UNIVERSITÉ PARIS XIII - SORBONNE PARIS NORD
École Doctorale Sciences, Technologies, Santé Galilée

Modélisation mathématique des maladies inflammatoires chroniques de l'intestin

Mathematical modelling of chronic inflammatory bowel diseases

---

THÈSE DE DOCTORAT
présentée par

Safaa AL ALI
Laboratoire Analyse, Géométrie et Applications (LAGA)

pour l'obtention du grade de
D̈OCTEUR EN MATHÉMATIQUES APPLIQUÉES

soutenue le <date> devant le jury d'examen composé de :

NOM Prénom, établissement ..... Présidente/Président du jury
NOM Prénom, établissement ..... Rapporteur/Rapportrice
NOM Prénom, établissement ..... Rapportrice/Rapporteur
NOM Prénom, établissement ..... Examinateur/Examinatrice
NOM Prénom, établissement ..... Examinateur/Examinatrice
NOM Prénom, établissement ..... Directeur/Directrice de thèse
NOM Prénom, établissement ..... Co-directeur/Co-directrice de thèse

Abstract

bigskip The Thesis Abstract is written here (and usually kept to just this page)
...

Keywords – Ulcerative colitis, bleeding, ulcer, severity assessement

Résumé

Ici, le résumé de la thèse en français …

Mots clé – Rectocolite hémorragique, saignement, ulcère, sévérité

Acknowledgements

Dedicated to/To my…

Contents

List of Figures

List of Tables

Lepton : List of listings

List of Abbreviations

AUROC	**A**rea **U**nder the **R**eceiver **O**perating **C**urve
CCV	**C**olor **C**oherence **V**ector
CD	**C**rohn **D**isease
CMYK	**C**yan **M**agenta **Y**ellow **B**lack
GIT	**G**astro**i**ntestinal **T**ract
HSI	**H**ue **S**aturation **I**ntensity
HSV	**H**ue **S**aturation **V**alue
IBDs	**I**nflammatory **B**owel **D**iseases
IC	**I**ndeterminate **C**olitis
KNN	**K**-**N**earest **N**eighbors
RBF	**R**adial **B**asis **F**unction
RGB	**R**ed **G**reen **B**lue
ROC	**R**eceiver **O**perating **C**urve
ROI	**R**egion **o**f **I**nterest
SVM	**S**upport **V**ector **M**achines
UC	**U**lcerative **C**olitis
UCEIS	**U**lcerative **C**olitis **E**ndoscopic **I**ndex **S**core
WCE	**W**ireless **C**apsule **E**ndoscopy
YCbCr	**Y**(Luma) **C**hrominance-**b**lue **C**hrominance-**r**ed
YIQ	**Y**(Luma) **I**n-phase **Q**uadrature

Chapter 1

General introduction

1.1 Inflammatory bowel diseases (IBDs)

Inflammation is an essential physiological process geared towards protecting the body from invading pathogens or tissue trauma. It becomes by triggering a strong immune response essential for the elimination of harmful stimuli. In order to avoid ongoing inflammation, this pro-inflammatory response must be reduced/scaled down to restore tissue homeostasis once the injury or the pathogen has been eliminated. Failures occuring during the resolution of the inflammation have been shown to be involved in the the pathophysiology of chronic inflammatory diseases such as asthma, rheumatoid arthritis (RA), or inflammatory bowel diseases (IBDs) [CGGCA21].

Inflammatory bowel diseases are chronic inflammatory illnesses that touch/affect the gastro-intestinal tract (GIT), in which the lining of the bowel wall becomes inflamed. The architecture of the epithelial membrane is thus destroyed, leading to bowel malfunction [JSW13]. There has been a worldwide increase in the incidence of IBD in recent decades [Mak+20]. In fact, IBDs display a strong increase in prevalence in Europe and North America in the second half of the twentieth century and are becoming widespread in the rest of the world [PJ16]. It is estimated that over 2 million citizens suffer from IBDs North America, while in Europe this number rises to 3.2 million (250 000 cases in France [1]) and many millions worldwide [Dee20]. In a recent publication [Bia+19], it is shown that IBDs patients are exposed to an increased risk of overall cancer. About 25% of IBD patients are diagnosed at the age of 20 years old [Ye+15].

IBDs significantly impact the life of patients who have to face many challenges related to the disease and its complications, including the effect of the disease

[1] AFA Website

on their psychological well-being and the ability to meet their individual needs for function and productivity [MJ+17]. They strongly develop stress, anxiety or depression and they need lifetime follow-up and treatments.

IBDs encompass two main classes/forms/entities: Crohn disease (CD) and Ulcerative colitis (UC). Crohn's disease was first described in the Journal of the American Medical Association in 1932 by Drs. B. Crohn, L. Ginsberg, and G.D. Oppenheimer as terminal ileitis (involving the ileum). Involvement of other segments of the digestive tract lesions was then made and the Crohn's Disease name attributed to this disease. As for ulcerative colitis, Dr. Wilks and Moxon were the first to present it as a non-infectious pathology [Kir01]. The incidence of CD and UC diseases varies considerably both within and between geographic regions. Differentiating Crohn's disease from ulcerative colitis disease is a challenging issue especially at the begining of the disease evolution/course. However, it is an important step because the appropriate treatments and potential complications vary for these two conditions. Indeed, the distinction between CD and UC is based on combination of clinical, radiological, endoscopic and pathological exams. In the table 1.1, there are some key indicators to differentiate between CD disease and UC disease. However, for around 10% to 15 % of IBD patients, common symptoms to CD and UC diseases may be present, which makes their diagnosis quite difficult. Those patients are classified as indeterminate colitis (IC) according to the Crohn'as and Colitis foundation [2].

Feature	**CD**	**UC**
Abdominal pain	common	variable
Depth of inflammation	transmural	mucosal
Diarrhea	less severe	severe
Distribution	Segmental, discontinuous spread ; less common rectal involvement; occurs in entire gastrointestinal tract	Diffuse, contiguous spread; always involves rectum; spares proximal gastrointestinal tract
Fistula and sinus tracts	common	rare

Table 1.1: Comparison between Crohn disease and Ulcerative Colitis disease [CLT04; KS04; Lan+07]

In the figure 1.1, the spatial evolution of lesions for each type of disease is indicated in dark red color. The Crohn's disease can affect any part of the digestive tract from mouth to anus in a noncontinous extention (for example, in figure 1.1,

[2] Crohn's and Colitis foundation website

lesions of CD are spreading in the colon and small intestine). Conversely, UC is confined to the colon and characterized by a continuous inflammation of the superficial colonic mucosa moving continuously proximal from the rectum.

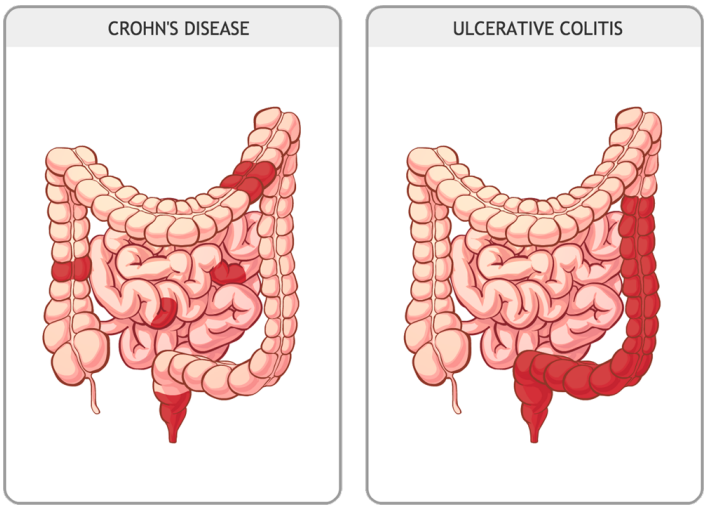


Figure 1.1: Crohn's and Ulcerative Colitis diseases extention in the gastrointestinal tract

1.2 Ulcerative Colitis disease (UC)

Ulcerative Colitis disease (UC) is a chronic/lifelong disease with a relapse remitting course that affects the lining of the large intestine (also called the colon) by inflammations and ulcers. The rectum is always affected by the disease which can extend proximally along a variable length of the colon. Depending on the colon parts involved in the inflammation, common forms of UC are found during a medical diagnosis procedure (cf. figure 7.1).

In the last decades, UC disease shows increased prevalence around the world (cf figure 2.3). Indeed, as recently posted by GlobalData, the estimated diagnosed prevalent cases of ulcerative colitis (UC) in eight countries (US, France, Germany, Italy, Spain, UK, Japan and Canada) will increase from 1.7 million cases in 2019 to 2 million cases in 2029, with an annual growth rate (AGR) of 1.36%.

