. Eftersom sjukdomen grundar sig just i cellers onaturliga celledelning innebär detta att det blivit en av de mest omfattande vardagliga sjukdomar vi hör om dagligen. Tjocktarmscancer är en form av cancer som orsakas i tjocktarmen och är extremt vanligt i dagsläget. Idag är koloskopi en vanligt förekommande metod för att undersöka för polyper eller så kallade tumörer, och i många länder är det ansett som ett rutin test som rekommenderas varje X antal år beroende på resultat från tidigare tester. Innan man kan leta efter polyper måste varje patient genomgå en rensning av avföring genom intag av laxeringsmedel. Det behöver sedan bestämmas om tjocktarmen är ren nog för att koloskopi bilderna/videon kan användas för att leta och hitta polyperna med god säkerhet

Denna studie baserar sig på att utveckla en mjukvara som kan detektera och klassificera de olika avföringstyperna i tjocktarmen för att tillslut beräkna hur stor procentandel avföring existerar i bilden. Tanken är att man i framtiden ska kunna bygga på detta projekt och slutligen kunna detektera, klassificera samt evaluera renheten av tjocktarmen.

Table of contents

1. Introduction	1
2. State of the art	2
2.1 Technical background	2
3. Justification	7
4. Objectives	8
5. Methods	9
6. Design	14
6.1 System Design	14
7. Development	27
7.1 Database	27
7.2 First stage: Elimination of black borders	27
7.3 Second Stage: Dilation of black edges	28
7.4 Third stage: Dilation of specular reflections	28
7.5 Fourth Stage: Intensity Adjustment	29
7.6 Fifth Stage: Color Segmentation	30
7.7 Sixth Stage: Training Data	31
7.8 Seventh Stage: Feature extraction & PCA	32
7.9 Eighth stage: K-means classification	33
8. Results	35
8.1 Comparison of doctors evaluation	35
8.2 Classifier comparison with doctors	37
8.3 Classifier compared with ground truth classification	39
8.4 Additional classified images	43

9. Discussion.....	45
9.1 Issues classifying the colon.....	45
9.2 Subjectivity in judgement.....	47
9.3 Images taken from different angles.....	47
9.4 The received Data	49
9.5 Dilation & Gaussian filtering of images.....	49
9.6 Evaluation of classifier.....	49
9.7 Alternative Feature selection methods	50
10. Conclusion	51
11. Future work	52
12. Acknowledgements.....	53
13. Bibliography	54

1. INTRODUCTION

The colon and rectum are a part of our intestines and can be explained as long hollow tubes traversing from the stomach to the anal opening. Our gastrointestinal tracts comprises of two different types of intestines, the small and the large intestine [1]. The small intestine attaches to the colon while the colon continues to attach to the rectum and finally the anus [2].

The colon is a part of the final stage in the digestive system [1]. Its objective is to absorb fluids and nutrients, ultimately storing the remaining waste products until it exits the body [3].

Colon or rectum cancer (also called colorectal cancer) is a cancer type that originates in either the colon or the rectum. Colorectal cancer is generally an adenocarcinoma meaning that it is formed in epithelial producing fluids and mucus [4]. In 2012, colorectal cancer was the third most common form of cancer in the world with 1.4 million diagnosed cases [5].

Colonoscopy has been approved as the accepted screening method and it has become a standard procedure to examine for polyps [6] [8].

Polyps are growth of tissue that can be found in the colon walls. They exist in different sizes where larger polyps indicate higher risk and they are typically described to have the form of a mushroom without a stalk [7]. These elements become more common to encounter in individuals older than 50 years, although not all polyps are dangerous [9].

The procedure for polyp examination begins with the patient goes through a bowel cleansing treatment [9]. After the treatment has been made the colonoscopy can begin. In order to continue to the next step, which would be searching for polyps, a decision has to be made if the bowel is clean enough to rely on the video footage. If this is not the case the patient will have to redo the bowel cleansing procedure and repeat the colonoscopy [10]. Measuring how clean the bowel is done with the help of the Boston Bowel Preparation Scale (BBPS) to qualitatively measure the cleanliness of the bowel [11]. This is done in order to decide if the video footage is reliable enough to use for detection of polyps. However a problem arises with the current BBPS method due to the subjectivity in the scoring system. Due to this, one can expect different scores for the same images, both between two different doctors and also by the same doctor if the two evaluations have been made at separate times.

This master thesis focuses on detection and classification of different stool types in the colon. Thus in a near future allowing for automatic calculation of the BBPS score. By automating the procedure, the idea is to reduce the subjectivity in the doctor's evaluations, henceforth reducing the margin error or difference between the assessments.

2. STATE OF THE ART

The human body is made of trillions of cells which enables cancers to be formed almost anywhere in the body e.g. in the colon, eye and brain. Due to the immense possibilities of where tumors can be produced in combination with the difficulties of targeting the damaged cells in, vast efforts are focused on early detection [12]. The following section will lift the theoretical information essential for the understanding of the project

2.1 TECHNICAL BACKGROUND

2.1.1 Colon

The colon or large intestine is divided up into four different parts as seen in Fig 1, the ascending colon, transverse colon, descending colon and finally the sigmoid colon [13]. The small intestine however, is comprised of duodenum, jejunum and the ileum [14]. Finally the cecum is the pouch located at the end of the ascending colon, however it is not included in the colon itself [13].

Most of the digestion that is performed is done in the small intestines. It is in this part of the digestive system where a large part of the nutrients are absorbed and finally enter the bloodstream [15]. The colon absorbs nutrients as well, however its primary functions are to help with waste storage, provide digestion through bacterial fermentation and finally reclaim and maintain the water balance [3].

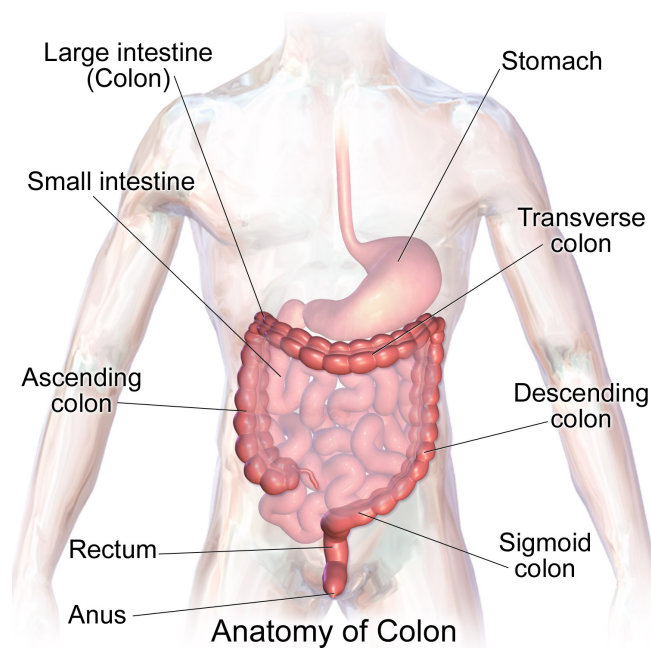


Fig. 1 Showing the large as well as the small intestine [16]. Illustration used with permission from Bruce Blaus.

2.1.2 Cancer in general

Cancer is a collective name used to describe a group of disease that shares similar characteristics. In all types of cancer, the process is initiated through the damages inflicted e.g. through sunlight [18] or other components that could directly damage the DNA in the cell. Once such damage has occurred to mechanisms related to the cell division [17] [19] and our repair mechanism fails to repair these faults. An uncontrollable proliferation process is started and this increase of cells in a concentrated region is what we know as a tumor [20]. Even though many cancers might produce tumors in the form of solid masses of tissues, cancers of blood such as leukemia's do not necessarily form such [4].

Tumors can either be benign or malignant. A benign tumor is cancer tissue that cannot spread to other parts of the body nor invade nearby tissue and generally does not grow back after it has been taken away [21]. However, malignant tumors unlike benign can spread to nearby tissue and also to other nearby organs [22]. The spreading of tumors to nearby organs is referred to a process called metastasis, which can be seen in Fig 2 below [23].

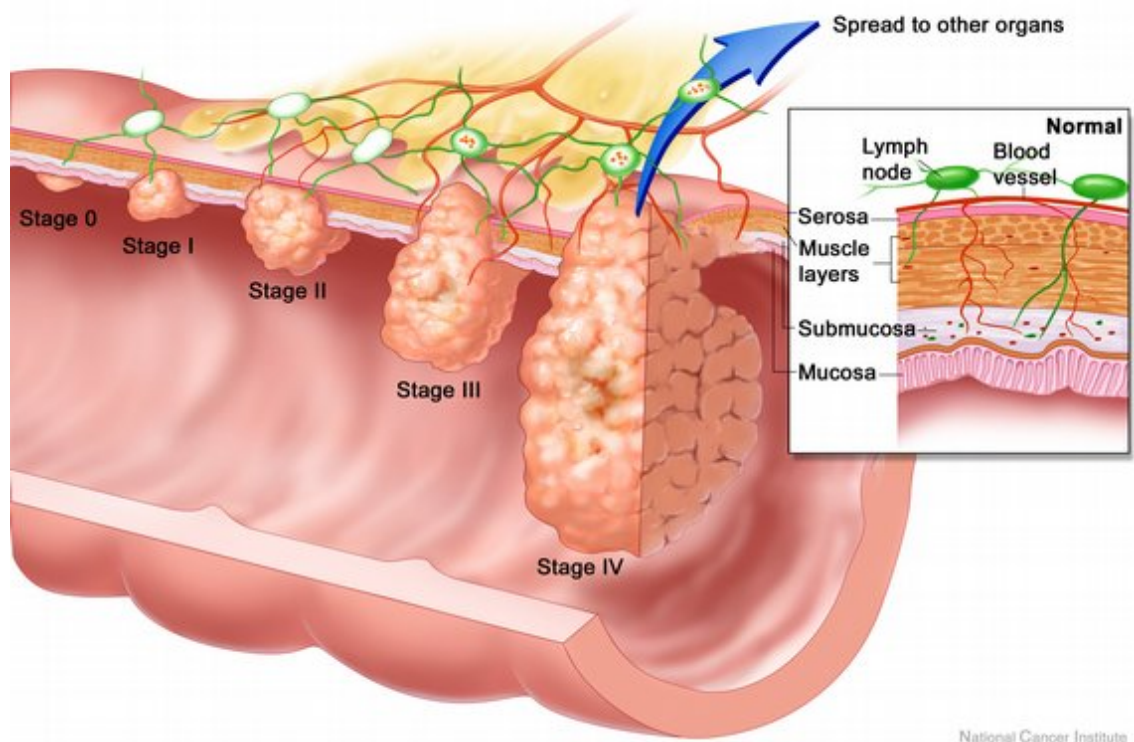


Fig. 2 Showing how colorectal tumor can spread if there is continued growth [24]. Illustration used with permission from Terese Winslow.

2.1.3 Difference between cancer cells and normal cells:

While a normal cell matures and attains its primary functions, it still lives under certain conditions. These conditions restrict cells from existing without a purpose and makes sure that damaged cells do not divide [25]. One of these very important processes is apoptosis also known as programmed cell death, which terminates the life cycle of a cell if any abnormalities are encountered or if the cell is just simply too old [26]. However cancer cells override these conditions and enter a state of uncontrolled cell division [28]. This results in new cells being formed but using the damaged cell as a template and ultimately allows for more damages and more uncontrollable divisions [4].

Cancer cells may in some cases influence normal cells that are in close proximity to the rogue cells [27]. By taking advantage of this, the cancer cells can help nourish the tumor, e.g. by forcing normal cells to aid in the formation of blood vessels to pass oxygen, nutrients and even help discard of waste products [4]. Besides affecting nearby entities, cancer cells have also the ability to occasionally hide from the immune system [29]. However to battle this, our bodies rely on the immune system whose primary objective is to defend the body from foreign attacks coming from pathogens [30], such as bacterias. It generally becomes more difficult with cancer cells because the disease generates from within, due to cells that have gone out of control. This is a problem just as big for the immune system as it is for doctors treating these diseases. The difficulties of determining and targeting sick cells while at the same time restraining from harming healthy cells is one of many controversies when discussing treatment, and remains a problem for our defense system just as it does for our doctors [4].

2.1.4 Colon cancer

Colon cancer often begins with elements called polyps that can be found in the colon. Polyps are growths of tissue that can exist in different sizes that commonly take the form of a mushroom without a stalk [7].

Adenoma polyps are however polyps that can have a higher risk of becoming cancerous [9]. The accepted screening method for polyps is colonoscopy. As the most efficient way to battle cancer is considered to be early detection, the need for colonoscopy is growing. However it is worth to mention that there are other alternatives to screen for colorectal cancer e.g. sigmoidoscopy and fecal occult blood test (FOBT) [31].

2.1.5 Bowel cleansing & colonoscopy

Before the colonoscopy can be performed, the patient needs to go through a bowel cleaning treatment [8]. This procedure is based on the patient taking a certain laxative e.g. polyethylene glycol (PEG), sodium phosphate, magnesium citrate or bisacodyl [32]. The objective of the treatment is to clean the bowel before the colonoscopy with respect to certain conditions. These conditions are that the procedure should have little effect on the gross and microscopic appearance of the colon while getting rid of as much fecal material as possible [8]. It should cause as little patient discomfort as possible and finally increase the detection of polyps in the

colon [33]. Despite the efforts a significant amount of inadequate cleansing still occurs, ranging through 10% to 75% shown in randomized controlled trials [34]. Although it has been proved that insufficient cleansings has had connections to certain patient characteristics, e.g. the use of antidepressants, history of constipation and patient disobedience regarding the necessary restrictions while on the treatment [34].

2.1.6 The Boston Bowel Preparation Scale (BBPS)

As colonoscopy has become the standard screening method for detection of cancerous elements in the colon, it is apparent that it has its obstacles. One of such is the fact of missing polyps and or lesions during colonoscopies [35]. Something that is directly related to how polyps can be missed while performing the procedure is the quality of the bowel cleansing performed before the procedure [33]. Thus, to improve the accuracy of the practice it was decided that a standardized manner of evaluating how well the bowel was washed through [11]. Henceforth a new system was inaugurated. The coined terms, “excellent”, “good”, “fair” and “poor” were first introduced. However, as the grading was considered subjective, the terms were exchanged with values ranging [0 3] applied to the three major parts of the colon, the right side (cecum, ascending colon), left side (descending colon, sigmoid colon and rectum) and the transverse colon (located in the middle) [11].

The following bullet points below are the exact regulations designed for the BBPS evaluation described and cited from the original produced article [11].

- “0, unprepared colon segment with mucosa not seen because of solid stool that cannot be cleared.”
- “1, portion of mucosa of the colon segment seen, but other areas of the colon segment are not well seen because of staining, residual stool, and/or opaque liquid.”
- “2, minor amount of residual staining, small fragments of stool, and/or opaque liquid, but mucosa of colon segment is seen well.”
- “3, entire mucosa of colon segment seen well, with no residual staining, small fragments of stool, or opaque liquid.”

Once the separate parts of the colon have received a partial score, they are summed up yielding a new complete BBPS score between [0 9] describing how clean the colon is [11].

2.1.7 Previous research

The research that has been previously done in the area is very limited. Most research that has been performed within the field has been within polyp detection.

An article was found titled, Color Based Stool Region Detection in Colonoscopy Videos for Quality Measurements [33]. Here they focus on detection of stools without distinguishing between the different stool types. The article describes a stool detection method based on observing planes made when plotting the cubic RGB space. The planes were made along the axis representing the red color channel where the range of values is [0 255]. This amounts for 256 planes that could possibly contain a stool pixel. If a stool pixel was encountered in a plane, the plane was selected. Each plane is then treated as a classifier to classify stool pixels. Finally, once all the stool has been detected a Boston Bowel Preparation score is calculated solely based on the percentage of stool existing in the image.

3. JUSTIFICATION

This master thesis has been centered on the field of medical image analysis and what it can provide for issues concerned in colonoscopy. Colorectal cancer is cancer originated in the colon or the rectum and was in 2012 considered to be the third most common form of cancer in the world [5]. Colonoscopy was decided to be the approved screening method and thus it turned into a standard procedure for examination of polyps. This project discusses the many subjective reasoning's behind the different practices and aims to deliver a program that would be a beginning to objectify the discussed practices.

The different practices include:

- Identification of what stool type has been located (solid, liquid or stain)
- Evaluation of the amount of stool located in the colon
- BBPS score assigned to each image and later summed up giving an entire score for the three different larger parts of the colon (right, left and transverse colon).

It is worth to mention that each of the stool types come with a certain problem for the doctor performing the surgery. Due to this, the stools are not equal and cannot be classed as one entity as this is a weighted problem. In this thesis, it will be seen how there at times can be a divergence in the three practices mentioned above. The intention of this project is therefore to work towards a goal that would reduce the subjectivity and hopefully aid in unifying future colonoscopy evaluations.

4. OBJECTIVES

The main objective of this research has been to design, develop and deploy a software capable of analyzing colonoscopy images to detect and classify stool in the colon. The detection and classification uses a k-means clustering algorithm with the help of features such as texture and color. After detection and classification the percentages of each stool type are calculated. This could therefore be a beginning to standardizing computation of the BBPS score and thus introduce a way to quantitatively measure evaluations.

The main objective is therefore to:

- Preprocess colonoscopy images
- Perform initial color segmentation
- Classify & detect stool types present
- Compute percentages of each stool type classified

The scope of the project is to create a software able to take an image as input and finally classify encountered stools. The result is presented as a repainted version of the original image where each class of the 6 classes (solid, liquid, clear liquid, dark liquid, stain, colon) receives a label that will be assigned a certain color. A plot is then performed in order to re-paint the image thus giving a visual presentation to the user.

5. METHODS

5.1.1 *Imdilate*

Imdilate is a built in matlab function used twice in the preprocessing block of the project. The function utilizes a structure element B where \hat{B} is the reflection of the structure element. In the case of this project, the structure element used was a disk of radius r . The dilation was performed in order to replace black edges as well as replace the continuous problem of specular reflections in the colonoscopy images.

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (1)$$

Binary Dilation

5.1.2 Gaussian filtering

After performing *imdilate* a Gaussian filtering was applied using the function *fspecial*. The Gaussian filter is a lowpass filter that attenuates intensities above a certain threshold. It was used to reduce the manipulation occurring on the regions being dilated as well as minimizing the problem if small specular reflections existed around the pixels being used for the dilation.

$$h_g(n_1, n_2) = e^{\frac{-(n_1^2 + n_2^2)}{2\sigma^2}}, \quad h(n_1, n_2) = \frac{h_g(n_1, n_2)}{\sum_{n_1} \sum_{n_2} h_g} \quad (2)$$

Gaussian low pass filter

5.1.3 *Imadjust*

Imadjust is the built in matlab function that was used for maximizing the contrast between the different elements in the images. The function maps the intensity values in an image I such that 1% of the data is saturated at both low and high intensities. The contrast maximization was used to create a mask applied to the image to be segmented in order to segment the colon.

5.1.4 *colorThresholder*

The matlab interface, *colorThresholder* was utilized for two different reasons in this thesis. It was firstly used as a means to find a suitable threshold in channels H and S in the HSV format to achieve the greatest segmentation of the colon. Secondly it was utilized to produce the ground truth images that were used to evaluate the performance of the classifier.

5.1.5 Principal Component Analysis

Principal component analysis (PCA) is a data analytical method used to reduce dimensions in the data set [36]. By finding a set of linear transformations of a group of correlated variables that agree to certain optimal conditions [38], one can reduce the amount of variables needed to represent the data. Such a condition could be to find the amount of uncorrelated variables that can represent the data without excessive loss of information. A common measurement of this would be to look at the amount of variables that reduces the reconstruction error. The goal of PCA is thus to look at the variables that amount for the largest variation in the data set [38]. In order to do this we have to take advantage of concepts such as eigenvectors, eigenvalues, and covariance matrices.

If \mathbf{X} is a vector of p random variables, then we are interested in finding a number of random variables n , where $n \ll p$. This is done by looking for a linear function $\alpha_1^T X$ that yields maximum variance where α_1^T is a vector of p constants.

$$\alpha_1^T X = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j \quad (3)$$

Linear function where we want to find a $j < p$ that maximizes the variance.

Assuming that the vector \mathbf{X} has a defined covariance matrix \mathbf{C} , we continue by calculating the covariance of matrix \mathbf{X} as seen in Eq.4

$$C = \frac{1}{N-1} \sum_{i,j}^N (x_i - \mu)(x_j - \mu) \quad (4)$$

Formula for calculating covariance matrix

The next step is to calculate the eigenvectors and eigenvalues of the covariance matrix in order to determine which principal components (eigenvectors) contain the most variation, as can be seen in Eq.5. A general assumption is that the principal component with the largest eigenvalue will contain the most variation in the data set [39].

$$(C - \lambda I_p)\alpha_1 = 0 \quad (5)$$

Formula for calculating eigenvector α_1 and eigenvalue λ from the covariance matrix C

Once the largest eigenvalues have been identified we are interested in the largest sum of eigenvalues which maximizes the variation in our new dataset of dimension $n \ll p$. The goal is

to add as many eigenvectors as needed to maximize variation but at the same time avoiding redundancy.

5.1.6 K-means

K-means is a clustering algorithm used for classifying data through clustering. It is considered to be the most widely used and effective clustering method due to its simplicity [40]. The algorithm works through the input of an integer c and a set of n data points which should minimize the objective function, in this case the squared error ε seen below in Eq.6 [41].

$$\varepsilon = \sum_{x \in X} \min_{c \in C} \|x - c\|^2$$

*Objective function to minimize in order to reduce the error
for values assigned to clusters. Where x and c are the respective data
points and clusters coordinates.*

(6)

The k means algorithm can either automatically generate the clusters necessary to reduce the objective function or these clusters can be calculated in advance [41]. Data points are then used as input for the algorithm and the data point is classified to the cluster that generates the smallest error ε . In the case of Eq.6 ε describes the Euclidean distance.

5.1.7 Local Binary Pattern

Visual descriptors or features are mathematical algorithms used for the purpose of describing and extracting information from images. The local binary pattern (LBP) method is one of such algorithms related to extracting texture from images. LBP is a local neighborhood threshold method, the idea is to within a window observe the surrounding pixels relation to the centering pixel within each window observed [42]. By doing this, one can use the alterations in relation to the middle pixel to describe the texture at a certain region. One of the greatest advantages of this method is that it is illumination invariant. This means that it is insensitive to light changes if all the nearby pixels within the observed window were affected equally [42].

As explained above LBP divides the image into grids (windows). More specifically within these grids it focuses on a 3x3 window resulting in a total of 9 pixel values where the center pixel will be used as reference as seen on Fig. 3 below. The center pixels R, G and B values are used to first calculate the difference against the surrounding pixels R, G and B values. However in this case, only one value is present for simplicities sake. This could for example represent an image in gray scale where values just like in the RGB space range between [0, 255].

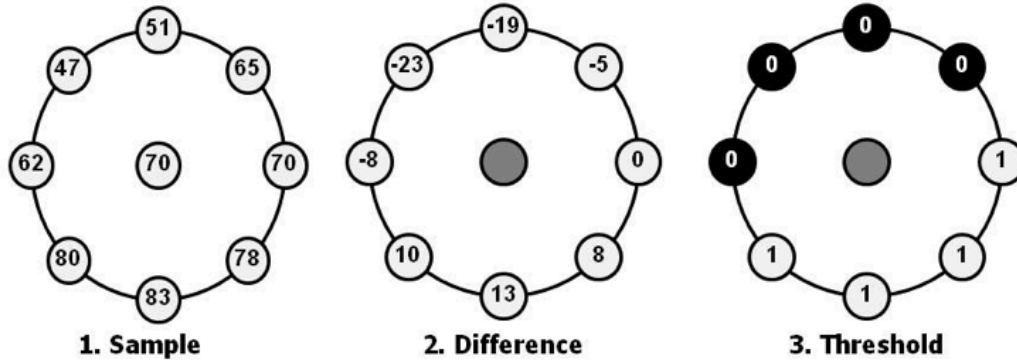


Fig. 3 LBP computation showing the 3x3 local neighborhood threshold method where the center pixel is used as a reference to generate an 8-bit binary representation of the region [43]. Illustration used with permission from Matti Pietikäinen.

When the difference has been calculated a binary coding is performed generating a certain pattern if the circle were to be traversed circularly, in this case the pattern obtained was [1 1 1 1 0 0 0 0]. Each pixel can attain a binary value [0 1], which means that there are $2^8 = 256$ total combinations that can be formed. These sets of combinations allows for a certain diversity, which can be used to describe the properties in each 3x3 window.

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Equation describing how binary coding of each pixel is performed in order to create the bitwise pattern. (7)

As was stated above, 256 distinct patterns can be generated through the binary mapping. Depending on the distribution of these binary values different features can be described within the image. One group of patterns contributing to large amounts of variation is the uniform pattern amounting to up to 90% of all patterns in a (8,1) neighborhood [42]. The uniform pattern includes consecutive amount of 1's or 0's when traversing circularly e.g. [0 0 0 0 1 1 1 1] while a non-uniform pattern would be e.g. [1 0 0 1 1 1 0 0].

Once the coding has been performed the label is calculated as can be seen in Eq.8. This is done by, taking the difference between center and the observed pixel, multiplying it with the binary constant s , and then multiplying by powers of two and finally summing over each of the surrounding pixel values, which ultimately yields the label.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (8)$$

Equation describing how the final descriptive value (label) is obtained from each 3x3 local neighborhood window

The next step is to calculate the histogram over all of the labeled values obtained by each 3x3 grid, observed in Eq.9, If a comparison is done of a histogram containing image patches of different sizes normalization is required, see equation 10.

$$H_i = \sum_{x,y} I\{f_l(x,y) = i\}, \quad i = 0, \dots, n-1$$

Calculating label frequency through the use of histogram presentation. (9)

$$N_i = \frac{H_i}{\sum_{j=0}^{n-1} H_j}$$

The following step is necessary when the image patches of compared histograms have different sizes (10)

6.1 SYSTEM DESIGN

The software architecture will be explained from two different perspectives to easily be able to parse through system design in a chronological and logical order. Henceforth, the two following diagrams are aimed to facilitate the understanding of the design.

- High-level design – The involving key blocks essential for implementation and improvement of the algorithm.
- Low-level design – This design describes the essential blocks needed and delves into each block in detail to describe the key components for each block containing the technical and specific programming features.

6.1.1 High-level design

The high-level design is a compressed version of the low-level design explaining the general program structure that defines the software displayed in Fig 4. Separated into four different main blocks, this scheme is a superficial explanation showing how an input image is finally outputted as classified.

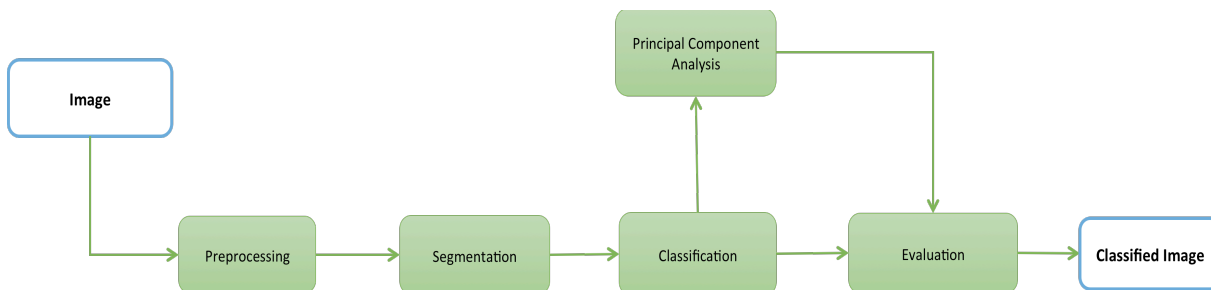


Fig. 4 Shows each step from the perspective of the high-level design architecture

For the sake of giving a broad overview of the software structure the following paragraphs will involve a general description of each block which later will be followed by a more extensive overview in the low-level design section.

- **Preprocessing**
 - **Elimination of black borders** – Removal of black frames surrounding image
 - **Dilatation of black edges** – In-painting of remaining black edges in each corner of the image.
 - **Dilatation of specular reflections** – In-painting of highly concentrated reflected light from regions surrounding mucous in the colon.
 - **Histogram intensity adjustment** – Maximization of contrast between bowel and non-bowels.
- **Segmentation**
 - **Color segmentation** – A color based segmentation, primary discrimination between stool and colon.
- **Classification**
 - **Feature extraction & PCA** – Extraction of LBP features as well as selection of most promising features.
 - **Training data** – Creation and structure of the training data.
 - **K-means classification** – Classification of each image using k-means cluster.
- **Evaluation**
 - **Evaluation** - Classifier comparison with doctor's evaluation to assess levels agreement.

6.1.1.1 First Stage: Elimination of Black borders

Each image received by the doctors may come in different formats. Some are obtained with a black frame surrounding the whole image, while others are received only with black edges at each corner. Figure 5 describes the process of eliminating the black borders from a high-level perspective. The first step to be performed before commencing the preprocessing was to normalize the R, G, and B intensity values. This was done by dividing each of the pixels color channels by the maximum value 255, yielding the new intensity range between [0 1]. Once this has been done the first stage in the preprocessing was performed where the goal was to eliminate the black border that frames each image. Although after the image has been processed, black edges remain in each corner, which will be taken care of in the next step.



Fig. 5 High-level diagram showing first level preprocessing of input images, elimination of black borders.

6.1.1.2 Second Stage: Dilatation of Black edges

As only the black edges remain from the previous preprocessing step the next stage was to eliminate the remaining black edges. Although if the black edges were to be removed the resulting image would not be rectangular thus the edges were replaced through dilation. The dilation uses close proximity pixels to paint the region to be dilated. In order achieve this a threshold was set in the V channel as seen in the high-level visualization in Fig 6.

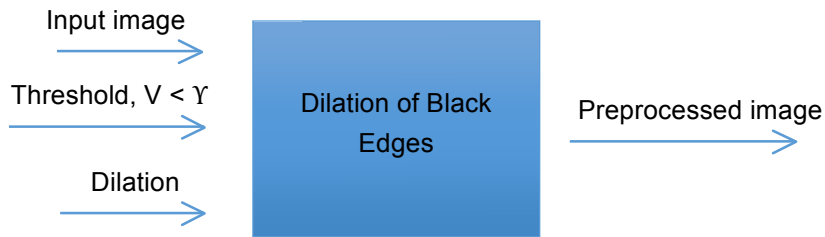


Fig. 6 High-level diagram showing second level preprocessing of input images, dilation of black edges

6.1.1.3 Third Stage: Dilatation of Specular Reflections

Once the image is rid of the frame remnants, the next step will be to take care of the reflections. The specular reflections are produced when the light from the colonoscopy camera is highly concentrated on regions surrounding a lot of mucous. These can be seen as white flares in the image. Such white flares turn out to produce a lot of problems in the fifth stage when the color segmentation is performed. The reflections can be removed through the usage of dilation similarly done in 6.1.1.2, the second stage. In this step the median of the image was subtracted with each point to search for high intensity values as shown in Fig 7.