To categorize the disease, doctors often use the video resulting from the colonoscopy exam in order to identify the colon parts affected by the disease as in figure 7.1:

(a) **Ulcerative proctitis** for the case of inflammation restricted/limited to rectum, and rectal bleeding may be the only sign of the disease

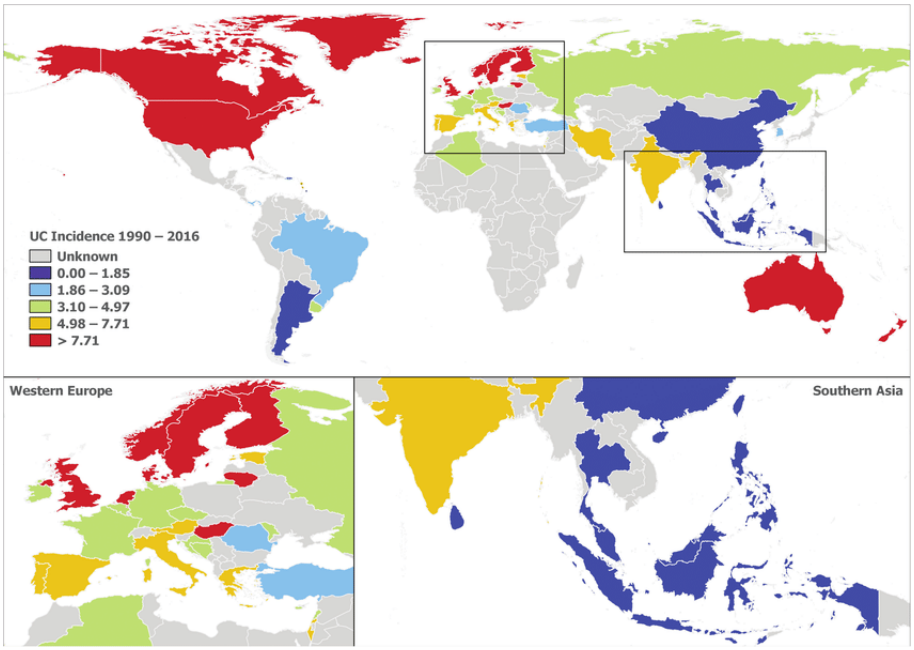


Figure 1.2: Worldwide map of the regional growth rates of UC disease. Countries in dark red represent the highest incidence of UC, while the gray ones indicate absence of data (Source University of Calgary)

(b) **Proctosigmoiditis** when disease/inflammation involves the rectum and sigmoid colon which is the lower end part of the colon. Signs and symptoms include bloody diarrhea, abdominal cramps and pain, and an inability to move the bowels in spite of the urge to do so (tenesmus).

(c) **Left-sided colitis** also recognized as distal colitis in the case of inflammation extending continuously from the rectum to the descending colon. Signs and symptoms include bloody diarrhea, abdominal cramping and pain on the left side, and urgency to defecate.

(d) **Extensive ulcerative colitis** is the case of a disease damaging rectum through up the transverse colon

(e) **Pancolitis** when the entire colon is damaged and causes bouts of bloody diarrhea that may be severe, abdominal cramps and pain, fatigue, and significant weight loss.

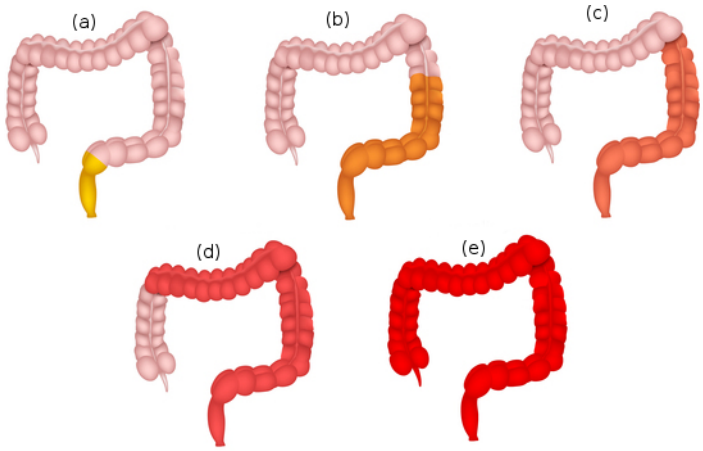


Figure 1.3: Main types of Ulcerative Colitis

During the disease evolution, the proximal extent of inflammation progresses, and after 20 years about 50% of patients have cumulative pancolitis (all the colon is affected/inflamed) [Cos+11].

Patients suffering fom UC usually exhibit multiple hallmark symptoms such as fever, abdominal cramping, intermittent bloody diarrhea, rectal bleeding, weight loss and general disconfort. Extra-intestinal manifestations as arthritis and skin lesions may be present. There is no cure, but the drug treatments currently available are designed to control pain and reduce the frequency and the duration of relapses, and thus relieve the symptoms in order to maintain the remission once the disease is under control. For the case of patients not responding to proposed treatments, surgery therapies such as total proctocolectomy (removal of the colon) and ileal pouch-anal anastomosis (IPAA) are introduced/used [TT08].

The exact causes of UC are still unknown. However, there are some factors that can be identified : dysregulated immune system response, heredity, smoking, environemental factors, alimentation..

The diagnosis of UC is based on some clinical and biological exams/findings:

- **Stool samples** to detect some markers of inflammation such as the calprotectin and C-reactive protein (CRP).
- **Blood tests** to screen anemia and/or inflammations.
- **Colonoscopy** to investigate the colon mucosa state.
- **Mucosal biopsy** from the colon inner lining tissues.

- **Medical imaging** like computed tomography (CT) scan, abdominal x ray, magnetic resonance imaging (MRI), radiography.

The location and spatial extent of the disease are significant indicators for the prognosis due to the association between the extent of lesions and the occurrence of complications such as colon cancer [Bál+18].

1.3 Mathematical modelling and medicine

In general, a model is a simplified representation of a phenomenon or process in which we are interested. In physics, for example, modeling is used to study the dynamics of fluids (liquids, gases and plasmas) and the associated internal forces. Models make it possible to describe the evolution of the state of the object under study by a set of laws, at a given level of approximation. In biology, we find the use of modeling to study the dynamics of a population or the spread of an epidemic.

In 2017, the US Food and Drug Administration (FDA) released a document affirming the importance of using mathematical models in the effective development and implementation of a robust pharmaceutical process [CMN17]. In the field of medicine, mathematical modeling is based on the use of data of various types collected upstream. The aim is to answer key questions such as analyzing data to develop a drug, treat a patient, predict disease progression over time, and monitor the effectiveness of the proposed treatment.

For ulcerative colitis disease, we find the use of deterministic mathematical models such as ordinary differential equations (ODE) and partial differential equations (PDE). This type of model makes it possible to understand the cellular interactions, complex in most cases, involved in the inflammation phase and thus makes it possible to explore different (medical) conditions without going through clinical experimentation, for example to compare several treatment strategies. Wendlsdorf and Bassaganya-Riera in 2010 [Wen+10], proposed a huge system formed of 29 ordinary differential equations to model the interactions between T-type immune cells that play an important role in the pathophysiology of chronic inflammatory bowel disease [PJ16]. Later in 2013, Lo and Arsenescu [LAF13a] proposed a simplified model of cellular interactions (initial phase of inflammation) to understand, depending on the type of inhibition of T cells in regulatory T cells, whether it is an inflammation or a bacterial infection. Most of this mathematical modeling work is concerned with the microscopic level, that is, the interactions between cells. At the macroscopic level (UC lesions), we rather find the intervention of the analysis of colonoscopy images via the use of binary classification methods such as k-means, Vector Support Machines (SVM), neural networks . The aim is to detect the pixels corresponding

to diseased or abnormal areas.

1.4 Thesis contributions

The work of this thesis concerns mathematical modeling at the macroscopic level. We worked on identifying mathematical models compatible with the observations collected in colonoscopy videos, describing their common properties for use in assessing the patient's condition and the spatial course of the disease. An interaction between image analysis techniques and PDE modeling was necessary to complete the work presented as well as a collaboration with the gastroenterology department of the Bichat-Beaujon hospital to obtain the medical data.

My research project proposes a set of tools for the analysis of colonoscopy videos and medical decision support: detection and quantification of lesions, representation of their spatial distribution, evaluation of the severity of the disease and its speed of propagation.

Chapter 2

Ulcerative Colitis disease

Abstract

This chapter resumes the ulcerative colitis disease medical aspects. First of all, we present the biological caracteristics of the disease. Then we show the prevalence world-wide map of the disease. Although the exact origin of the disease is unkown, some environemental and biological parameters can be a part that will be discussed. After, we will exhibit the different medical exams used to made the diagnosis of the disease. The end of the chapter is devoted to the therapies actual used in medical practice in order to control the symptoms of the disease, bring long-time remission and ameliorate the patient's condition of life.

Ulcerative Colitis disease (UC) is a lifelong disease with a relapse remitting course that affects the lining of the large intestine (also called the colon) by inflammations and ulcers. The rectum is always affected by the disease which can extend proximally along a variable length of the colon. In general, UC affects the innermost lining of the colon and the rectum, and develops inflammation like bleeding and sores (ulcers) as shown in Figure 2.1. The first report that described the disease dates to 1850 [DF11].

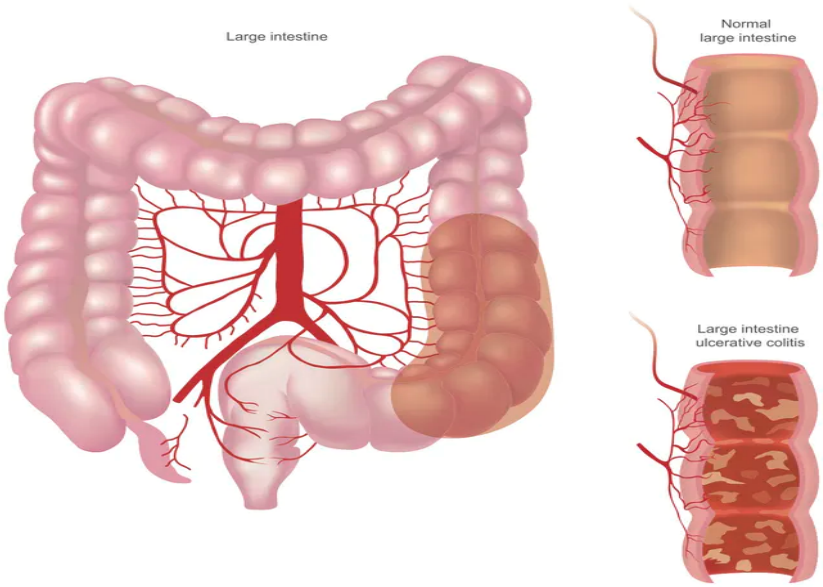


Figure 2.1: Normal colon and UC irritated colon (Source Getty Images)

2.1 Biology

Ulcerative colitis is considered an idiopathic and auto-immune disorder that affect the colonic mucosa.

namely caused by a derulated action of the immune system againt itself. It be sometimes detected by symptoms and complications outside the colon. The main role of the immune system is to protect the body from external bodies such as germs, viruses, bacteria, and other dangerous substances.

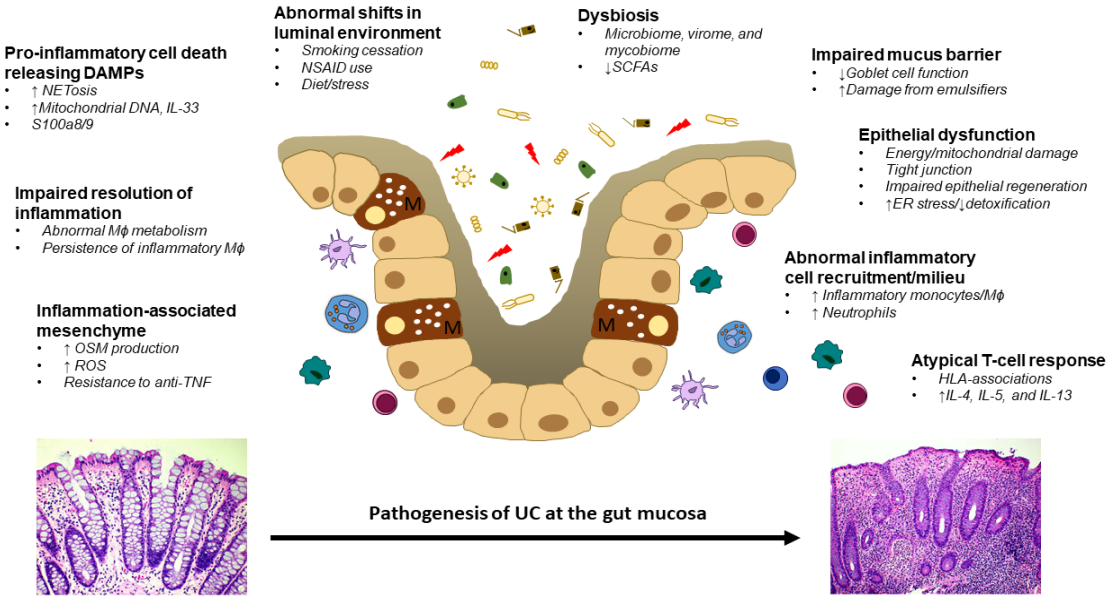


Figure 2.2: Molecular mechanisms during inflammation caused by UC [PKH20]. The mucosal and submucosal layers are damaged because of inflammation. The inflammation is restricted to the surface mucosa unlike the case of CD.

2.2 Epidemiology

The UC disease can develop at any age even it is almost diagnosed for age ranging from 15 to 30. It is likely to be more common among european white people, especially those from Ashkenazi Jewish communities, and blacks. In the last decades, UC disease shows increased prevalence around the world (cf Figure 2.3). The UC almost affects an equal number of women and men. It is shown that the industrialised country like North America, Western Europe, Southern Asia and Australia have the highest rates of UC growth. Indeed, as recently posted by GlobalData, the estimated diagnosed prevalent cases of ulcerative colitis (UC) in eight countries (US, France, Germany, Italy, Spain, UK, Japan and Canada) will increase from 1.7 million cases in 2019 to 2 million cases in 2029, with an annual growth rate (AGR) of 1.36%.

2.3 Risk factors and symptoms

2.3.1 Risk factors/Possible causes

Despite the fact that the exact origins of the disease is not kown, combination of some genetic and environmental factors may increase the risk of having the disease in the life. These factors include:

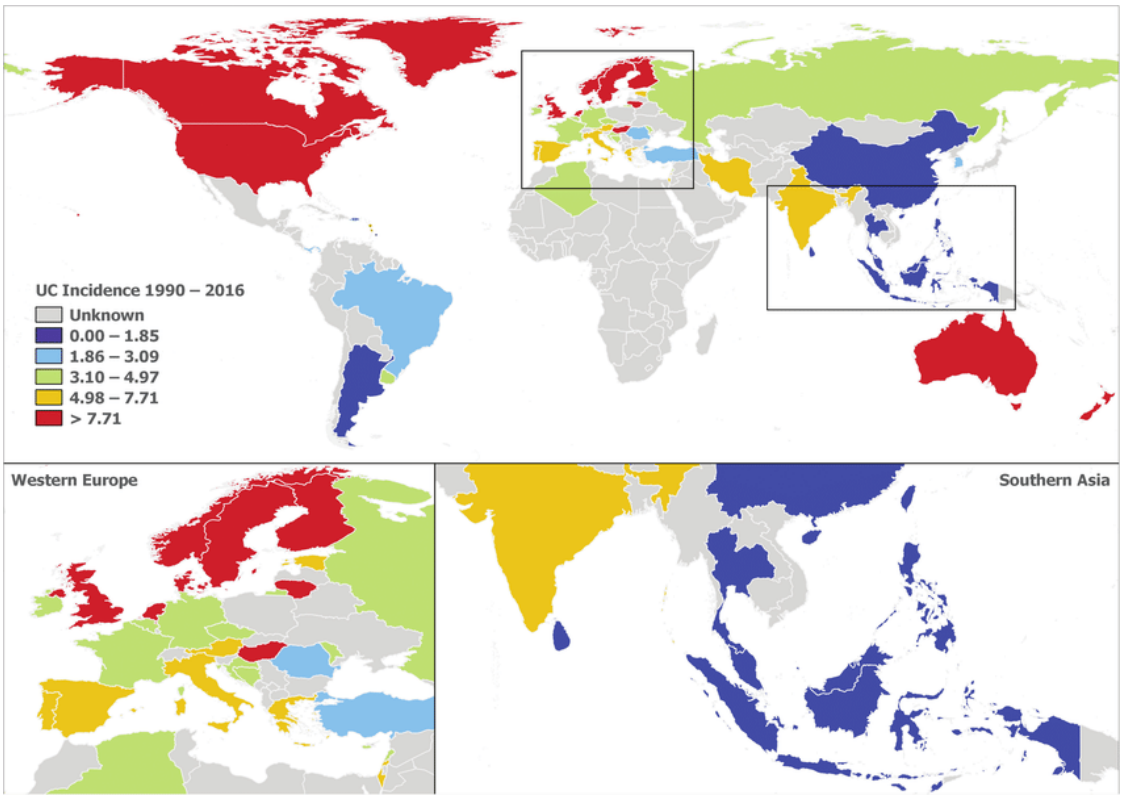


Figure 2.3: Worldwide map of the regional growth rates of UC disease. Countries in dark red represent the highest incidence of UC, while the gray ones indicate absence of data (Source University of Calgary)

- **Age** there are peaks of UC diagnosed cases at the ages 15 to 30 and then again for the range 50 to 70
- **Ethnicity** Ashkenazi Jews have a rate of ulcerative colitis that is three to fi ve times higher than that of other ethnic groups [Ord+12]
- **Family history**: reports have shown that up to 20% of UC patients have a family member of relative experienced with UC or IBD. The risk of UC is particularly high in fist-degree relatives: 5·7–15·5% of patients with ulcerative colitis have a fi rst-degree relative with the same disease [Mon+87; FMM80]
- **Genetic**
- **Environmental factors**
- **Alimentation and Lifestyle**
- **Alternative suppositions**

2.3.2 Symptoms and signs

Inflammation caused by UC the disease tends to occur multiple times over the course of a person's life, causing recurring signs and symptoms with remission times. In gact, these symptoms usually evolve over time rather than suddenly. Patients suffering fom UC usually exhibit multiple hallmark symptoms such as:

- Fever
- Irregular heartbeats
- Abdominal pain and cramping
- Intermittent bloody diarrhea
- Rectal bleeding
- Loss of appetite, weight loss
- Urgency to defecate or controversy inability to defecate despite the necessity
- General disconfort and fatigue
- Failure to grow up in the case of children

The most UC patients experience come and go phases of active symptoms. However, if the disease have affected the rectum and the colon, then the patient may often suffer from severe reccurring symptoms

2.3.3 Complications

The UC may generate some complications for the colon such as perforation or abnormal dilation (megacolon). In addition, extra-intestinal manifestations as arthritis, osteoporosis (bone loss), skin and eyes inflammations may be present.