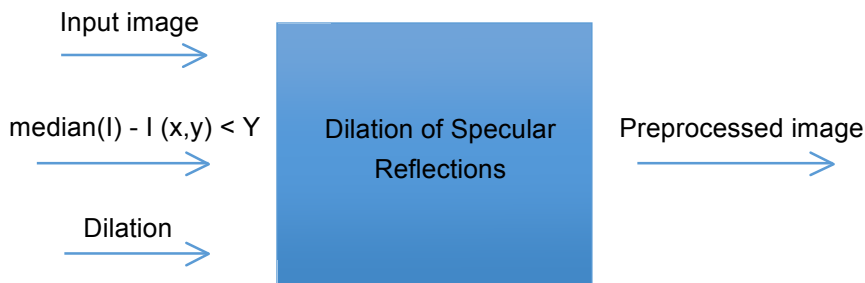


Fig. 7 High-level diagram showing third level preprocessing of input images, dilation of Specular Reflections where I is an arbitrary image.

6.1.1.4 Fourth Stage: Intensity Adjustment

In order to ease the segmentation done in 6.1.1.5, contrast maximization will be performed as seen in Fig 8. The Matlab function *imadjust* can be utilized to enhance the contrasts. This results in a prolonged intensity histogram in each of the separate channels R, G and B.



Fig. 8 High-level diagram showing fourth level preprocessing of input images, histogram intensity adjustment.

6.1.1.5 Fifth Stage: Color segmentation

The second block involves the color segmentation seen in the high-level design in Fig 9. The goal was to segment as much of the colon as possible but simultaneously avoiding segmentation of small parts of stools that have almost fused with the colon in color. An evaluation of which color format would be best suited for the segmentation was performed and finally the HSV format was elected with a threshold set in channel H and S.

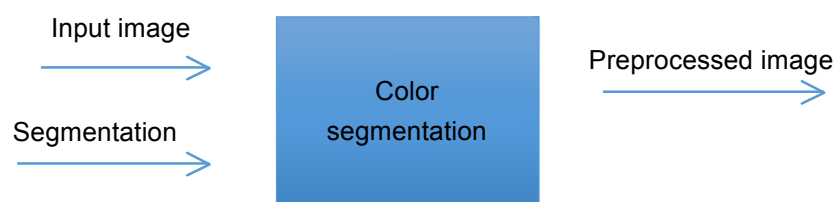


Fig. 9 High-level diagram showing segmentation of input images.

6.1.1.6 Sixth Stage: Training data

The training data is constructed by first performing an initial manual cropping of each stool type in the selected training images. Then, an automatic fragmentation is performed that cuts the image into perfect squares of any given size as seen Fig 10.

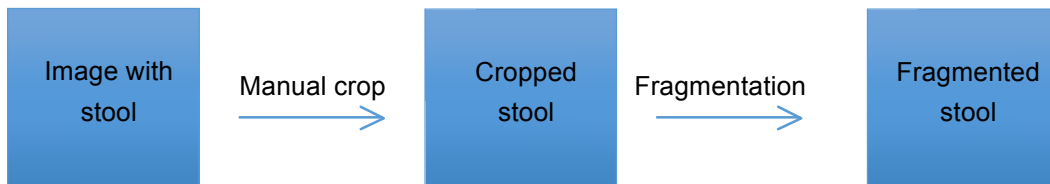


Fig. 10 High-level design, creation of training data

6.1.1.7 Seventh Stage: Feature extraction & PCA

Before the classification step is initiated the most prominent features need to be extracted and used in the classification of images. Fig 11 describes the process of the feature extraction and PCA performed in order to improve classification.

The texture feature, local binary pattern (LBP) was selected. LBP extracts 84 variables that describe the texture in the image. Thus a principal component analysis will be performed to look at what features are the most essential for classification.

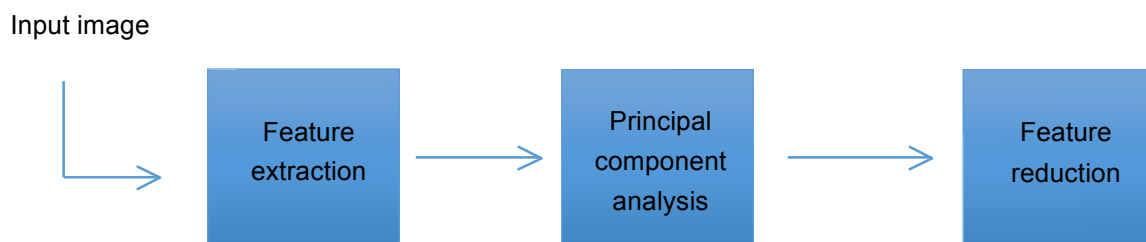


Fig. 11 High-level design of feature extraction and principal component analysis

6.1.1.8 Eighth Stage: K-means classification

Once the most important features will have been selected using principal component analysis, the next step will be to classify the images, seen in Fig 12. The images are classified using a k-means classification where each pixel in the image is assigned a class label. In order to visualize the classification, every class label will be assigned to a specific color. By doing this, each image can be repainted giving a visual representation of the classification.

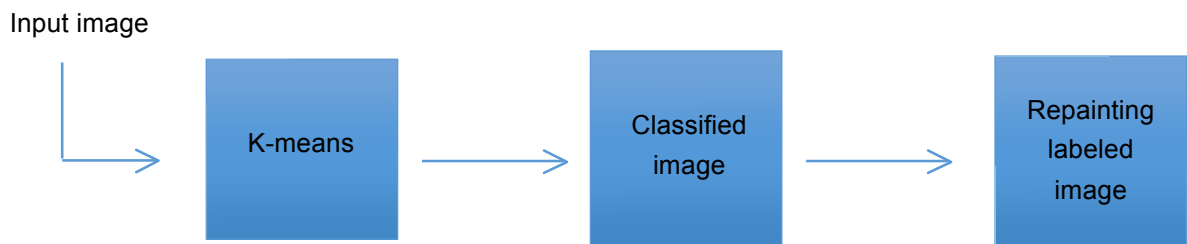


Fig. 12 High-level design K-means classification

6.1.1.9 Ninth Stage: Evaluation

One of the most common methods of evaluating classifiers is the use of k-fold cross validation. Although as reference images are not obtainable the classifier will be assessed through a study comparing the evaluations of the medics to the classifier in 79 images, see Fig 13. The classifiers performance will also be tested when classifying a real situation (non complete segmentation) compared to the classification of ground truth images (full segmentation).

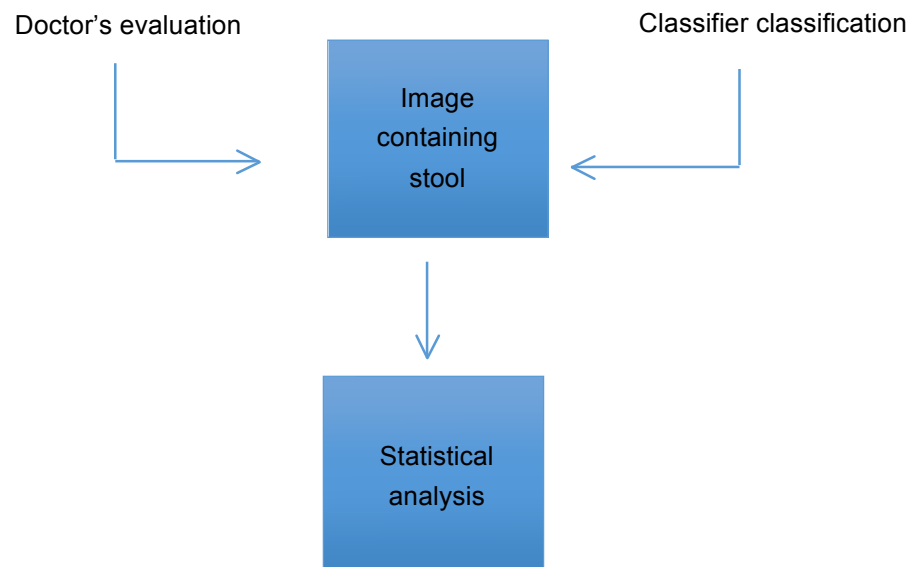


Fig. 13 High-level design showing evaluation of classifier compared to the medics.

6.1.2 Low-level design

The preceding unit concerned the high level design of the program structure. In this section, the program structure will be explored in a more detailed fashion, decomposing each of the above subunits in order to view the program architecture from a low-level design and thus in detail go through each of the steps involved. A diagram presenting each of these steps involved can be found below Fig. 14. The diagram divides the main blocks of the project into green boxes while the blue boxes further describes the process occurring within each block.

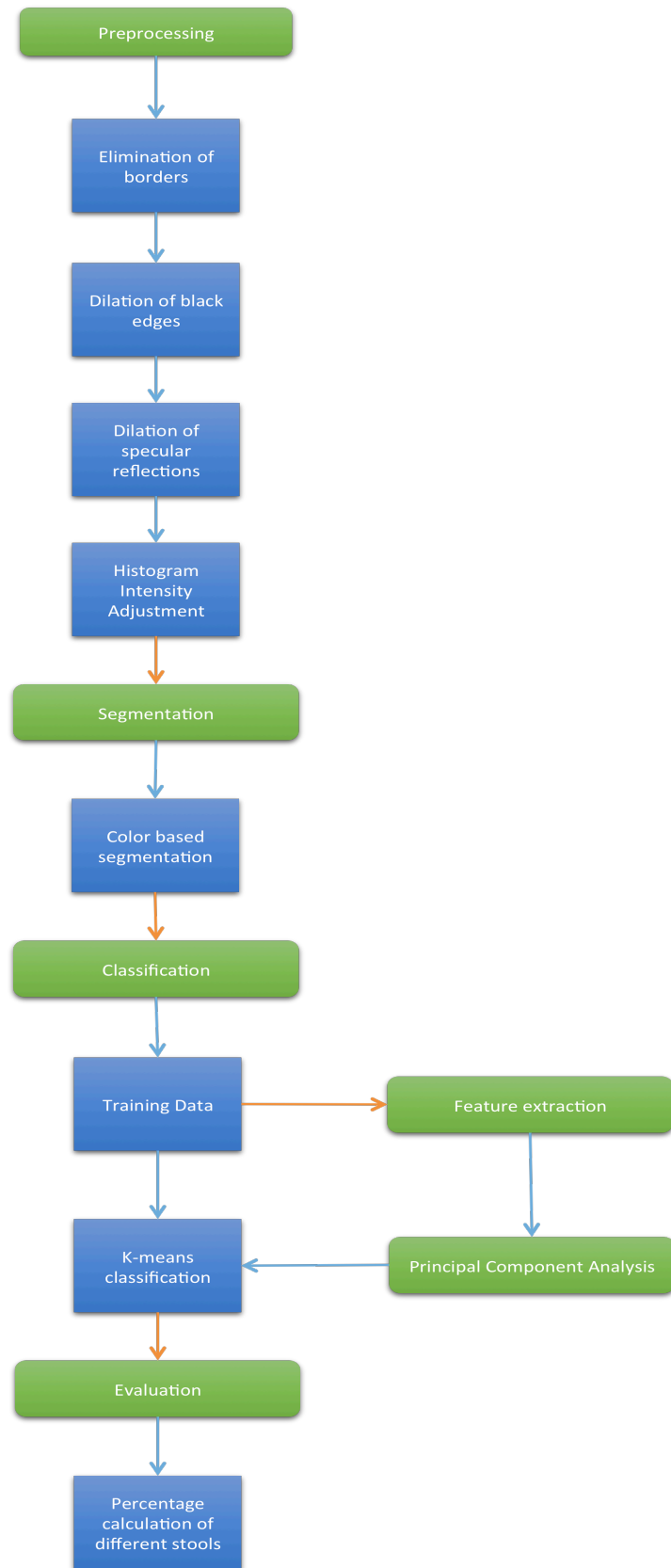


Fig. 14 Describes each step from the perspective of the low-level design architecture

6.1.2.1 First Stage: Elimination of black borders

The images obtained from the doctors have been received in different formats and therefore some images require an extra preprocessing step. The black borders are black frames surrounding the images. Before the black frames were dealt with, a normalization of the intensity range was initially performed. This involved dividing the each pixel in all of its R, G and B color channels with the maximum value 255 resulting in a new range [0 1]. By normalizing the intensity values the idea was to approach a range of values more similar to the other feature that will be used, in this case LBP see section 6.1.2.7. If not normalized it was expected that the Euclidean distance used to classify the pixels to its cluster described in section 6.1.2.8 would be largely affected hence leading to possible misclassifications.

For the elimination of the black frames, the images will initially be converted into the HSV format. The HSV (hue, saturation, value) format is a geometric reformation of the conventional RGB (red, green and blue) format. In the HSV format the V channel represents the brightness of the image. A threshold will be set in order to locate and remove the black border.

$$V < 0.03 \quad (11)$$

Threshold in V channel in the HSV format for locating and eliminating the black pixels

6.1.2.2 Second Stage: Dilation of Black edges

As the black edges have the same color characteristics as in 6.1.2.1 the same threshold can be used to locate them.

Once the black edges have been located, an in-paint/dilation using the pixels outside the black edges will be used. This leaves the edges looking like as they were originally part of the image but blurred. When the dilation is done a Gaussian filter will be applied. The idea of the Gaussian filter will be to reduce the manipulation of the image that had been made by the in-painting. Moreover, avoiding problems that could occur if small specular reflections exist among the pixels used when performing the in-paint, thus attenuating any high intensity added.

6.1.2.3 Third Stage: Dilation of Specular Reflections

Specular reflections are created when the light from the colonoscopy camera is highly concentrated on regions surrounding a lot of mucous. These are perceived as strong white flares in the image. Elimination of such will be done by calculating the median of the entire image and later subtracting each pixel with the median. If the value is larger than the threshold seen in Eq. 12 the pixel is considered to be a reflection.

$$\text{median}(I) - I(x,y) < 0.75 \quad (12)$$

Detection of specular reflections

For pixels exceeding threshold, dilation will be used to paint the regions with the neighboring pixels and finally a Gaussian filter will once again be applied.

6.1.2.4 Fourth Stage: Intensity Adjustment

The last step in the preprocessing stage performs a histogram intensity adjustment, to increase the contrast between the stool and colon. Applying the adjustment to each of the separate R,G,B channels saturating 1% of the data at low and high intensities, thus improving the segmentation in the next stage. The contrast maximized image will be utilized to create a mask used in the segmentation.

6.1.2.5 Fifth Stage: Color Segmentation

The objective of the second block is to perform a color based segmentation using the built in Matlab interface, colorThresholder. Five of the training data images were selected as references to identify the optimal threshold. Once the threshold was identified it was tested on a set of test images containing large variations in the appearance of both the stools and colon. The segmentation is performed through the use of a threshold set in the channel H and S in the HSV format. This is done in order to at the greatest extent only segment the colon. Different stool types have proven to have distinct levels of difficulties when segmenting. Therefore the segmentation focuses on segmenting as much of the colon as possible instead of focusing on segmenting the entire colon. In the case of the segmentation, the color-maximized image was used to solely create a logical matrix or also called a mask that was applied to the image to be segmented this can be seen in Fig. 21.

The idea when investigating the threshold was to perform a segmentation that only targeted the stool in images containing different stool types. Once a common threshold was encountered a logical matrix, “mask” was created where the regions containing colon were left untouched and regions containing stool were segmented. The following threshold was used:

$$\begin{aligned} 0.522 < H < 0.001 \\ 0.221 < S < 1.00 \end{aligned} \tag{13}$$

Thresholds set on channel H, S in the HSV format in order to segment colon, leaving stool intact.

After segmenting only the stool from the image, a logical NOT operation was performed on the logical matrix in order to reverse the previous segmentation.

$$Mask = \sim M \tag{14}$$

Logical NOT Matlab operator (\sim) used to invert mask.

This provides instead an inverted mask, segmenting the previous non-segmented regions. Thus, the new inverted mask will be targeting the colon while leaving the stool unaffected.

6.1.2.6 Sixth Stage: Training data

As described earlier the training data is made by first executing a manual cropping of different stool types in 19 unique images. Out of the 19 images used for stool training, 18 out these will be used for producing the training data for the colon. The 19th image was eliminated from the data used to train the colon because sufficient training data had already been generated in comparison to that of the stools. Furthermore, the 19th image appeared to be very different from the earlier elected images and it was suspected that it would introduce an exceptional case that might mislead the classifier. The stools that were cropped were the ones that appeared to be the most representable for each of the class types. As images labeled by clinicians could not be obtained. The images were labeled together with the help of PhD student Alain Sánchez, who was the primary contact to the doctors at the hospitals of Basurto, Cruces and San Eloy. Once the stools have been manually cut out, the images are further automatically fragmented into square matrices of size 80x80 as seen in Fig 15. The goal of the automatic fragmentation was to reduce the risk of contamination. By fragmenting the cropped images it was easier to adapt the amount of fragments that were going to be used for the feature extraction. This was important incase of the suspicion that misclassifications would be due to cross contamination between class types in the training data. Once the features were extracted, the mean of each of the features for every class type was calculated and used as clusters in the k-means clustering. The features extracted were the 84 LBP features as well as the R, G and B mean values.

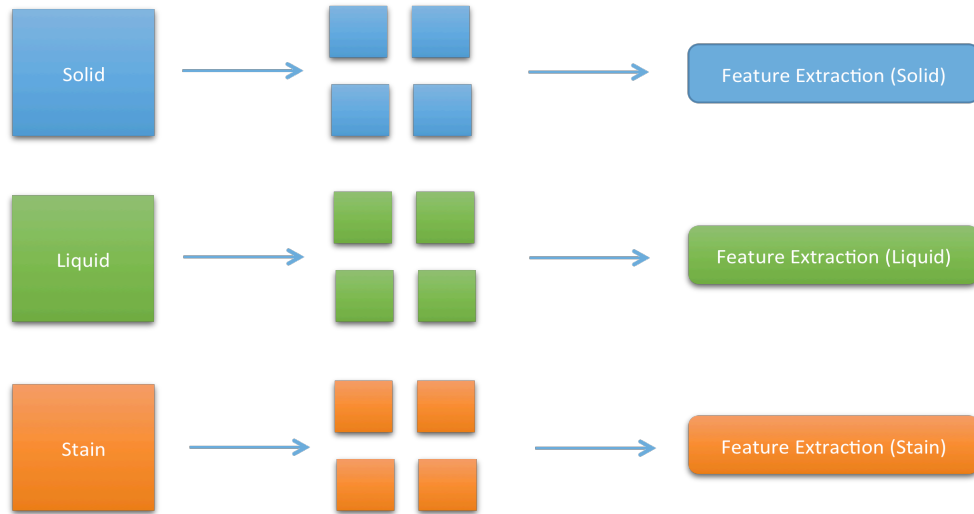


Fig. 15 Low-level design. Creation, fragmentation and feature extraction of training data.

6.1.2.7 Seventh Stage: Feature Extraction & PCA

A 4x4 window is parsed through the whole image and multiple 8-bit binary maps are created, as described in the methods section 5.1.7 Local Binary Pattern. Each of these 8-bit maps translates into a series of patterns. The pattern is then used to calculate a one-value label. A histogram is finally used to look at the frequencies of these one-valued labels.

The frequencies of the one-value labels are counted and used to describe the texture in the image. The LBP extraction results in 84 features, which most likely are redundant and therefore are not characteristic for our dataset. If this is the case a principal component analysis will be performed in order to reduce the amount of variables to the most important ones. Observing the eigenvalues of the covariance matrix containing the 84 LBP features of the fragmented images of each class, we can look at the variation between the features in the data set. Thus, to find the optimal amount of features resulting in most variation, classification tests will be run using k numbers of LBP factors where $k = 1, 2, \dots, N$.

6.1.2.8 Eighth Stage: K-means Classification

Once the most promising features have been found tests will be run in order to select the features that to the largest extent can distinguish between the different class types. Initial classifications using the selected features will be run using k-means executed using 6 clusters, solid, liquid, liquid clear, liquid dark, stain and colon. In order to evaluate the number of features that give best performance, 14 images containing large differences in appearances for both the colon as well as the stools have been selected. The initial discrimination will be to visually distinguish between the best classifications once the cases have been reduced the classifications will be compared quantitatively against each other.

6.1.2.9 Ninth Stage: Evaluation

The idea of the evaluation is to compare the performance of the classifier to each of the doctors separately and also joined. An assessment of the error between the evaluations of the doctors was made since a great difference was observed in their evaluations. Lastly an estimation of the classifier will be made in the case of where no colon is present (ground truth images) as well as when a normal segmentation is performed in order to observe how the colon cluster affects the overall classification. As for the ground truth images, 52 were selected where the images discarded were either clean colons or had bad resolution.

7. DEVELOPMENT

The following chapter will explain the process of how one starts with an unprocessed image and ends with a classified result using the software developed in this master thesis. The section includes an explanation of each of the stages divided up in the same amount of blocks as was done in the design except for evaluation which will be discussed in results. Moreover, the following segment contains a subsection named database where the data that was obtained is described.

7.1 DATABASE

The data was obtained in two separate occasions, the first part of the data collected contained, 18 colonoscopy images containing stools as well as a large set of images consisting of clean colons with encountered polyps. The second part of the data included an additional 181 images. Out of a total of 199 images there were 13 duplicates, 81 clean colons, 19 used for training and a few images with bad resolution leaving 79 total stool images. 52 of these were selected as ground truth images. Amongst the 79 images almost all of them contained mixed stool types were the dominating stool types were liquid and stain, thus few cases of solid were received. It is important to take into account that the idea of the colonoscopy is for the patient to be received with a clean colon, making it difficult to encounter good quality images containing stools..

7.2 FIRST STAGE: ELIMINATION OF BLACK BORDERS

As was explained in the design, the first step was to normalize the R, G and B intensity values that range between [0 255]. The normalization was done by dividing each of the intensities with 255, resulting with values ranging in between [0 1]. When the normalization has been performed the elimination of the border, shown in Fig 16, is attempted by first converting the image into the HSV format. After the conversion, a threshold was set on the V channel as mentioned in section 6.1.2.1.

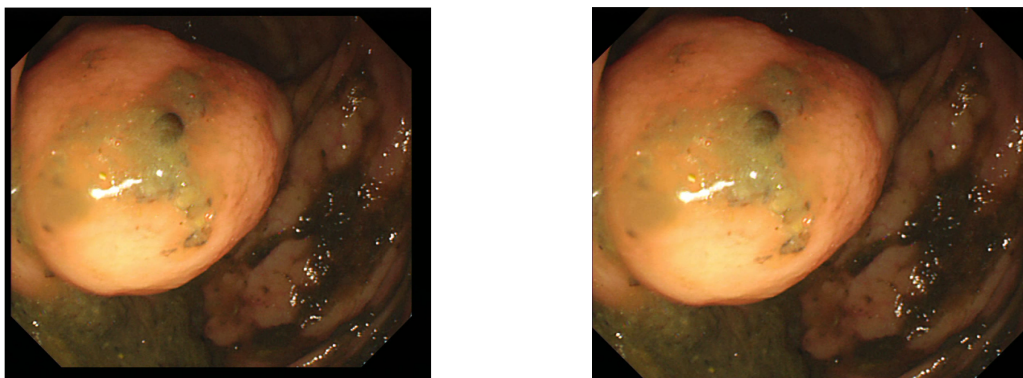


Fig. 16 Showing elimination of black borders

7.3 SECOND STAGE: DILATION OF BLACK EDGES

Once the borders had been eliminated the next step was to perform a dilation of the black edges. This was done by first locating the black edges with a threshold set in the V channel in the HSV format just as was done in the previous step 7.2. Once located, a disk with radius 10 was created using the matlab function *strel*. The disk was then used together with the built in matlab function *imdilate*. This paints the earlier allocated pixels with nearby pixels, as depicted Fig 17. When this was done, a Gaussian low pass filter was applied using the matlab function *fspecial*. This is done in order to reduce regions manipulated by the previous dilation step.

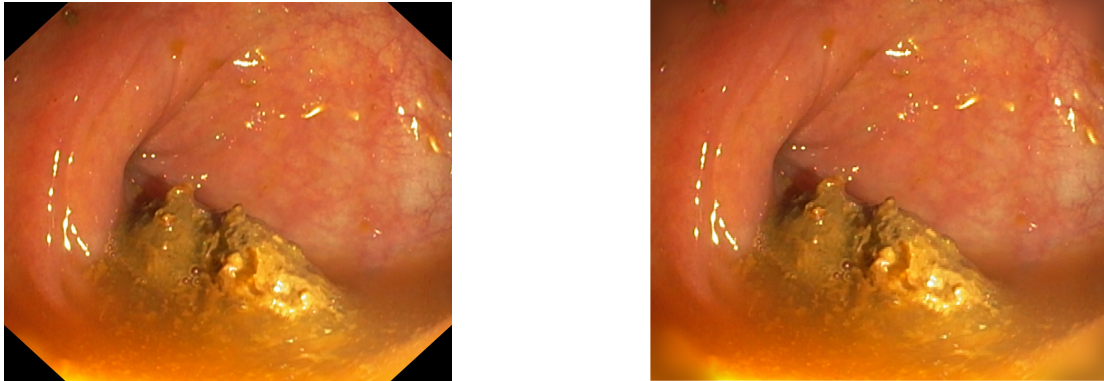


Fig. 17 Showing dilation of black edges

7.4 THIRD STAGE: DILATION OF SPECULAR REFLECTIONS

The specular reflections can be perceived as white flares in the image as seen in the right image in Fig. 18. These flares occur when the light of the colonoscopy camera is highly focused on regions surrounding mucous. As it is not easy to manage the camera due to limited space, this is a common phenomenon that produces problems both in the segmentation and in the classification. In order to allocate these flares, the median of the image was calculated. Furthermore, each pixel was subtracted by the median and if it exceeded a certain threshold the pixel was considered a flare. The identified pixels finally went through dilation and a Gaussian filtering as explained above in 7.3.

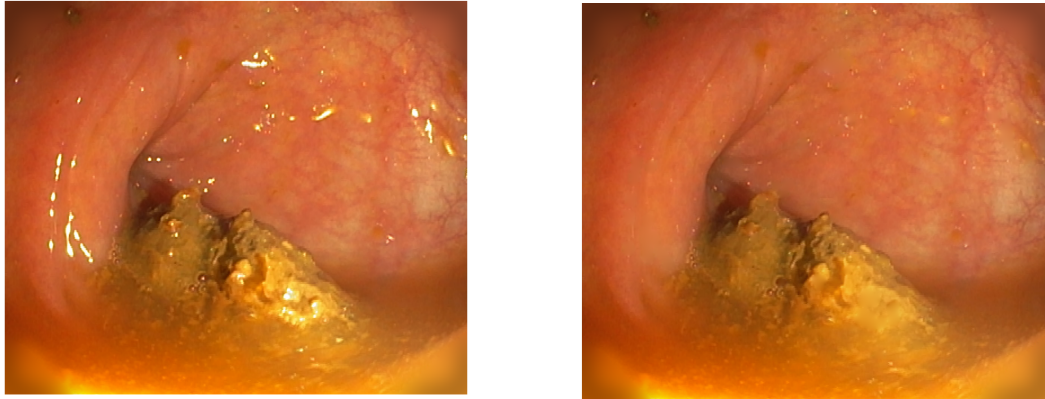


Fig. 18 Dilation of specular reflections

7.5 FOURTH STAGE: INTENSITY ADJUSTMENT

The final step in the preprocessing was to maximize the contrast between the colon and the stool seen in Fig 19. This was done in order to improve the color segmentation performed later in 7.6. It was done by performing a histogram intensity adjustment on each of the RGB color channels resulting in a saturation of 1% of the data at low and high intensities. Therefore extending the histogram, making it easier to isolate specific color ranges, this can be seen in Fig. 20.

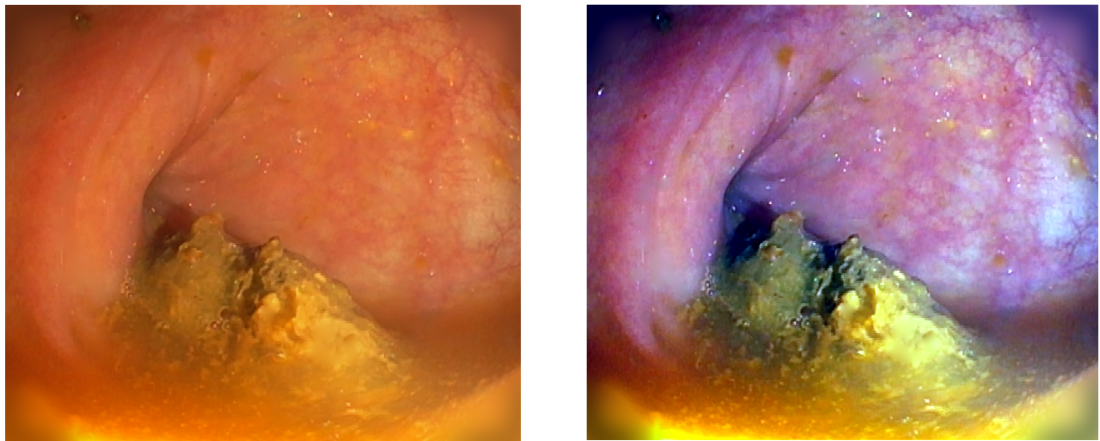


Fig. 19 Results after contrast maximization can be seen in the right image.

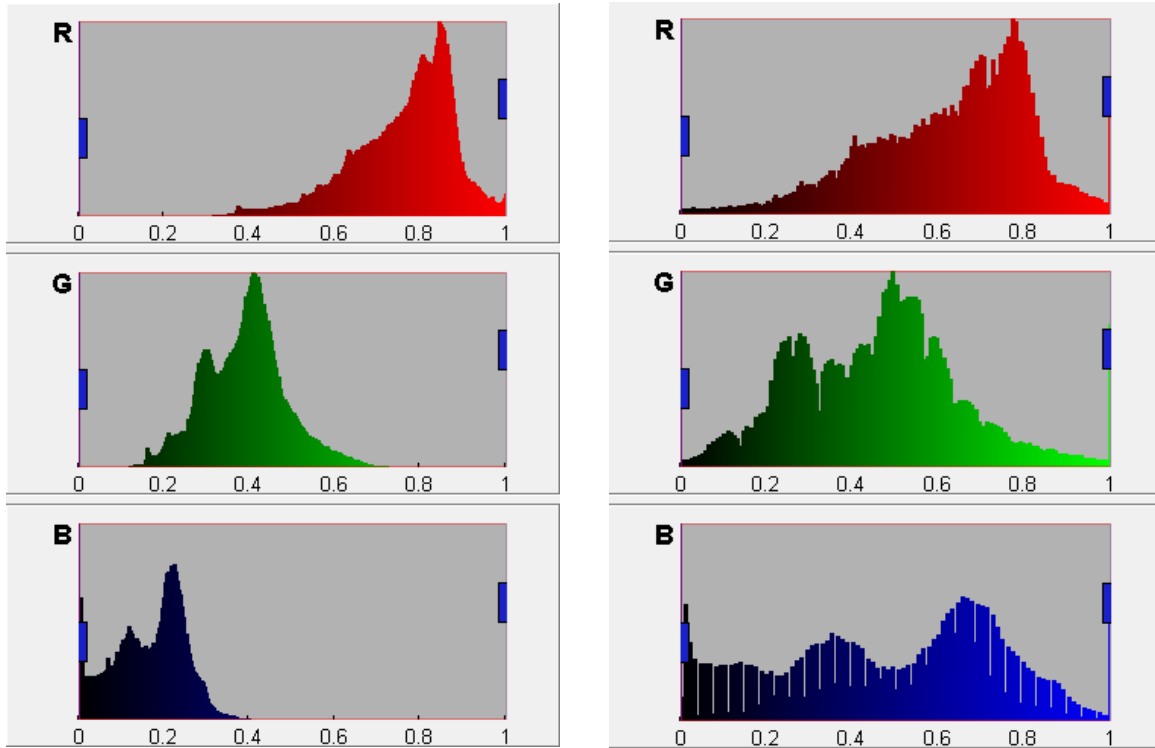


Fig. 20 RGB histogram before (left) and after (right) contrast maximization

7.6 FIFTH STAGE: COLOR SEGMENTATION

Once the preprocessing had been done the next step was to perform the segmentation. The main goal of this step was to find the range of the color spectra that contains the colon in order to eliminate as much of the colon as possible. This allows for two advantages:

Advantage 1: By initially segmenting the image, the classification time reduces in relation to how much of the colon has been segmented.