2.4 Diagnosis

The diagnosis of UC is based on some clinical and biological exams/findings:

- **Stool samples** to detect some markers of inflammation such as the calprotectin and C-reactive protein (CRP). The calprotectin is a small protein, mainly secreted by polymorphonuclear cells in the lumen of the digestive tract and is excreted in the stool. It is an objective indicator of inflammatory activity systematically integrated in therapeutic trials. It plays crucial role in defferent situations such as:
  - Diagnosis assistance to differentiate between functionnal or irritable bowel syndrome (IBS) pain and those of origin IBDs. An increased dosage value of calprotectin can be sign of various diseases such as

IBDs, infectious and parasitic enteritis, colorectal cancer, obverdose of the nonsteroidal anti-inflammatory drugs

- Tracking/Following up IBDs Measure of the response for IBDs treatments
- Postoperative

The fecal calprotectin is recognized as a reliable and non-invasive marker of the inflammation of the digestive tract. It is an important indicator in the case of recidive CD disease i.e. remission phase and active UC.

The CRP is generally secreted by the liver after a stimulation with interleukin 6 (IL-6). In CD, the CRP is correlated with the clinical activity of the disease while in the case of UC, the increase growth in CRP value indicates a severe form. However, it is less specific than the fecal calprotectin to indicate inflammation of the intestinal tract [CMH15].

- **Blood tests** to screen anemia as seen in Figure 2.4 and/or inflammations.

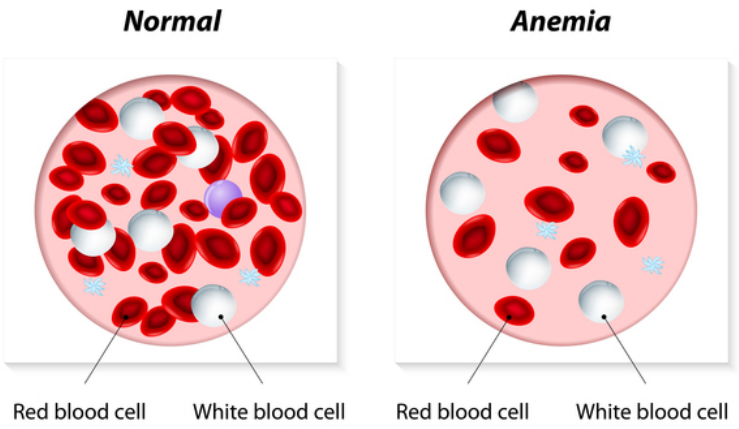
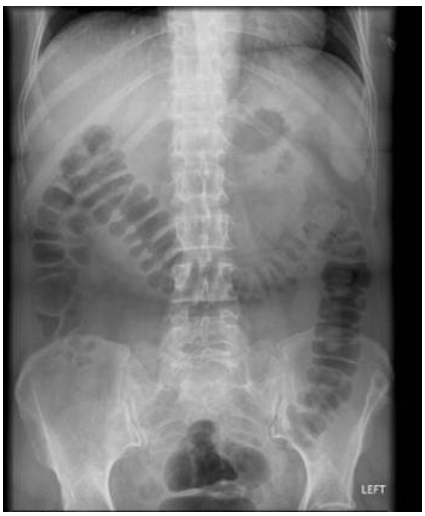


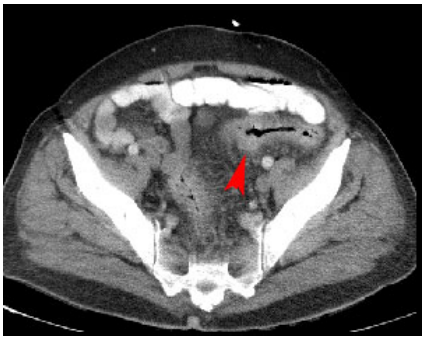
Figure 2.4: Anemia signs shown during a blood test (Credit Designua/Shutterstock.com)

- **Mucosal biopsy** from the colon inner lining tissues.
- **Medical imaging** there are many type of images can be recommended during the UC diagnosis procedure like those presented in Figure 2.5:
  - Computed tomography (CT) scan which produces digital images of the interior of the colon and documents the size and location of found abnormalities
  - Abdominal x ray that permits to the physician a view about the contours of the bowel,
  - Magnetic resonance imaging (MRI),

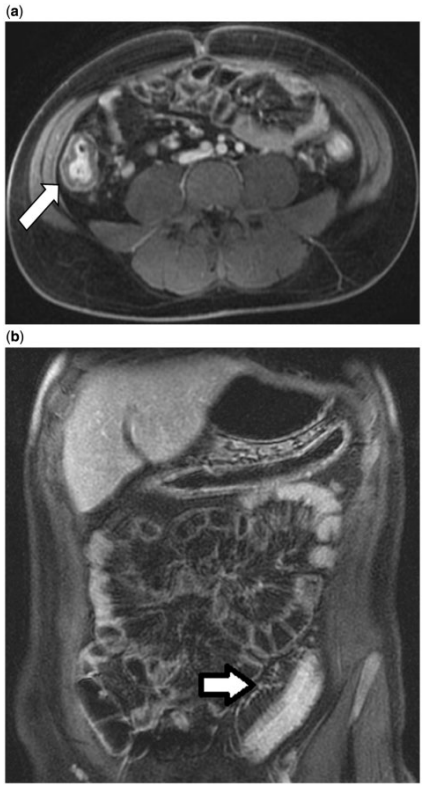
– Radiography.



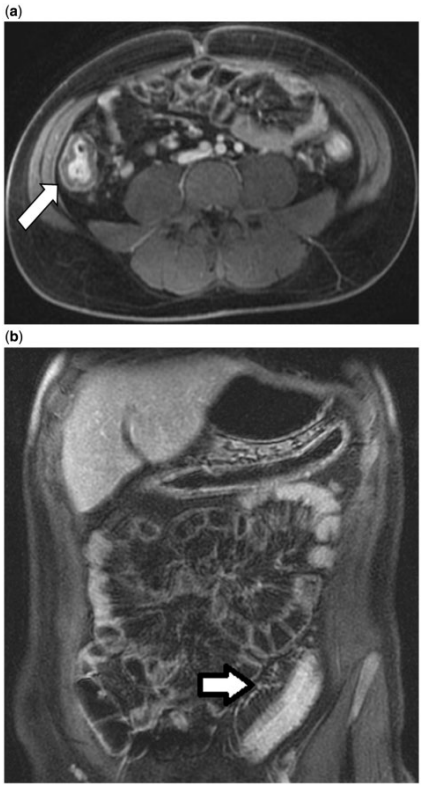
(a) CT scan showing inflammation and bowel wall thickening in ulcerative colitis



(b) Abdominal X-ray



(c) MRI

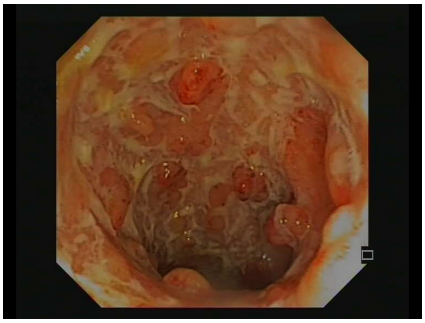


(d) Radiography

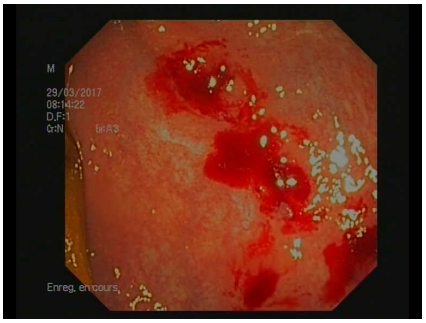
Figure 2.5: Examples of medical imaging tests used for the diagnosis of the UC

- **Colonoscopy** to investigate the colon mucosa state. It consists of the introduction of a medical instrument called endoscope for which a camera is

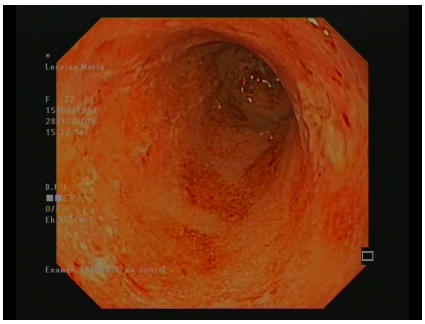
attached to visualise the intestinal inner wall. The video countains images showing signs of inflammations caracterized by lesions such as bleeding and ulcers as presented in Figure 2.6. The chapter 3 illustrates the complete procedure, the steps to do before and after.



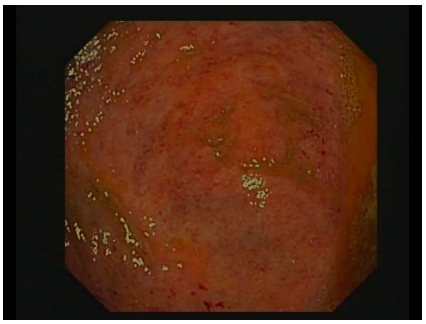
(a) Ulceration



(b) Bleeding



(c) Erosion



(d) Fibrosis

Figure 2.6: Endoscopic abnormalities that are signs of UC disease found during a colonoscopy video

The location and spatial extent of the disease are significant indicators for the prognosis due to the association between the extent of lesions and the occurrence of complications such as colon cancer [Bál+18].

2.5 Classification

Depending on the colon parts involved in the inflammation, common forms of UC are found during a medical diagnosis procedure (cf. Figure 7.1). To categorize the disease, doctors often use the video resulting from the colonoscopy exam in order to identify the colon parts affected by the disease as in Figure 7.1:

(a) **Ulcerative proctitis** for the case of inflammation restricted/limited to rectum, and rectal bleeding may be the only sign of the disease

(b) **Proctosigmoiditis** when disease/inflammation involves the rectum and sigmoid colon which is the lower end part of the colon. Signs and symptoms include bloody diarrhea, abdominal cramps and pain, and an inability to move the bowels in spite of the urge to do so (tenesmus).