Advantage 2: As the colon is a lot more diverse in its characteristics the better the segmentation is the better the classification gets.

In order to perform the segmentation the images were converted to the HSV format with thresholds (see Eq.13) set on channels H and S. The thresholds were found using the built in matlab interface colorThresholder. The goal of the segmentation as stated in the design section 6.1.2.5, was to exclude as much colon as possible without segmenting the stool, see Fig 22. Although what is important to note is that the goal was not find a general scope of the spectra that would segment the entire colon. This is because it would most likely also lead to a segmentation of stains that are practically fused with the colon

A suitable threshold was found in the channels H and S and a logical matrix “mask” was produced.

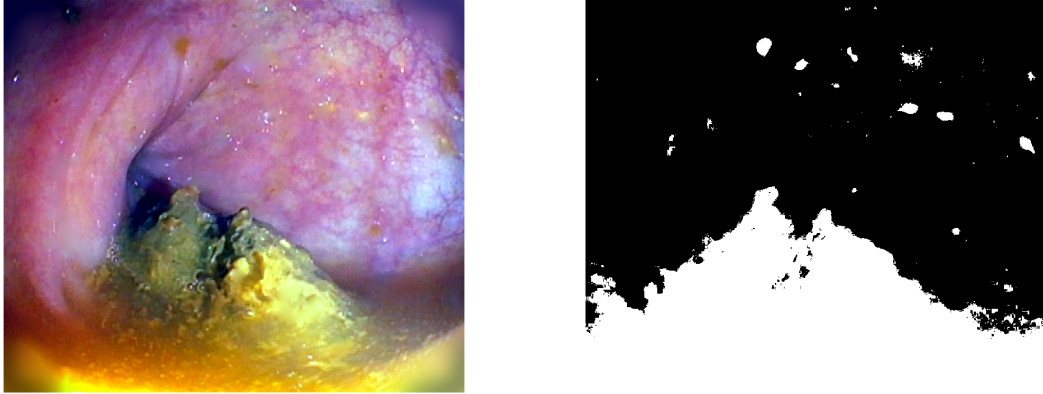


Fig. 21 Mask “logical matrix” created from contrast maximized image using threshold set at H,S channel

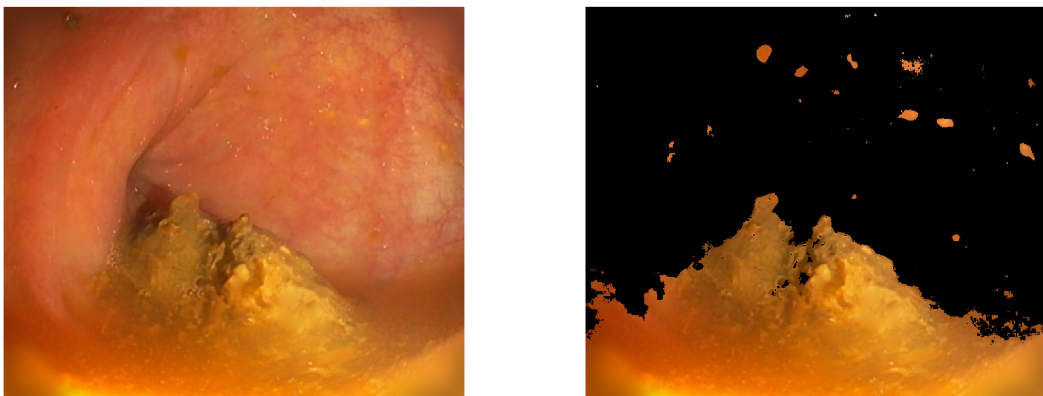


Fig. 22 Results when applying masked to the earlier preprocessed image

7.7 SIXTH STAGE: TRAINING DATA

As the objective for the doctors is to receive patients with clean bowels due to the bowel cleaning procedure that is gone through before the surgery. This limited the amount of images that could be obtained by the doctors, which also restrained the possibility of receiving representable images for training. Due to the lack of reference images containing isolated stool types, the classifier was instead trained on image fragments achieved through manual cropping of 19 images. As for the colon, the same set of images were used where only the clean colon

was cropped, although only 18 were utilized of the 19 images. After the manual cropping, the data was automatically fragmented into 80x80 fragments. Due to the lack of images and differences in stool sizes, the total of number of fragments differed between the class types. For instance, while solid stool generated 54, 80x80 fragments, stain generated 15, 80x80 fragments.

7.8 SEVENTH STAGE: FEATURE EXTRACTION & PCA

After the training data had been made the cluster centers to be used in the classification could now be calculated. The texture features were then extracted from the training data resulting in an 84-dimension feature vector. Due to the unlikeliness that high variation was maintained across all of the 84 features a principal component analysis was performed in order to identify the features with most variation in the data set. By identifying the features with highest variation it was expected that they would best describe the differences between the class types and thus lead to a better classification. Firstly, 84 LBP features were extracted from the 80x80x3 training data fragments, producing one 20x20x84 matrix for each of the 80x80x3 fragments. As explained above in section 7.7, the amount of training fragments for each of the class types differed. Therefore a code was written to utilize matlabs random number generator to select ten fragments from the earlier generated 20x20x84 matrices in each class. The fragments were then concatenated into the same matrix and the covariance matrix was computed. In order to observe the variation, the respective eigenvectors and eigenvalues were calculated. The first ten eigenvalues were selected and proportionally adjusted to that of the largest eigenvalue,

$$\lambda = [1, 0.338, 0.237, 0.184, 0.129, 0.107, 0.101, 0.095, 0.089, 0.087] \quad (15)$$

Ten largest eigenvalues calculated from covariance matrix containing all classes computed LBP features.

The values were later plotted maintaining the proportionality to the largest value resulting in the plot below Fig. 23. Indicating that most likely the optimal amount of variables to be used in the classification would be between 4-6 to variables. Classification tests were performed using the 4-6 LBP variables on 14 test images. It was found that 5 variables was the most efficient due to a large drop in the classification of the colon if more variables were implemented.

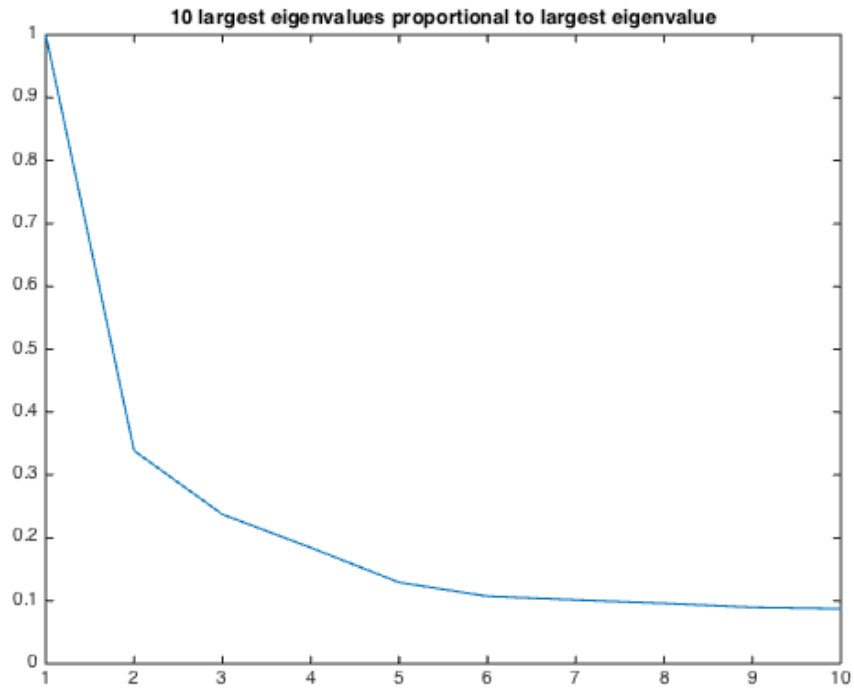


Fig. 23 showing 10 largest eigenvalues proportional to the largest eigenvalue

7.9 EIGHTH STAGE: K-MEANS CLASSIFICATION

Finally the classification was performed using the optimal amount of LBP features that described the most variation in the dataset. Finally 5 LBP features were used together with the mean values of each of the class types R, G and B values. Classification was performed using 6 clusters representing, solid, liquid, liquid clear, liquid dark, stain and finally colon. The choice of dividing up the liquid into two further subclasses was due to the variation observed in the characteristics and appearances across the images already obtained. It was suspected that one of the reasons why 6 classes seem to work better than 4 classes is due to the large variations in the colon results in frequent misclassifications.

Classification of the image is done pixel-wise and thus assigning a class label value between [1 7] representing each of the class types. The 7th value is assigned for the pixels that had been segmented and therefore not been classified. By separating the two different colon classes it facilitates performance evaluations of how well the classifier can classify the colon and the frequency of how well the segmentation can segment the colon.

In order to understand the classification the class labels are presented below:

- 1 – Solid - Red
- 2 – Liquid - Light yellow
- 3 – Liquid clear - Dark yellow
- 4 – Liquid dark - Orange
- 5 – Stain - Cyan
- 6 – Colon classified - Green
- 7 – Colon segmented - Blue

Fig. 24 below shows a classification example of the image that has been used throughout the development section to show the process from beginning to end. Moreover, the k-means clustering was run with maximum of 40 iterations in order to avoid local minima.

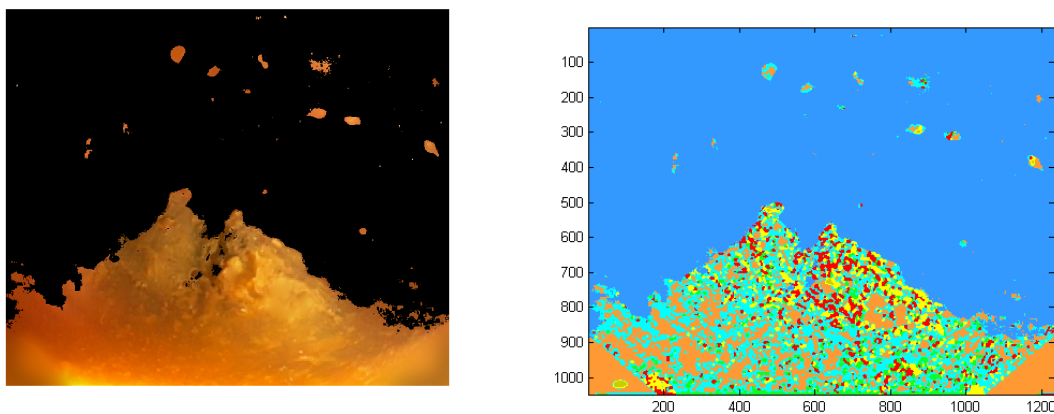


Fig. 24 Classification (right) of segmented (left) test image

Both medics assessed the classified image. Table 1 contains the percentages of each stool type, the BBPS score they estimated individually as well as the classifiers evaluation of stool percentages.

Stool type	Doctor 1	Doctor 2	Classifier
Solid	20%	20%	3.2%
Liquid	5%	25%	14.02%
Stain	0%	0%	12.4%
BBPS score	1	1	-

Table. 1 Showing percentages depicted by classifier as well as evaluated by clinicians.

8. RESULTS

The following section contains a statistical analysis of the classifier compared to the doctor's evaluations. To compare the classifier with that of the medics the evaluations will be compared by looking at the standard deviations and errors. This will be done between the classifier and each of the doctors individually as well as compared with the two doctors assessments combined. Preferably, when evaluating a classifier it would be more informative observing the positive, negative, false positive and false negative rates and ultimately performing a ROC plot to observe the area under the curve (AUC). Another way to evaluate the classifier would be to observe overlap classifications between the different classes. However, as there does not exist any true reference images ensuring the different stool types the standard deviations and errors were analyzed. As for the doctors evaluation only two were received limiting the information that could be extracted by the comparisons.

8.1 COMPARISON OF DOCTORS EVALUATION

The images received by the doctors contained a large variation in the appearances of the different stool types. The variation in appearance led to a deviation in not only percentages evaluated by the medics but also the different stool types present see section 9.2 Subjectivity in Judgment for more information. The bar graphs in Fig. 25 show the differences in percentages evaluated by each class over a set of 79 images all containing stools. It is worth to mention that the doctors are generally less interested in the exact percentages of each stool and more interested in the location in combination with quantity when determining the BBPS scores.

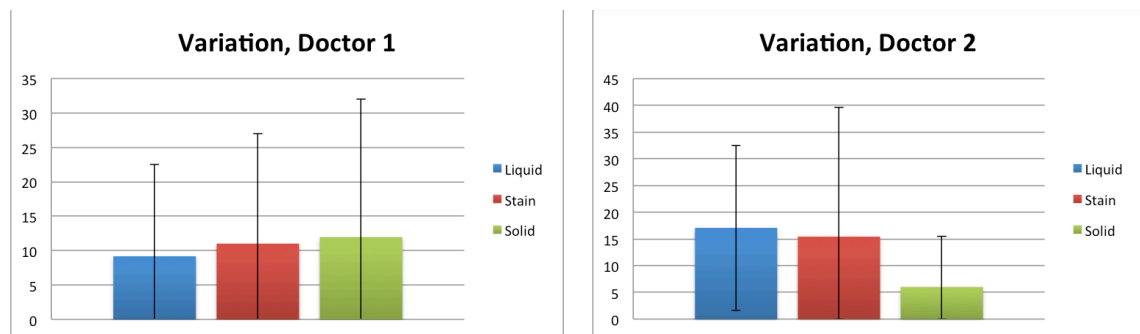


Fig. 25 Bar graph showing variation in the doctor's evaluations over a set of 79 images. Y-axis contains mean values for each class type and each bar contains its respective standard deviation.

While reviewing the evaluations it was observed that doctor 1 was more prone to select stain when an image was for example taken from a bad angle. It was found that doctor 2's approach in such situations was to decide if the stool type was either a single isolated stool type or mixed. This can be observed in Fig. 26 where the mean error between doctors was highest in stain. Additionally solid stools are more likely to be encountered in bigger quantities. Thus if the doctors disagree in such cases it is expected to see a large increase in the error, furthermore as

solids are underrepresented in the data set (see section 7.1 Database) such misclassifications will tend to add large errors.



Fig. 26 Absolute mean error measured between doctor's evaluations. Y-axis contains the difference between stool estimations for each image and finally the mean of the differences for each class type over all images.