(c) **Left-sided colitis** also recognized as distal colitis in the case of inflammation extending continuously from the rectum to the descending colon. Signs and symptoms include bloody diarrhea, abdominal cramping and pain on the left side, and urgency to defecate.

(d) **Extensive ulcerative colitis** is the case of a disease damaging rectum through up the transverse colon

(e) **Pancolitis** when the entire colon is damaged and causes bouts of bloody diarrhea that may be severe, abdominal cramps and pain, fatigue, and significant weight loss.

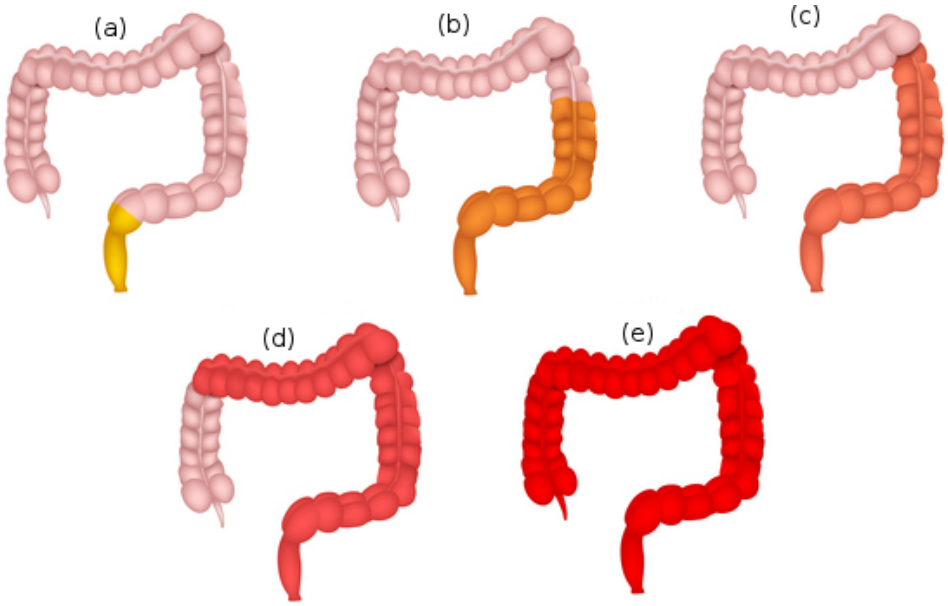


Figure 2.7: Main types of Ulcerative Colitis

During the disease evolution, the proximal extent of inflammation progresses, and after 20 years about 50% of patients have cumulative pancolitis (all the colon is affected/inflamed) [Cos+11].

2.6 ?? Endoscopic Scores ??

2.7 Management of UC

Patients living with UC disease go through a series of alternate phases of relapses when the disease is active (and so on the symptoms) and phases of remission. During the remission stage, the patient almost little to no symptoms. Thus the actual proposed therapies try to remain the remission phase the longest possible time, i.e. for years if possible and offer to patient the possibility going back to the daily activities.

The treatment of the UC depends of its severity i.e amount of inflammation and extent in addition to the patient's nature, then it is not unique.

2.7.1 Changes in diet and lifestyle

maintaining a high calorie diet or a lactose-free diet can improve symptoms.

2.7.2 Symptomatic drug therapy

2.7.3 Anti-inflammatory therapy

There is no cure to the UC disease, but the drug treatments currently available are designed to control pain and reduce the frequency and the duration of relapses, and thus relieve the symptoms in order to maintain the remission once the disease is under control in order to maintain an optimal quality of life for each patient. For the case of patients not responding to proposed treatments, surgery therapies such as total proctocolectomy (removal of the colon) and ileal pouch-anal anastomosis (IPAA) are introduced/used [TT08].

- **Aminosalicylated derivatives**
- **Corticosteroids**
- **Immunosuppressants**
- **Antidiarrhoeals**
- **Antispasmodics**
- **Biotherapies, anti-TNFs**
- **Biosimilars**
- **Integrin** $\alpha 4 - \beta 7$

2.7.4 Surgery

For some UC patients, the different methods of disease management fails, then a surgery as removing the irritated parts of the large intestine becomes a necessity. There are about 30% of UC patients undergo a surgery during their life. The surgery is especially prescribed to some cases like:

- Avoidance of the side effects of drugs
- Control or stop of the symptoms that are difficult to manage
- Prevention from the colon carcinoma (CRC)
- Elimination of bleeding

A colectomy is a surgery operation to manage and treat the large intestine conditions such as inflammatory bowel diseases, diverticulitis, colorectal cancer, bowel obstruction and uncontrolled bleeding. The principle is to remove the affected parts of the colon, which will may be totally damaged by diseases in some cases. We distinguish a variety types of colectomy (cf Figure 2.8):

- **Total colectomy** when the entire colon is removed.
- **Partial colectomy** (also known as subtotal colectomy) in the case of removing of the majority parts of the colon. The sigmoid colon is only left and it is connected to the small intestine.
- **Hemicolectomy** which depends on the part of colon that will be removed is known as:
  - **Left hemicolectomy** in the case of amputation of the left colon segment
  - **Right hemicolectomy** in the case of amputation of the right colon segment
  - **Transverse hemicolectomy** in the case of amputation of the transverse colon segment
- **Sigmoid colectomy** in the case of resection of the sigmoid colon
- **Proctocolectomy** when the colon and the rectum are totally removed.

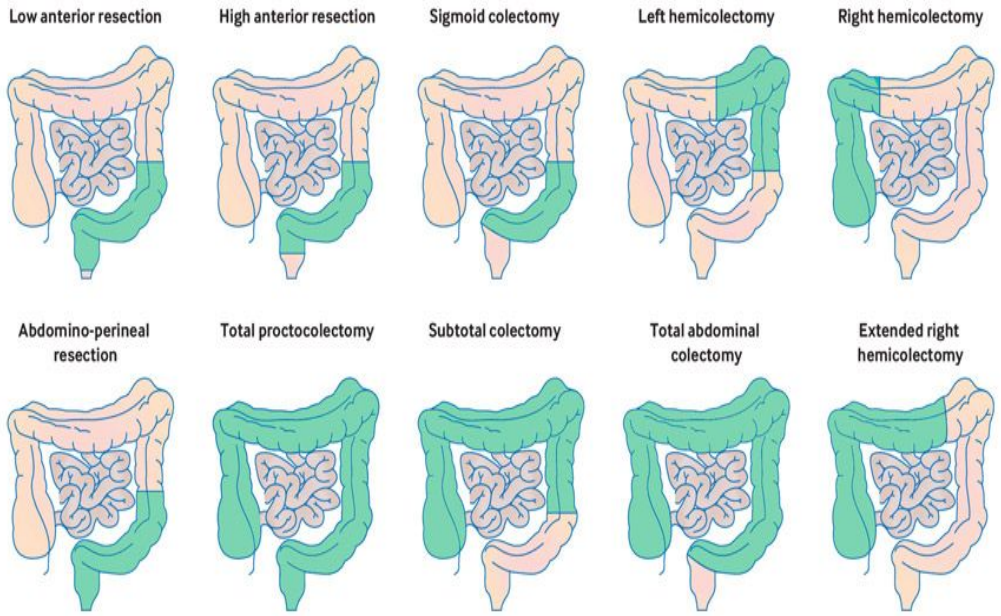


Figure 2.8: Types of colectomy surgery operations. Removed parts are indicated in green color whereas the remained colon parts are in light pink color (Source https://drdavidford.com/procedures/colectomy-colon-removal-by-dr-david-w-ford)

It has been shown that around 10% to 15% of patients diagnosed with UC undergo/follow a surgery operation within the first 10 years of disease progression.

Consequently, to proceed with a colectomy operation, the doctor needs to know the affected parts of the colon to be able to decide the necessity or not and the type of the of special medications and surgery intervention as well.

However, surgery intervention can lead to some complications. *During surgery, the small intestine will either be diverted from an opening in the abdomen (an ileostomy) or used to create an internal pocket connected to the anus called an ileoanal pocket.*

2.8 Prognosis

2.9 Conclusion

Chapter 3

Exploration of the digestive tract: wireless capsule endoscopy and colonoscopy

Abstract

The diagnosis of inflammatory bowel diseases entails a set of clinical, biological and endoscopic measurements. Inflammation does not manifests similarly for all patients, therefore blood tests and biomarkers such as CRP and faecal calprotectin cannot reveal the inflammation's state of the intestinal mucosa. Consequently, for a precise grading of the disease severity, the screening of the gastrointestinal tract is necessary. The doctors usually use two main techniques for IBD: wireless capsule endoscopy (WCE) or colonoscopy.

Due to its ability to screen the upper part of the gastrointestinal tract, the wireless capsule endoscopy is more suitable for the diagnosis of Crohn's disease while colonoscopy is convenient for ulcerative colitis disease.

The obtained videos contain images presenting the lesions that characterize the digestive tract inflammatory diseases like bleeding, ulcers and the vascular pattern ~~appearance/state~~. These images will be later used to assess the disease severity through the endoscopic scores ~~that~~ will be presented later in the next chapter.

We presented in chapter 1 the inflammatory bowel diseases which are characterized by the appearance of lesions such as bleeding and ulcerations at the lining wall of the large intestine. Those lesions can affect any part of the digestive tract from the mouth down to the end of the rectum (or the anus) (cf Figure 3.1).

Consequently, the doctor needs to visualize the inner of the digestive system to be able to detect abnormalities caused by the inflammation generated by IBDs to make the best diagnosis. To do so, "endoscopy" turns out to be a convenient procedure to visualize the interior of the human body, especially the digestive tract. Endoscopy is a medical operation consisting of the insertion of an endoscope into the openings parts of the body like the mouth, anus, or any possible small incisions in the knee or abdomen for example. The word "endoscope" is made of two Greek roots: "endo" which means "inside", and "skopos", which means "to target or watch". In other words, endoscope means looking inside the body.

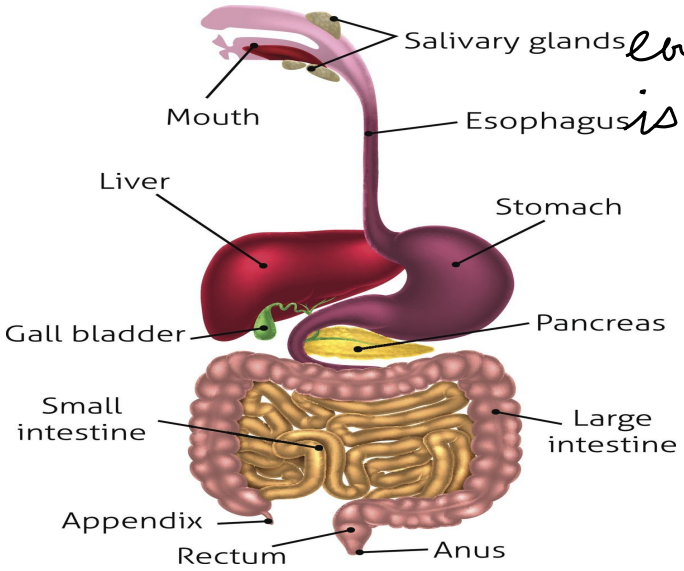


Figure 3.1: Diagram of the digestive tract

In general, the doctor prescribes an endoscopy examination to a patient presenting some unusual symptoms such as:

- swallowing trouble,
- persisting stomach pain,
- anaemia,
- diarrhoea,
- nausea and vomiting,
- chronic pain,

- unexplained loss of weight,
- abdominal pain,
- blood in the stool,
- black or tarry stools

There are several types of endoscopy examinations depending on the body part or organ the doctors want to explore to make his diagnosis. In Table 3.1, we recall some types of endoscopy examination and the organ to be explored. In general, the endoscopy examination is used not only to explore the interior of the human body. It is also to remove some tissue samples (biopsies) to be analyzed to confirm the diagnosis in cancer, UC and CD diseases or other conditions, or for treatment purposes like cauterizing a bleeding vessel, or removing the appendix (appendectomies) during laparoscopic surgery. It can be also used to remove abnormalities such as tumours or polyps found during the investigation of the digestive tract, also called colonoscopy. More applications can be found in references [Met96; Buc+18; Che+21; Dav+21]. We will give more details throughout the chapter's sections.

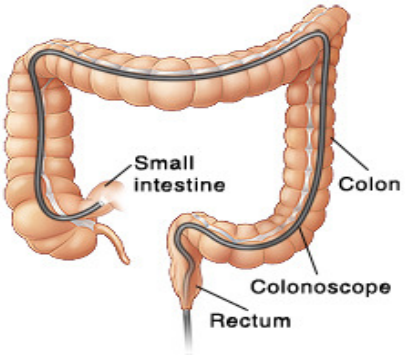
Area/tract	Organ	Type of endoscopy
Ear	Ear	Otoscopy
Respiratory	Nose	Rhinoscopy
	Lower Respiratory tract	bronchoscopy
Larynx	Larynx	Laryngoscopy
Gastrointestinal	Esophagus, stomach, and duodenum	Esophagogastroduodenoscopy
	Small intestine	Enteroscopy
	Colon	Colonoscopy
	Rectum and the lower part of the colon (i.e. the sigmoid colon)	Sigmoidoscopy
	Bile duct, rectum	Rectoscopy
	Anus	Anoscopy
Through small incisions	Abdominal or pelvic cavity	Laparoscopy
	Interior of a joint	Arthroscopy
	Organs of the chest	Thoracoscopy and mediastinoscopy

Table 3.1: Some types of endoscopy

In this thesis, we are interested in the analysis of endoscopy videos obtained while exploring the gastrointestinal tract for the diagnosis of ulcerative colitis. In other words, we are interested in the colonoscopy examination (or the sigmoidoscopy while visualizing the lower part of the colon and the rectum). From a medical point of view, we can split up the gastrointestinal digestive tract into four main different parts which are the oesophagus, the stomach, the small intestine and the colon (or the large intestine), [Bas+12].

To visualize the state of the intestinal wall, the doctors use an embedded camera which can be either sitting inside a pill-size capsule, called Wireless Capsule Endoscopy (cf Figure 3.2a) or attached to a long, thin flexible hose, called colonoscope (cf Figure 3.2b).

(a) Wireless Capsule Endoscopy device



(b) Colonoscope

Figure 3.2: Medical devices used during the exploration of the digestive tract.

In the next sections, we will present in detail the two procedures, the tools that are used during the examination of the gastrointestinal tract and some side effects of each one.

3.1 Wireless Capsule Endoscopy

The Wireless Capsule Endoscopy (WCE, [Idd+00]), also called Capsule endoscopy or capsule enteroscopy, is proposed for the diagnosis of the small intestine including the duodenum, the jejunum and the ileum (cf Figure 3.3), because these parts can not be reached by traditional gastrointestinal tract endoscopy as the upper endoscopy or by the colonoscopy (details about these procedures are listed in Table 3.1).

3.1.1 Medical device

The first WCE tool was developed by the Given Imaging and got the approval of the Food and Drug Administration (FDA) in 2001. In general, WCE is a tiny medical device of approximately 26×11 millimetre size with the outer casing isoplast. It is made of three components/parts which are presented in Figure 3.5:

1. a capsule endoscope (large pill size instrument whose architecture is made of 8 elements as shown in Figure 3.4)
2. a system of sensors or electronic circuits with adhesive sleeves and belt to be attached to the abdomen of the patient, a battery and the data recorder device
3. a personal computer workstation

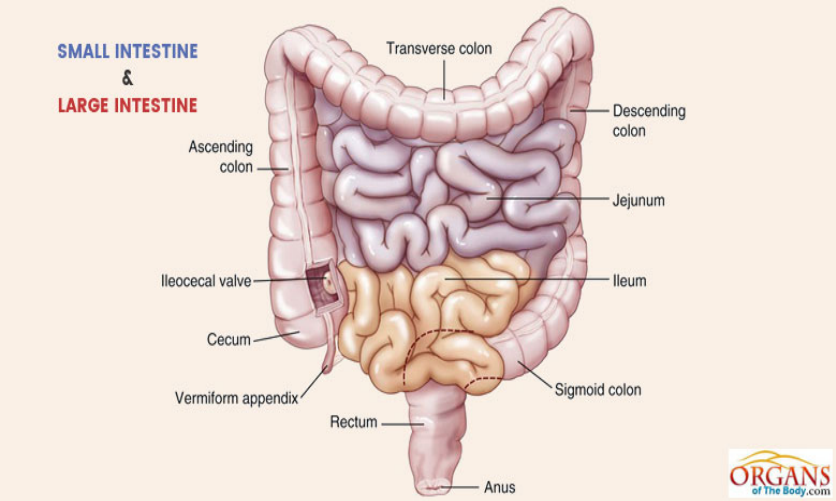


Figure 3.3: Anatomy of the small and large intestine (cf https://www.organsofthebody.com/small-intestine/)

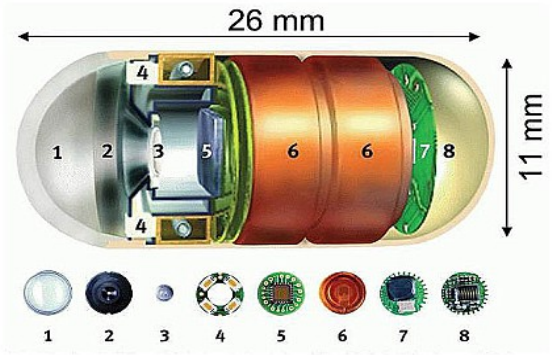


Figure 3.4: The architecture of the WCE [OEI18]

In Figure 3.4 we enumerate the different eight parts composing a standard capsule endoscopy device, which are according to reference [OEI18]:

1. Optical Dome used for liquid filtration and gastrointestinal enzymes balancing,
2. Lens Holder which tightly holds the lens,
3. Lens,
4. light-emitting diode (LED) responsible for the light illumination around the passage area of the body for easier identification of the affected tissues,
5. Complementary Metal Oxide Semiconductor (CMOS) Sensor able to detect tiny objects with 140-degree accuracy and very high-quality images,
6. Battery made up of silver oxid, made to be unharmful to the body and it is used to feed the CMOS detector, LED and transmitter. And it can work

for about 8 hours,

7. Application Specific Integrated Circuit (ASIC Transmitter),
8. Antenna composed of coated polyethene, it ensures the information/images communication between the belt receiver and the capsule.