8.1.1 Distribution of data in doctors evaluations

To present the distribution of the data, a box plot presentation was used. The box plot shows the most common percentages evaluated over the 79 images elected. It also gives an idea of the errors depicted above by showing the outliers in the dataset. The box plot in Fig. 27 reveals a number of difficulties when the clinicians evaluate the colonoscopy footage. One of such difficulties is the polarization created by the stain class, due to the stains being defined as an intermediate between liquid and solid. In Fig. 27 it can be seen that doctor 1 has large outliers (indicated by the red plus sign) in both the solid and stain class while doctor 2 has large outliers in the stain class although large values are in fact encountered frequently in the liquid class.

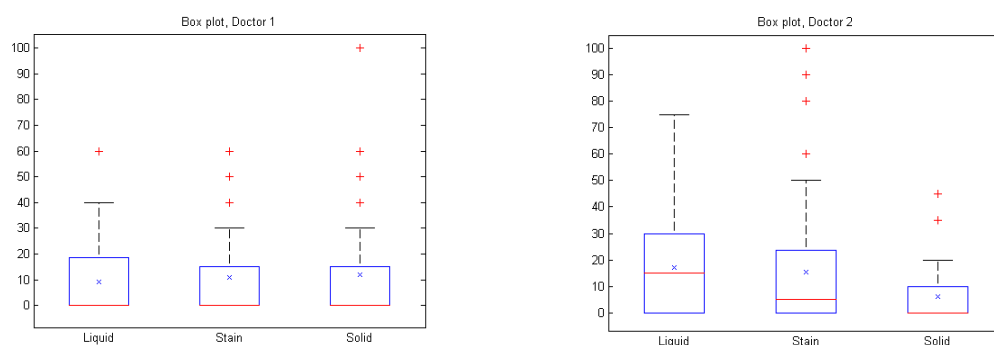


Fig. 27 Representing a box plot diagram of the doctor's evaluations. The y-axis contains the stool percentage estimations made by the doctors.

8.2 CLASSIFIER COMPARISON WITH DOCTORS

The bar graph in Fig. 28 displays the mean error for each class and their standard deviations. The representation illustrates a pattern not too surprising, due to most images that were obtained contains the class types liquid and stain. A reason for this is that the purpose of performing the colonoscopy to begin with is on patients having clean colons, thus limiting the data set too fewer distinct cases. Moreover, as solid stools create more visual hindrances, it is less common to encounter solids stools in patients. Seeing that there are also less clear cases of solid stools in the received data set it is therefore more difficult to evaluate the differences between the doctor's estimation and that of the classifier.



Fig. 28 Bar graph showing variation in k-means classification in all 79-stool images.

In the mean error diagram in Fig. 29 the classifier is compared to each of the clinicians individually while as in Fig. 30 it is compared to the combined evaluation of both doctors. A similar trend can be seen between all of the evaluations showing a dominating error in the classification of the liquid class. Though, as the dataset involves more liquid and stain classes in combination with the disagreement between the doctors amongst these classes, it is expected to see a larger error here.

Another factor explaining the error in Fig. 29 lies in the visual descriptors or features used in the classification. Even though the solid class is less represented in the dataset, its error is not as large as what could be expected if compared to the distribution of solid stools encountered in Fig. 26 doctor 1, which is an indication that the classifier is able to cope with classifying solid stool.

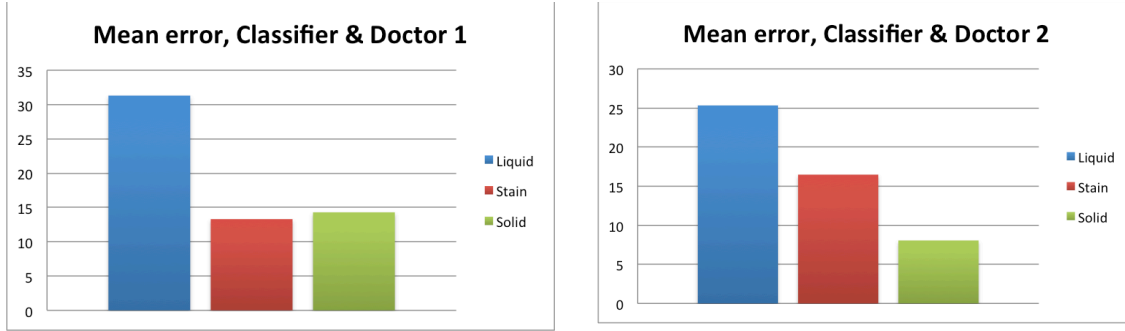


Fig. 29 Absolute mean error, comparing classifier to that of the doctors individually. Y-axis contains the difference between stool estimations for each image and finally the mean of the differences for each class type over all images.

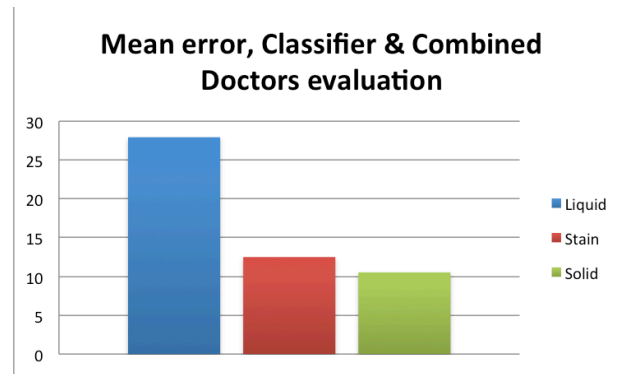


Fig. 30 Absolute mean error measured between classifier and the combined doctor's evaluation. Y-axis contains the difference between stool estimations for each image and finally the mean of the differences for each class type over all images.

8.2.1 Distribution of data in classification in all stool images

The box plot in Fig. 31 shows the distribution of the classified classes in all 79-stool images and as expected, it can be seen that the liquid class is dominating with frequently high estimations resulting in a large mean value. The plot also reveals the skewness of the dataset. In Fig. 31 it can be seen that the classifier tends to classify high values for the liquid. The high frequency of larger classifications for liquid indicates the possibility of misclassifications.

The maximum values for each of the stool types (liquid, stain, solid) as can be seen above in the clinician's box plot in Fig. 27 were, [60 60 100] for clinician one and [75 100 45]. Seeing that the classifiers maximum value for liquid is close to 80% and it is not considered an outlier, it is suspected that there are a large number of classifications between the values 50%, as it belongs in the upper quartile, up to close to 80%. Amongst these classifications it is evident that there lie misclassifications of what actually are stains.

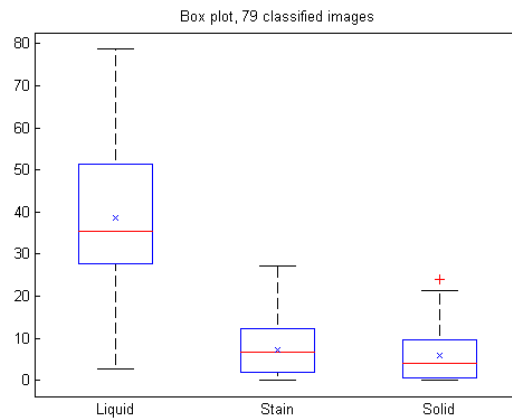


Fig. 31 Representing a box plot diagram of the classification of all 79-stool images. The y-axis contains the stool percentage estimations of the classifier.

8.3 CLASSIFIER COMPARED WITH GROUND TRUTH CLASSIFICATION

The bar graph seen in Fig. 32 shows 52 classified images that initially went through a normal segmentation based on the threshold explained in the design section 6.1.1.5 followed by a normal classification. For the ground truth images a manual segmentation was performed in each of the images to see how the classifier is affected when no colon is present, this can also be seen as testing the “best scenario case”. The bar graph in Fig. 32 shows a smaller mean value in all classes for the ground truth segmentation, which is expected. This drop in the mean can either be that the classifications have been distributed over all classes or it could reflect the reduction of misclassification of the colon, as all the colon has been segmented. As for the standard deviation, it is possible that we are observing either a lot of medium-misclassification or also not as frequent misclassification but instead large outliers when a misclassification does indeed occur.

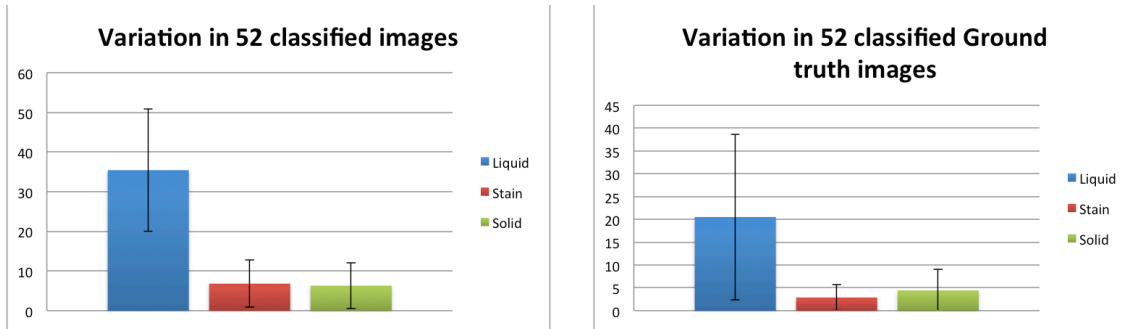


Fig. 32 Bar graph showing variation in classification of 52 images and the respective 52 ground truth images produced by the same set. The Y-axis represents the mean and each bar contains the classes standard deviation.

As predicted previously above the, liquid is once again the dominating class, strengthening the theory of the liquid class being frequently encountered in the stool images. The mean error plot seen in Fig. 33 indicates that the stain and solid classifications are likely to be not too far apart between the different classifications. However, the liquid class indicates the possibility of larger spread in the estimation of the stool present.



Fig. 33 Absolute mean error, comparing a normal classification to the classification of 52 ground truth images. Y-axis contains the difference between stool estimations for each image and finally the mean of the differences for each class type over all images.

8.3.1 Distribution of data in classification of images selected for ground truth comparison

If we observe the distribution of the data by looking at the box plot in Fig. 34 below we see that the misclassifications amounting for the error in the liquid class in Fig. 33 is most likely due to frequent large classifications. This frequency does however reduce in the ground truth image classifications.

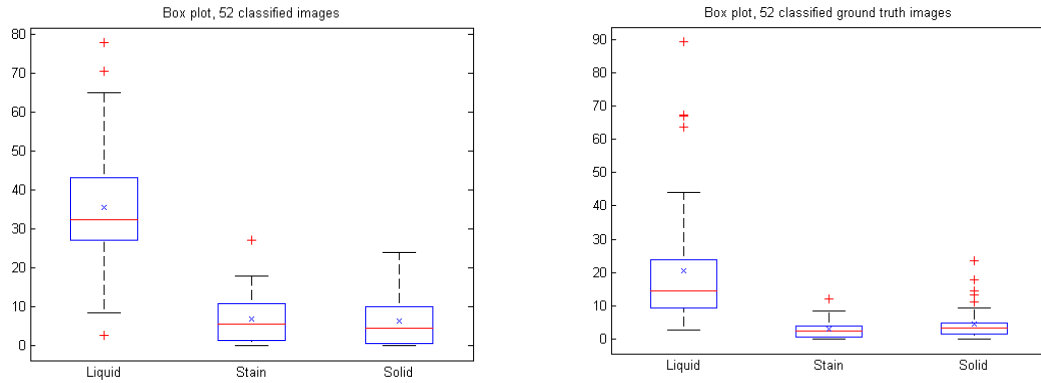


Fig. 34 Box plot diagram of the normal classification of the as well as ground truth images. The y-axis contains the stool percentage estimations of the two classifications.

Table 2 and 3 contain data collected from the classifiers prediction as well as the doctor's evaluations for the ground truth and the original image. Indications can be seen that there is in fact a possibility that the classifications are being distributed across the other classes in the ground truth classification see table 2 and 3. The spread could possibly be misclassifications although by observing the images classified it can be seen that the classifier has slightly improved its classification of the class type stain, see Fig. 35. Finally its possible to state that the colons diversity does in fact affect the classification of stains indicating that the two clusters contain samples that are in close proximity of each other. Nevertheless, In order to certify this it would be required to have reference images to observe positives, negatives, false positives, false negatives and finally all overlaps in stool classification which would basically be to look at all the different false negatives cases.

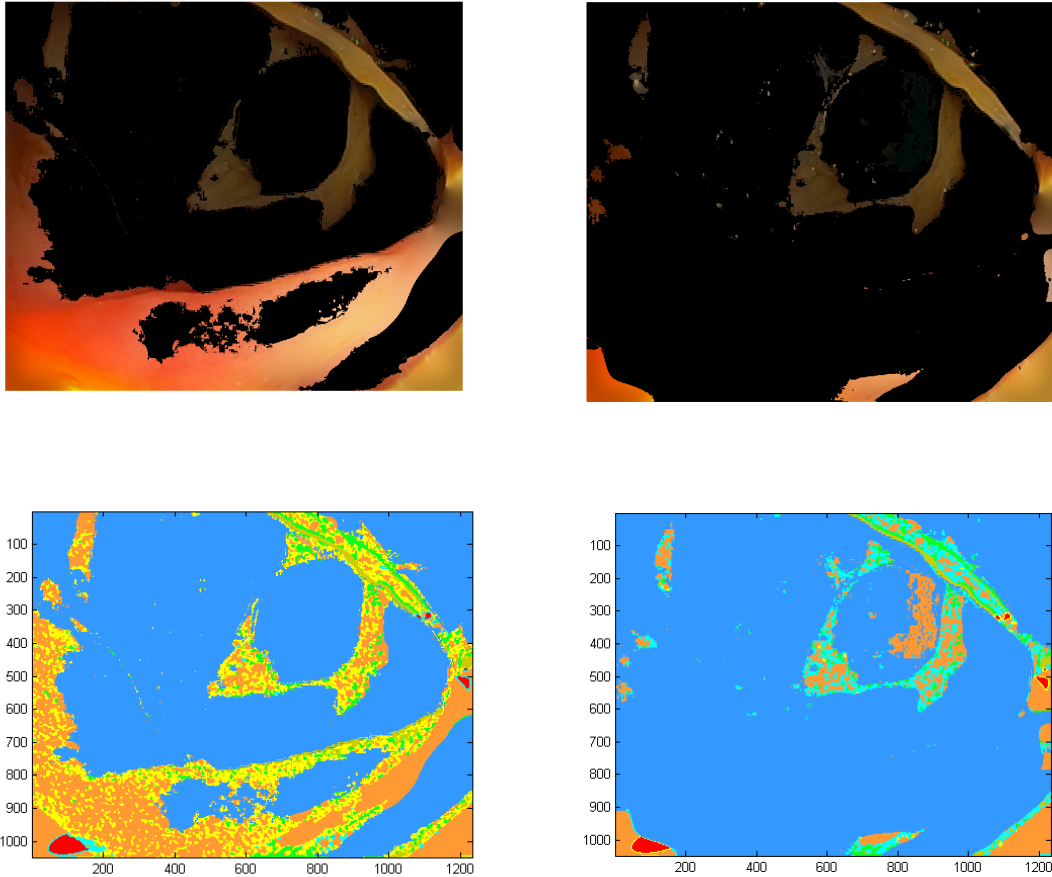


Fig. 35 Showing the normal segmentation (upper left corner) and classification (down left corner) versus ground truth segmentation (upper right) and classification (down right) of image 192.

Ground truth Image 192	Classifier	Doctor 1	Doctor 2
Solid	0.32%	0%	10
Liquid	8.2%	50%	0
Stain	4.58%	0%	80

Table. 2 Percentage estimation of both doctors as well as the classifier of ground truth image 192

Normal segmentation Image 192	Classifier	Doctor 1	Doctor 2
Solid	0.41%	0%	10
Liquid	31.59%	50%	0
Stain	0.54%	0%	80

Table. 3 Percentage estimation of both doctors as well as the classifier of normal segmented image 192.