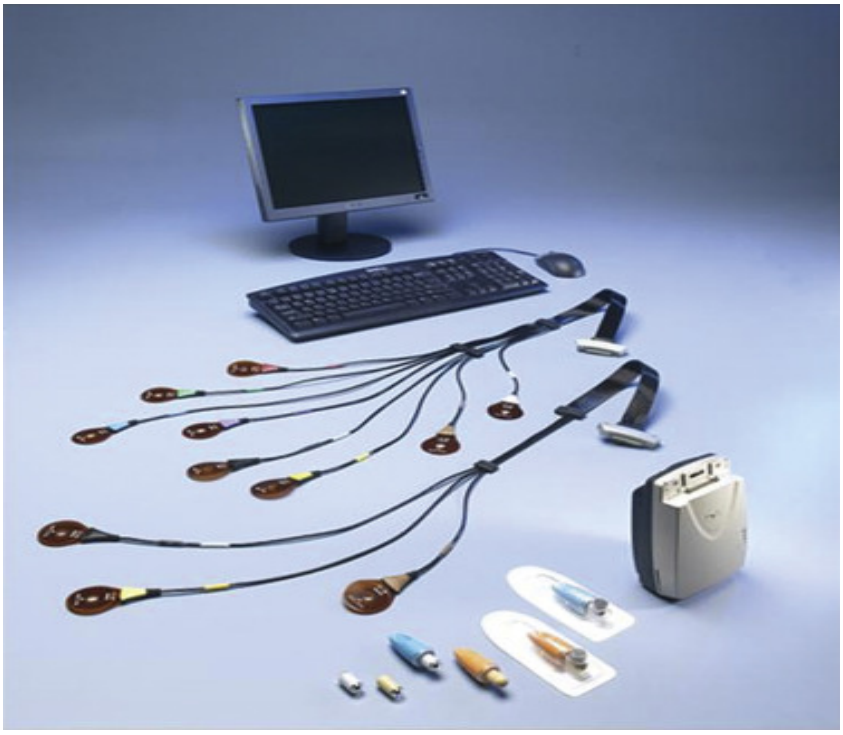


Figure 3.5: The complete system for WCE device/tool. It consists of the capsule device, the sensing system with sensing pads and the data recorder, the battery pack, and the personnal computer workstation.

Actually, there are many companies that have developped approved WCE devices (see Figure 3.6):

- **EndoCapsule**; Olympus America, Inc, Center Valley, Pennsylvania
- **PillCam**; Given Imaging, Ltd, Yoqneam, Israel
  - PillCam SB (as SB2,SB2EX,SB3) is almost recommended for the detection of obscure gastrointestinal bleeding [1]
  - PillCam ESO is used for the investigation of the oesophagal diseases
  - PillCam COLON is adapted for the case of the colonic neoplasias inspection
- **MiroCam**; Intromedic Co Ltd, Seoul, Korea

[1]Obscure Gastrointestinal Bleeding (OBG) is a syndrome grouping together patients who have bleeding from the small intestine since gastroscopy and colonoscopy are normal

- **OMOM**, Chongqing Jinshan Science and Technology Group

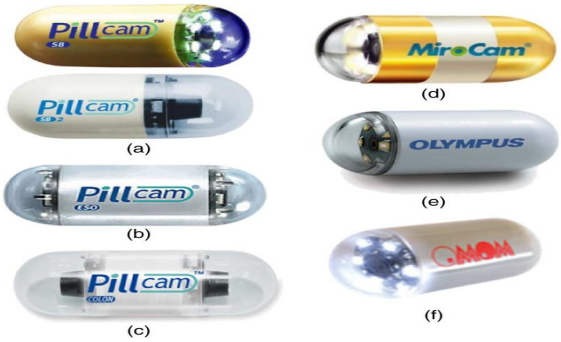


Figure 3.6: Available clinical WCE devices [CMD11]. (a) PillCamSB and PillCam SB2, (b) PillCam ESO, (c) PillCam COLON, (d) MiroCam , (e) EndoCapsule, and (f) OMOM

Figure 3.7 shows the different types of WCE which differ by many parameters such as size, weight, image resolution and price. They have almost approximately the same rate of frames per second, fps (in the middle of 2 to 3 fps).

Capsule Endoscopy

Devices used to perform endoscopy operations

				
Capsule	PillCam® SB 3 Given Imaging	EndoCapsule® Olympus America	MiroCam® IntroMedic Company	OMOM® Jinshan Science and Technology
Size	Length: 26.2 mm Diameter: 11.4 mm	Length: 26 mm Diameter: 11mm	Length: 24.5 mm Diameter: 10.8 mm	Length: 27.9 mm Diameter: 13 mm
Weight	3.00g	3.50g	3.25-4.70g	6.00g
Battery life	8 hours or longer	8 hours or longer	11 hours or longer	6-8 hours or longer
Resolution	340x340	512x512	320x320	640x480
Frames per second	2 fps or 2-6 fps	2 fps	3 fps	2 fps
Field of view	156°	145 °	170°	140°
Communication	Radio frequency communication	Radio frequency communication	Human body communication	Radio frequency communication
FDA approval	Yes	Yes	Yes	No
Price per capsule	$500	$500	$500	$250

Figure 3.7: Features characterizing some available WCE devices

In general, the computer workstation of the WCE system is equipped with convenient software for the review and the interpretation process of the obtained images during the realization of the colonoscopy. For example, the Given Imaging company uses the system RAPID v 6.5, while Olympus America uses WS-1 EndoCapsule.

In the following sections, we present the procedure of WCE examination to explore the digestive tract.

3.1.2 Before procedure

The exploration of the digestive tract can be viewed as driving a car to explore new places. If the weather is cloudy, the visualisation will be troubled, the driver will not be able to distinguish right directions and hence it is very difficult to drive and see the country. However, if the weather is sunny, the driver will be able to drive easily and see all its neighbours. The car corresponds to the capsule endoscopy device. The first case, i.e. cloudy road driving, is like an unclean digestive tract (containing residual). The second case corresponds to the cleared digestive tract.

As result, to get a better visualisation of the gastrointestinal tract, the doctor must see it clean as much as possible, otherwise, the diagnosis will not be accurate due to the obscuration of some parts of the intestinal wall caused by persisting residuals. Therefore, the patient who will undergo a WCE examination will be asked to fast (food and drink) at least 12 hours before the procedure. Moreover, if the doctor sees that the fasting procedure is not sufficient, he can propose some laxatives to clean up the digestive tract. The patient should inform his doctor if he has some swallowing disorder, previous abdominal surgery or history of bowel obstructions, or chest pain. Moreover, nausea and vomiting disorders must be declared. He should indicate the presence of a pacemaker or defibrillator as well.

3.1.3 During the procedure

On the day of the procedure, the patient will have attached adhesive patches to his abdomen. Each patch contains an antenna with wires that connect to a recorder (see Figure 3.5). He wears the recorder which will be placed on the belt put around his waist. The recorder is firstly connected, then the patient swallows the capsule endoscopy (cf Figure 3.4) with a cup of water. The camera takes pictures during his passage through the digestive tract at the rate of two to four images by second, depending on the type of WCE devices (cf Figure 3.7). The sensor device transmits the obtained pictures via a radio frequency (RF) system [Eri02] to the data recorder. The capsule will be rejected naturally through the digestive tract for time-varying between eight to seventy-two hours approximately.

3.1.4 After the procedure

Once the capsule is ejected from the patient body, he must follow the doctor instructions as the removal of attached patches and the recorder from his body and packaging them to return all the equipment. After two hours of the capsule

rejection, the patient can take clear liquids (such as water). And, for another two hours later, he can take light lunch. These instructions are usually discussed with the doctor before the WCE examination. Some unusual accidents can arrive during this type of medical technique. For example, if the capsule was not rejected within two weeks, the patient should contact his doctor who might order an X-ray image to detect the localisation of the capsule and then be able to remove it from his body.

During the course of the WCE in the body, the camera takes 50 -60 thousand images with a rate of 2 to 4 images per second. The data recorder has the ability to store of 5500-6000 JPEG images, with about 10 GigaBytes (GB) drive capacity [OEI18]. The obtained endoscopic images are saved on the recorder and transferred to the computer station which will connect the images together to create a video. This video will be reviewed by the expert doctor in order to detect digestive abnormalities and then make his treatment. The results are in general diffused to the patient within a week after the WCE examination.

Main advantages of WCE:

- easy and safe procedure
- non-invasive
- easy movement through the digestive tract
- detection of the oesophagus and small intestine conditions
- high-quality images of the gut interior
- short recovery time

Main disadvantages of WCE:

- non-reusable, very expensive (500 $)
- uncontrollable once swallowed by the patient
- precautions with pregnant patients
- precautions with patients having pacemakers or other electromedical tools
- inability to collect biopsies during the procedure
- battery may stop before the end of the procedure

The review and analysis procedure of a WCE video is a very time-consuming task that requires a lot of caution and experience. There are a large number of images to be analyzed to find abnormalities such as lesions in the case of CD disease or erosive esophagitis; angiodysplasia; suspicious nodular ulceration corresponding to the histology of adenocarcinoma in the small intestine and many other conditions of the upper parts of the digestive tract. The coloured

.

images are similar for some cases which make their detection and or localisation difficult. Therefore, many researchers have proposed methods to detect automatically the lesions that can be found during a WCE video. These works will be discussed in chapter 6.

The technique of WCE is suitable for the diagnosis of Crohn's disease which can affect all the parts of the digestive tract. However, this is not the case for the exploration of the colon parts to assess the severity of the UC disease. Indeed, during the movement/the transit of the swallowed capsule endoscopy throughout the digestive system, it is quite difficult to detect its location regarding the colon, the speed for which it moves is not constant and sometimes the loss of battery can affect it.