8.4 ADDITIONAL CLASSIFIED IMAGES

The following subsection involves additional segmentations and classifications of images, seen in Fig 36, 37 and 38 as well as the evaluated percentages of stool estimated by both of the clinicians and the classifier. The evaluations for each respective image can be seen in tables 5, 6 and 7. Table 4 below describes the color labels used in the classification of each class type and also the color label marking the part of the colon that has been segmented.

Class type	Color label
Solid	Red
Liquid	Yellow
Liquid clear	Dark Yellow
Liquid dark	Orange
Stain	Cyan
Colon (classified)	Green
Colon (segmented)	Blue

Table. 4 Displaying color labels for each of the classes.

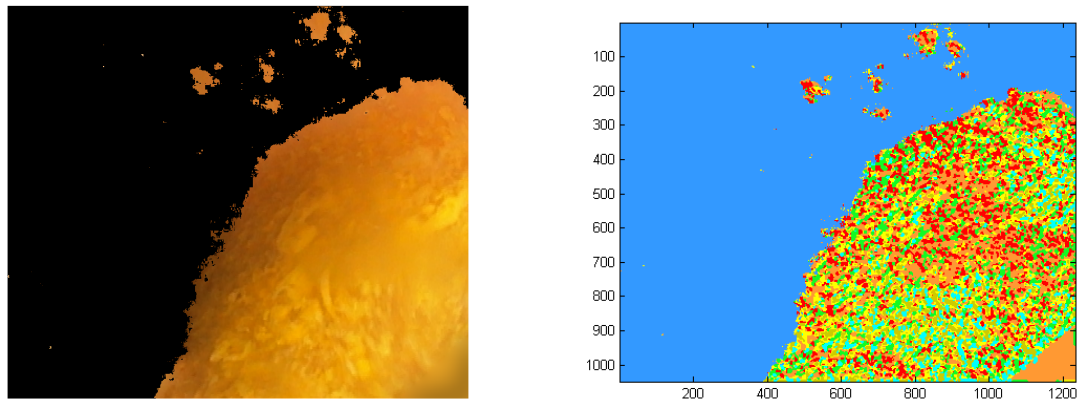


Fig. 36 Showing segmentation and classification of image 155

Image155	Classifier	Doctor 1	Doctor 2
Solid	8.34%	0%	0%
Liquid	22.5%	30%	30%
Stain	3.65%	10%	0%
Colon	65.48%	-	-

Table. 5 Displaying the estimated percentages of each stool present in image 155.

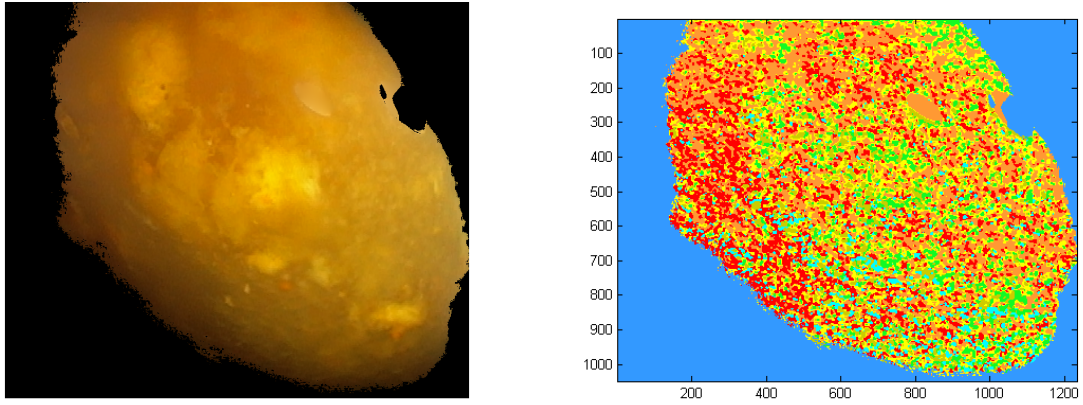


Fig. 37 Showing segmentation and classification of image 171

Image171	Classifier	Doctor 1	Doctor 2
Solid	15	0%	10%
Liquid	42.98	40%	75%
Stain	9.12	40%	0%
Colon	37.95	-	-

Table. 6 Displaying the estimated percentages of each stool present in image 171.

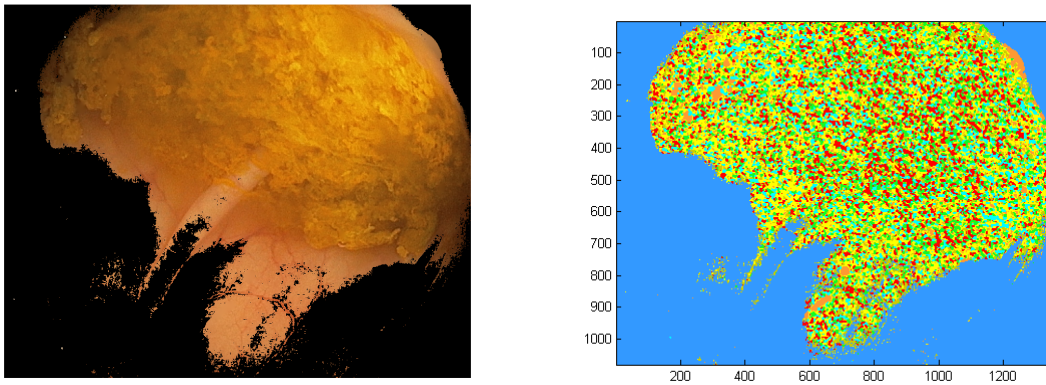


Fig. 38 Showing segmentation and classification of image 182

Image182	Classifier	Doctor 1	Doctor 2
Solid	10.92%	10%	10%
Liquid	28.36%	30%	30%
Stain	9.13%	0%	0%
Colon	51.57%	-	-

Table. 7 Displaying the estimated percentages of each stool present in image 182.

9. DISCUSSION

The following section will discuss issues that have been encountered during the course of the project as well as the limitations that have been faced.

9.1 ISSUES CLASSIFYING THE COLON

As the classifier has problems with classifying the colon a decision was made to look into these misclassifications. What was encountered was a huge difference in the possible values that could be attained in each of the separate color channels. Four images are displayed below in Fig. 39 and Fig. 40 with their respective original image as well as a 2D plot containing the RGB values over line segment. The line segments plotted were chosen over homogenic regions to ensure minimal noise such as reflected light or traces of stool. The four different images preview, two clean colon images Fig. 39 and two colon images containing stool Fig. 40.

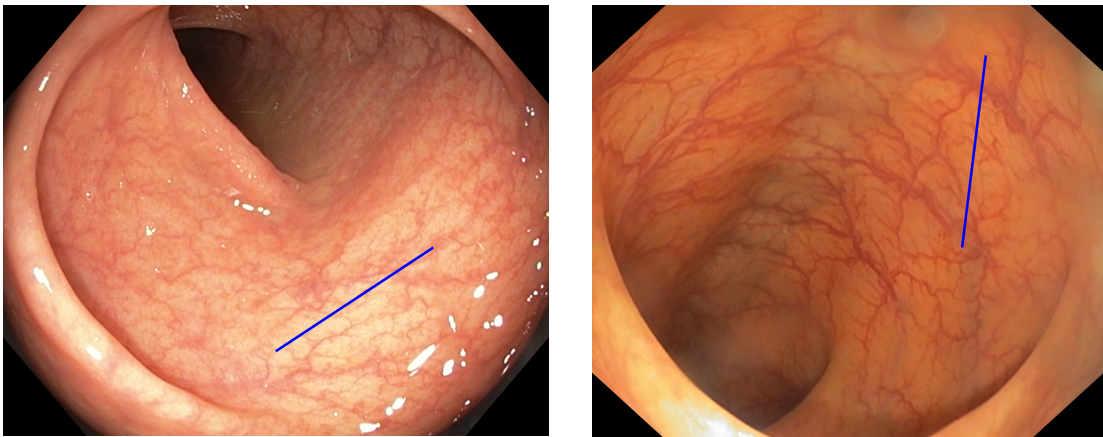


Fig.39. Clean colons that would both receive a BBPS score of 3. *The blue lines correspond to the line segment plotting the RGB values while traversed.*

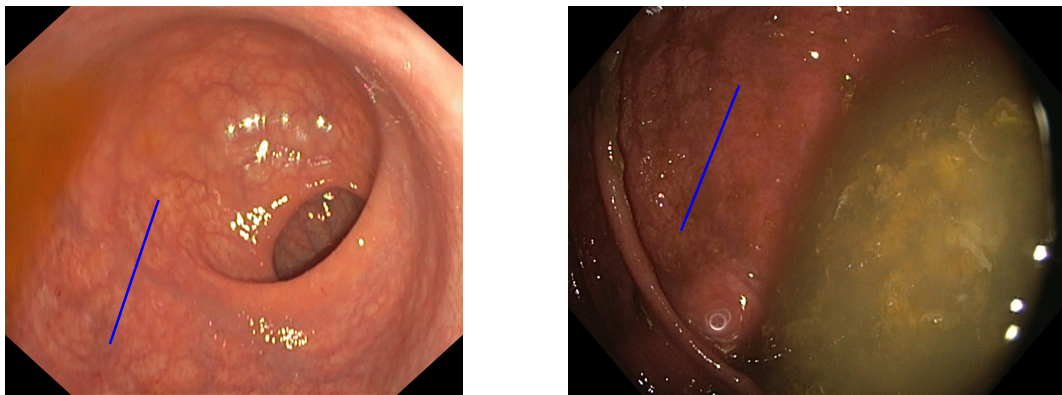


Fig. 40 Colon that both contain stool. *To the left BBPS value 3 and to the right BBPS value of 2.*

Figures 41 and 42 are the RGB values encountered when traversing the blue line in Fig. 39 and Fig. 40. The graphs below are in the same order as the ones above.

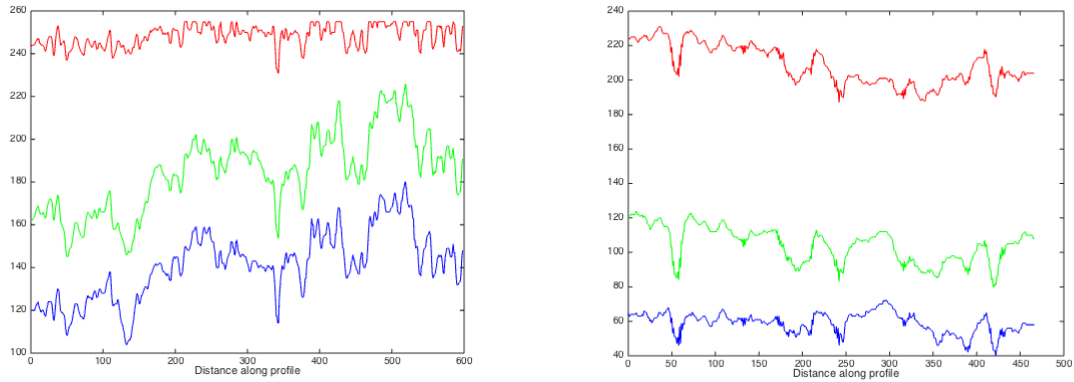


Fig. 41 Showing RGB values on the line segment of the two clean colon images

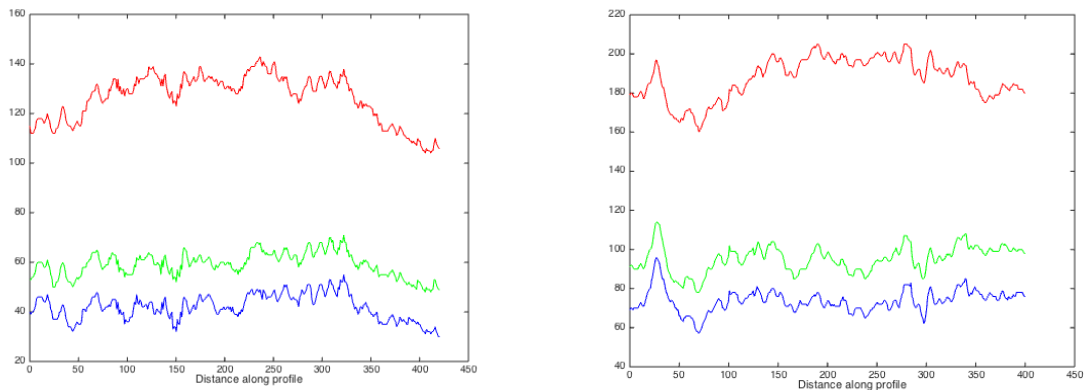


Fig. 42 Showing RGB values on the line segment of the dirty colon images

The variation in the RGB space in the tissue shows the large quantity of R, G and B combinations possible. This in turn, most likely increases the cluster size and therefore increases the probability that misclassifications occur. A reason for the difference in colon is that the images received originate from different patients. Another reason that could explain the variation is if there was an attempt to remove stool residue during the procedure. This would most likely leave the lumen with a darker color similar to Fig. 39 (right).

9.2 SUBJECTIVITY IN JUDGEMENT

The goal of this project since the very beginning was to create a program that could hopefully in the future reduce the subjectivity in the assessment of colonoscopy videos/images. However due to the broadness of the standards used to evaluate such data, it has become evident that observing differences in doctor evaluations is something to be expected. Moreover, in the evaluation sent to the doctors 13 duplicates existed in the data. Even though the 13 images obtained the same BBPS score, doctor one evaluated the images 6/13 with different percentages while doctor two did the same with a probability of 4/13. The outcome was expected although the idea was to show that when analyzing a large amount of images we lose our ability to recognize images we have already seen. If this idea now instead were to be applied on a multitude of images simulating a full day of work, we would expect to see a scattered distribution. A doctor's goal is to make sure that the BBPS score is the same over all evaluated images, which was the case for the duplicate images. However, when comparing the two doctors evaluations, 62 out of the 199 images contained a different BBPS score. This shows that there is an obvious gray zone where the doctors have difficulties of agreeing with each other.

9.2.1 Stain vs. liquid

Another disagreement seen amongst the doctors is the fine line between the two different stool types, stain and liquid. This is due to that a stool is considered a stain when it is in between liquid and solid, which is to say when it is not fluid or hard resulting in a gooey appearance. Due to this broad definition it is common to find a disagreement in the evaluations and thus making it difficult to compare doctors evaluation with the classifiers classification.

9.3 IMAGES TAKEN FROM DIFFERENT ANGLES

An interesting result came up showing the classifier sensitivity to uneven illumination changes. This can be seen in the classified images in figures 43, 44 and the results from both the classifier and the two doctors evaluation shown in table 8. As the LBP texture features are illumination invariant this does not affect the calculated texture features unless the illumination is uneven. Meaning that only some pixels within the LBP grid parsed over the image have increased RGB values while the others remain the same resulting in a different 8-bit binary pattern. However, a more common possibility for the difference in classification could be due to that in the k-means clustering the mean of the R, G and B channel are used to classify the image. This is most likely the reason to why the classification is sensitive to color changes and also why the colon can appear to be so different.

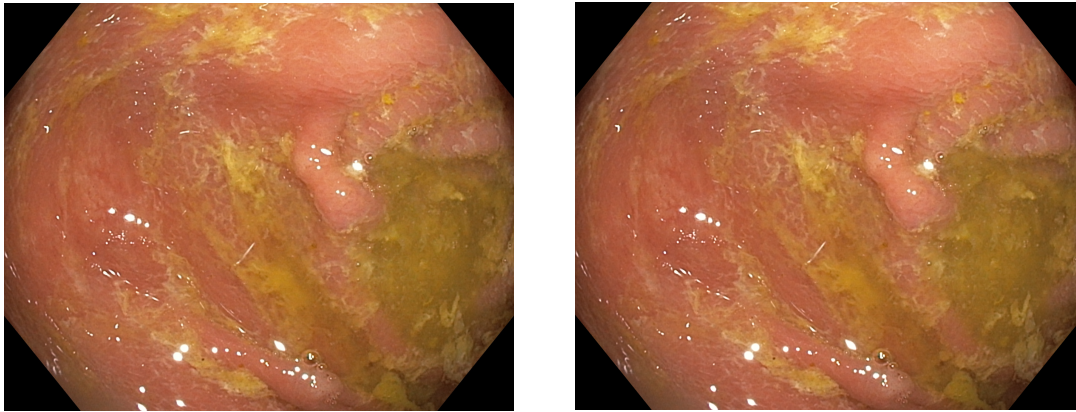


Fig. 43 Two images taken from a slightly different angle containing stains. Left image 168, right image 169 (see classification below in Fig. 44)

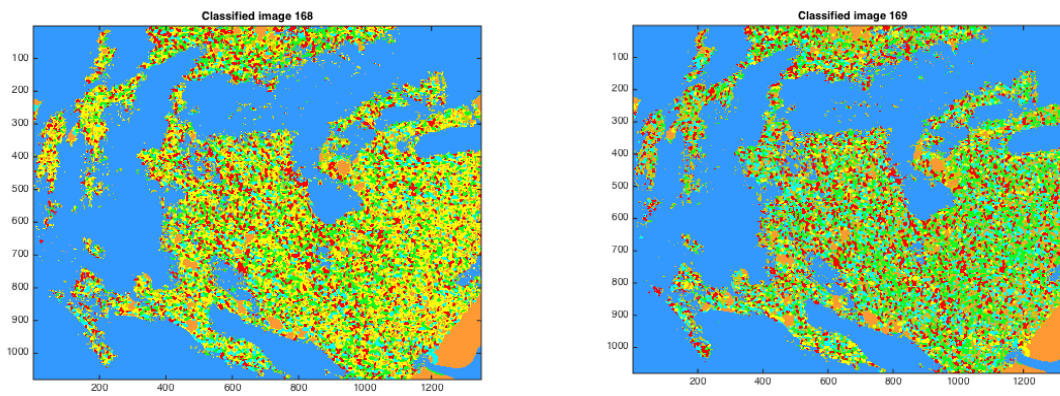


Fig. 44 Classified images 168 and 169, see table 8 below for detailed evaluation by doctors and classifier

	Image168 (Classifier)	Image169 (Classifier)	Image168 Doctor 1	Image169 Doctor 1	Image168 Doctor 2	Image169 Doctor 2
Solid	8.91%	9.6%	0%	0%	0%	0%
Liquid	27.5%	21.5%	50%	50%	10%	10%
Stain	6%	9.7%	0%	0%	20%	20%

Table. 8 Representing the percentages depicted by classifier as well as evaluation by clinicians in images 168 & 169

9.4 THE RECEIVED DATA

When working with real samples coming from patients we have very little control of what data we can obtain. One of the greater limitations with the work was the dependence of images from the medics e.g. the time it took to receive them as well as what images that could actually be obtained. Another limitation was that once the images were received it could not be avoided that most of them either simply just had bad resolution or other technical errors occurred during the colonoscopy. This limited the data that could be used in training and testing, which in turn affected the classifier and also limited the amount of different cases that could be analyzed.

9.5 DILATION & GAUSSIAN FILTERING OF IMAGES

Dilation was used at two different instances during the preprocessing stage, dilation of black edges and specular reflections. A suggestion of using a missing information class instead of performing the dilation was made by professor Carolina Wählby. The reason for this is that there is a risk that false texture information is added and could therefore create false classifications. In the case of the dilation and Gaussian filtering of the black edges, a missing information class would most likely have been a better choice, as these regions do not need to be classified. As for the specular reflections the situation is slightly different. Due to large quantities of mucous on the colon there are a lot of issues with light reflections. By introducing a missing information class, this could lead to large parts of the image not being classified. For example in situations where there is a lot of clean colon present we also expect to see a lot of mucous, thus large amounts of reflections. In such cases the classifier might not be able to classify essential parts of the image. Another reason was that the idea in the project was to work from the assumption that the error (e.g in percentage calculation of stools) caused by the dilation would be less than avoiding classifying the white flares. It was therefore decided to implement dilation and Gaussian filtering in order to get rid of the white flares.

9.6 EVALUATION OF CLASSIFIER

In order to evaluate a classifiers performance in the long term it is essential to have reference images. In the case of this project the optimal case would have been to have images labeled by different clinicians. The reason for this can be seen in the box plot in Fig. 27 where two doctors had evaluated the same amount of 199 images but with strikingly different evaluations of what stool types that were actually present. However, if labeled images were to be obtained this would open up for the possibility of observing the frequency of positives, negatives, false positives and finally false negative classifications. By being able to extract such data we could look at the similarity as well as the dissimilarity in classification between the ground truth and normal classified images. By observing the amount of correct classified pixels divided by total pixels, also known as the Jaccard coefficient. We could begin to observe the frequency of correct classifications. Also as previously described in section 8 Results. It would also be

interesting to observe the overlap of misclassification between each of the stool types and once again observe the frequency. This would reveal the bias of the classifier and would lead to the possibility of finding further reasons to why certain misclassifications occur.

9.7 ALTERNATIVE FEATURE SELECTION METHODS

In this master thesis the feature selection or feature decomposition method used was principal component analysis. By taking 10 training data fragments randomly selected from each class type and later extracting the 84 LBP features in order to finally aggregate them into one data matrix. The covariance matrix of the aggregated data was calculated and the eigenvalues were examined for features that amounted for most variation. There is always the possibility that a larger variation does not necessarily have to mean that a feature will be better at distinguishing between the classes. It could also simply mean that there is a large difference between the classes but it may not be ideal for discerning between them.

Alternative selection methods are e.g. wrappers, filters and embedded feature selection methods. In this project 84 LBP features were extracted, the goal of the feature selection is to find an efficient way to examine which of the 84 LBP features are the most essential. There are different approaches of how to examine or select the features that are to be tested. Examples of such are forward selection, where more variables are added gradually while as in backward elimination, one starts with a large subset and eliminates variables that are considered excessive [44]. There are also different ways to evaluate if the selected variables will improve the classification. In a k-means clustering, one way would be to measure the difference in the Euclidean distance when a subset of \mathbf{S} variables are used and compare it to $\mathbf{S}-1$ or $\mathbf{S}+1$ [44]. After evaluating a range of different subsets the idea is to find a subset \mathbf{S} that minimizes the objective function seen in for example equation 6. In addition to minimizing the objective function it would also be interesting to compare the classifier with the ground truth classification and perform a k-fold cross validation to further study the performance.

10. CONCLUSION

The goal of the project was to introduce a software that could detect and classify stools in colonoscopy footage. During the course of this project there have been a lot of problems that proved that the task was more difficult than initially expected. These issues have however opened up for new ideas and hints on how the original problem should be dealt with. A necessary addition would be to encounter images that could be used as references in order to be able to quantitatively measure the performance of the classifier. This limitation increased the difficulties of finding the root to various issues that were confronted.

There is still a lot that can be done within the field to improve the classification. It's apparent that in order to improve it, more visual descriptors will have to be implemented e.g. color visual descriptors. The use of local binary pattern has proven to be efficient in the classification. Although in order to improve, it is evident that the mean values of the RGB color space should be replaced. The reason for this is that the advantage of using LBP is due to its illumination invariant properties. In colonoscopy videos it is difficult to require specific standards of how images should be taken and thus it's important that this property is maintained through the features used to classify the colon.

One of the largest restrictions in the project was the variation in the tissue. The colon can easily vary a lot in its appearance making it difficult to build a classifier solely based on texture and mean RGB values. A trend could be seen in the classifications, the more colon still present after the segmentation, the more susceptible the classifiers were to performing misclassifications.

Finally, through performing this project many things have been learnt and thus there are a lot of things that can still be done. In section 11. Future work, some ideas of what could either be implemented or should be looked into are presented.

11. FUTURE WORK

To improve the classification there are different things that need to be done, the most important part would be to create a set of reliable reference images that have been segmented and labeled together with the help of the professional opinions of different clinicians experienced in the field. By having approved labeled reference images it would be possible to quantitatively compare the classifier run with different set of visual descriptors and measure performance directly.

As for the visual descriptors one possibility that could be interesting to implement is the GIST as color descriptor that extracts a 4x4 grid where the orientations histograms are extracted [45].

Other important aspects would be to disregard the mean RGB values as a feature for classification. Issues concerned are for example overlaps in RGB space between images, also different illuminations such as change of lighting or more presence of mucous tends to greatly affect classification, see section 9.3 Images taken from different angles.

Lastly other possibilities could be to look at classifiers with different decision boundaries for example smooth linear, higher power or orthogonal boundaries.

12. ACKNOWLEDGEMENTS

I would like to thank the University of Deusto, more specifically the institution DeustoTech for giving me the opportunity to perform my master thesis in their department. I would also like to thank my supervisor Maria Begoña Garcia Zapirain Soto, my subject reader Carolina Wählby for important advice, PhD student Alain Sánchez for aiding when in need and with concerns regarding the preprocessing. I want to thank the doctors in hospitals Basurto, Cruces and San Eloy for providing me with the images and the evaluations and lastly many thanks to PhD student Alan Dolhasz for the advice of how to further expand on evaluating the classifier.

13. BIBLIOGRAPHY

- [1] Kenneth L. Bontrager, *Textbook of Radiographic Positioning and Related Anatomy*, 8th ed. Louis, Missouri: Mosby, 2014.
- [2] Medical University of South Carolina. (2012, March) Digestive Disease Center. [Online]. <http://www.ddc.musc.edu/public/organs/colon-rectum.html>
- [3] Healthline. (2014, November) Healthline. [Online]. <http://www.healthline.com/human-body-maps/descending-colon>
- [4] National Cancer Institute. (2015, February) National Cancer Institute. [Online]. <http://www.cancer.gov/about-cancer/what-is-cancer>
- [5] World Cancer Research Fund International. (2015, January) World Cancer Research Fund International. [Online]. <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/colorectal-cancer-statistics>
- [6] Leticia Angulo-Rodríguez, "Automated System for Quantifying the Level Of Preparation in Colonoscopy," *Information Models and Analyses*, vol. 1, pp. 226-235, 2012.
- [7] Harvard Health Publications. (2016) Drugs.com. [Online]. <http://www.drugs.com/health-guide/colon-polyps.html>
- [8] Saltzman, "Bowel preparation before colonoscopy," *Gastrointestinal Endoscopy*, vol. 81, no. 4, pp. 781-794, 2015.
- [9] National Cancer Institute. (2014, November) National Cancer Institute. [Online]. <http://www.cancer.gov/types/colorectal/screening-fact-sheet>
- [10] Cleveland Clinic. (2015, January) Cleveland Clinic. [Online]. https://my.clevelandclinic.org/health/diagnostics/hic_Colonoscopy_Procedure
- [11] Lai, "The Boston Bowel Preparation Scale," *Gastrointestinal Endoscopy*, pp. 620-625, 2009.
- [12] Levin, "Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology," *Cancer Journal for Clinicians*, vol. 58, no. 3, pp. 130-160, December 2008.
- [13] Vinay Kumar Kapoor. (2013, August) Medscape. [Online]. <http://emedicine.medscape.com/article/1949039-overview>

- [14] PubMed Health. PubMed. [Online].
<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0022231/>
- [15] PubMed Health. (2013, September) PubMed Health. [Online].
<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0072487/>
- [16] BruceBlaus. (2013, October) Sections of the colon.
- [17] Hoeijmakers, "review article Genome maintenance mechanisms for preventing cancer," *Nature*, pp. 366-374, May 2001.
- [18] Setlow, "The Wavelengths in Sunlight Effective in Producing Skin Cancer: A Theoretical Analysis," *Proceedings of the National Academy of Sciences*, pp. 3363–3366, September 1974.
- [19] Friedberg, "feature DNA damage and repair," *Nature*, pp. 436-440, January 2003.
- [20] PubMed Health. PubMed Health. [Online].
<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0024119/>
- [21] Pancreatic Cancer Action. (2016) Pancreatic Cancer Action. [Online].
<https://pancreaticcanceraction.org/about-pancreatic-cancer/diagnosis/faq/difference-benign-malignant-tumour/>
- [22] American Brain Tumor Association. (2014) American Brain Tumor Association. [Online].
<http://www.abta.org/brain-tumor-information/diagnosis/malignant-benign-brain-tumors.html?referrer=https://www.google.se/>
- [23] Seyfried, "On the origin of cancer metastasis.," *Critical Review Oncogenesis*, pp. 43-73, 2013.
- [24] National Cancer Institute. (2012, August) Colon Cancer Expanded Version.
- [25] Genetics Home Reference. (2016, May) Genetics Home Reference. [Online].
<https://ghr.nlm.nih.gov/primer/howgeneswork/genesanddivision>
- [26] DeSalvo, *GOD?* Boulder, Colorado, USA: Pithy Publishing Inc, 2006.
- [27] Patil, "Carcinoma cells induce lumen filling and EMT in epithelial cells by soluble E-cadherin-mediated activation of EGFR," *Journal of Cell Science*, pp. 4366-79, December 2015.
- [28] Cooper, *The Development and Causes of Cancer*, 2nd ed. Sunderland, Massachusetts,

USA: Sinauer Associates, 2000.

- [29] Malladi, "Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT," *Cell*, vol. 165, no. 1, pp. 45-60, March 2016.
- [30] Parkin, "An overview of the immune system.," *Lancet*, pp. 1777-89, June 2001.
- [31] National Cancer Institute. (2015, March) National Cancer Institute. [Online].
http://progressreport.cancer.gov/detection/colorectal_cancer
- [32] Ernstoff, "A randomized blinded clinical trial of a rapid colonic lavage solution (Golytely) compared with standard preparation for colonoscopy and barium enema.," *Gastroenterology*, pp. 1512-16, June 1983.
- [33] Muthukudage, "Color Based Stool Region Detection in Colonoscopy Videos for Quality Measurements," in *Lecture Notes in Computer Science*, D Hutchison, Ed.: Springer-Verlag, 2011, vol. 7087, pp. 61-72.
- [34] Ness, "Predictors of inadequate bowel preparation for colonoscopy.," *Am J Gastroenterol*, 2001.
- [35] Heresbach, "Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies," *Endoscopy*, pp. 284-290, April 2008.
- [36] Jolliffe, *Principal Component Analysis, Second Edition*.: Springer-Verlag.
- [37] Ringnér, "What is principal component analysis?," *Nature*, vol. 26, no. 3, pp. 303-304, March 2008.
- [38] Jackson, *A User's Guide To Principal Components*.: John Wiley & Sons, 1991.
- [39] Wold, "Principal Component Analysis," *Chemometrics and Intelligent Laboratory Systems*, pp. 37-52, 1987.
- [40] Finley, "Supervised k-Means Clustering," Department of Computer Science, Cornell University, 2008.
- [41] Arthur, "k-means++: The Advantages of Careful Seeding," Stanford University, Stanford,.
- [42] Ojala, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE*, vol. 24, no. 7, pp. 971 - 987, July 2002.
- [43] Pietikäinen. (2010, March) Scholarpedia. [Online].
http://www.scholarpedia.org/article/Local_Binary_Patterns

- [44] Guyon, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research* 3, pp. 1157-1182, 2003.
- [45] Douze, "Evaluation of GIST descriptors for web-scale image search," Santorini, 2009.
- [46] Ojala, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.