Consequently, for the diagnosis of UC disease affecting the colon, the gastroenterologists rather use the colonoscopy which allows visualizing the colon, and it is recognized as the method of reference for evaluating the disease severity, and hence making treatment decisions and assessing treatment response [Naw+14; Goe18]. Colonoscopy procedure [Pro+18] is adapted to check the entire colon and rectum (which are approximate 1200-1500 mm in length). As mentioned above and in the Table 3.1, if the diagnosis aims to only visualize the rectum and the lower colon parts, we call the procedure, a sigmoidoscopy. In the next section, we will discuss the procedure of the colonoscopy and the instructions to follow to obtain a successful examination of the colon.

3.2 Colonoscopy

Colonoscopy is a worthwhile non-surgical medical examination for screening abnormalities of the colon mucosa as bleeding, ulcerations, vascular pattern appearance in the case of inflammatory bowel diseases and detecting adenomas or polyps which may turn into cancer in the case of colorectal cancer. Polyps are small growth with a stalk protruding from the mucous membrane, ..??. Adenocarcinomas present 96% of colorectal cancer cases.

In general, to have an efficient colon examination using colonoscopy, we compel the combination of three elements depending on patient preparation, equipment structure and gastroenterologist/physician experience [ARG15]. The first element concerns the patient bowel preparation i.e. the cleansing of the colon, to eliminate all possible solid matter residue (stool) which can obscure the visibility of the state of the colonic lining during the examination and hence prevent precise diagnosis. So some preparations are recommended before the realization of the colonoscopy procedure. The second element is the medical equipment used to establish the examination as the endoscope (Figure 3.8), the camera resolution that impacts the quality of obtained images/videos, the medical tools

used for polyp removal and/or biopsies extraction. The third key element for a high-quality colonoscopy is the expertise and the competence of the doctor doing the colonoscopy.

3.2.1 Medical instrument: endoscope

The endoscope device used during a colonoscopy examination is called colonoscope. In general, it is a cylindrical long thin tube (cf Figure 3.8). It can be handled by a single individual: the gastroenterologist or the specialist realising the examination. It consists of a flexible tube and proximal housing connected by an umbilical cord to a light source [SP21]. The eyepiece, control bottom for angulation and suction are designed to be in the proximal housing usually handled by the left-right of the doctor while the right hand manages the advancement and withdrawal of the endoscope during the procedure. The endoscope used for examination of the colon in the case of children differs from the endoscope used for adults. The endoscope length varies between 160 and 180 cm, and its diameter between 1 to 1.2 cm [SP21].

As in the case of WCE, a variety of companies integrate the production of endoscope instruments and many challenging issues were to be resolved. Some companies include in the handle part of their scope a dial to increase or decrease its rigidity. This feature leads to the facilitation of controlling the instrument and its passage across the turning angles of the colon. These angles concern the corner connecting the ascending colon part to the transverse colon, the corner connecting the transverse colon to the descending one and the corner linking the latter to the sigmoid colon. Afterwards, we will see how it is difficult to localise these angles during a colonoscopy examination and hence correctly identify possible existing lesions in these areas. If the scope rigidity is high, this may lead to pain and an increase in the gut wall perforation rate. So the doctor handling the scope should be aware of this feature.

The control bottom (cf Figure 3.9) is also adapted to manipulate the images obtained during the colonoscopy as recording, taking screenshots, zooming. In some cases, the doctor fixes lesions and take a picture to analyze later. The upward and downward of the instrument are controlled by a large wheel placed rightly near the handle. While the small wheel distally located from the handle is used to move the scope in the right and/or left directions. The multiple angulation movements of the scope permit wider visualisation of the interior of the colon.

From Figure 3.8, the end of the colonoscope presents working channel ports and suction allowing the insertion of accessory instruments useful for better visualization of the colon and better therapeutic interventions. These instruments

include two or three lenses to obtain clear images, a two LED lights to illuminate the field of view of the mucosa through a fiberoptic light bundle. And, if the doctor needs to collect some tissue samples, he may use a special instrument as biopsy forceps (Figure 3.14), or to remove polyps, he rather uses snare forceps (Figure 3.15). Additive instruments include needles and clip appliers. These instruments can be used to stop bleeding in a vessel or cauterize bleeding tissue. Other channels are used for water/air insertion during the colonoscopy. This procedure also called irrigation means the insertion of water or another medication to clean the tested organ, which is where the colon is. This step is necessary in the case of present stool in the colon. This will be detailed in subsection 3.2.2.

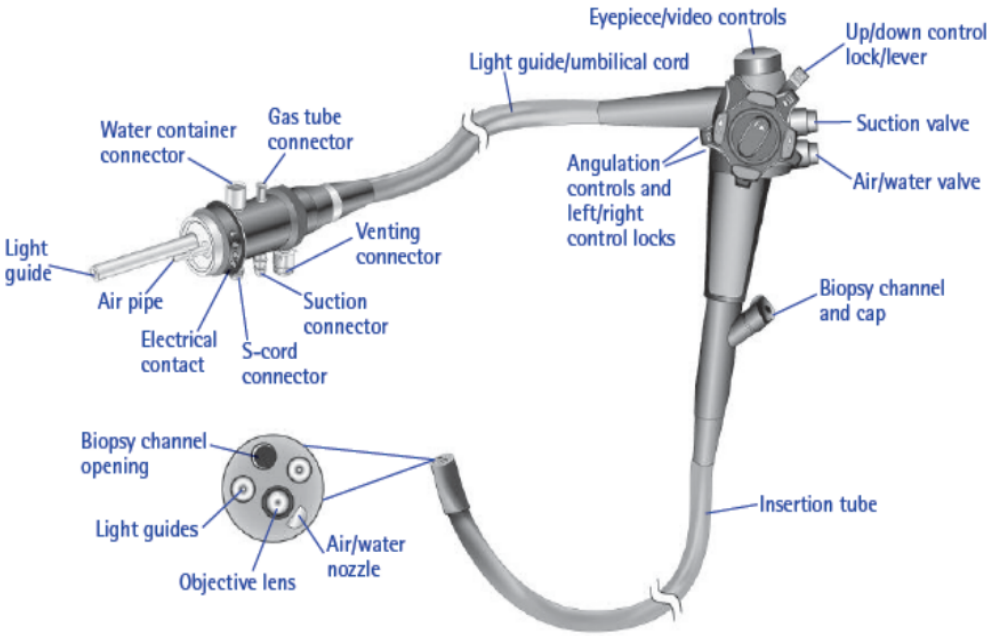


Figure 3.8: Flexible endoscope instrument used to explore the interior of the colon

3.2.1.1 Table of available endoscopes + compagnies ?

? Need of information in this thesis from doctors: type of endoscopes ..

In the following sections, we will exhibit the steps to follow before and after the procedure. Then we figure out the colonoscopy examination course, the tools that doctors may use to remove polyps or biopsies, the acquisition of images of the colon interior. We end this section by presenting some measures to do after the procedure to reduce some possible side effects that the patient may feel after the procedure.

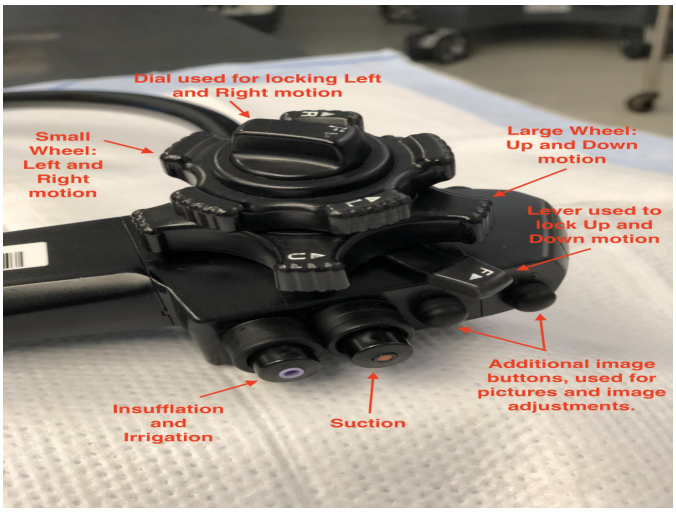


Figure 3.9: Zoom on the endoscope part that will be handled by the doctor and related to the external monitor machine (cf [SP21])

3.2.2 Procedure before colonoscopy

At least two weeks before the upcoming colonoscopy, the patient should inform his doctor about all treatments that he is currently taking. This involves diabetes, blood fluidizing, lung conditions heart problem medications and any treatment or vitamins containing iron. And, if he has any allergies cases for some medications, he should mention it. The doctor may adjust the dosages of these treatments or stop them temporarily before the colonoscopy examination regarding the patient state. For some patients having an artificial heart valve, the prescription of antibiotics may be required.

As we precise it, to perform a successful colonoscopy, we must clean out the colon from any possible solid residual because it masks the visibility of the lining of the colon. Hence, the doctor may indicate to the patient to have a special diet a day or many days before the examination. In most cases, the patient is required to follow a low fibre regime. This consists of stopping the consumption of solid food, gelatin (Jell-O) food container and milk products. This type of food is hard to be digested and more likely to cause constipation and so to get in the way of the camera during the passage of the endoscope along the colon.

In Figure 3.10, we show different images from colonoscopy videos. We remark that the first image (from the left) contains a lot of residual matter that obscures parts of the lining of the colon. The last image (at the right) demonstrates a very good cleansing of the colon because there is no residual substance which brings a clear view of the state of the mucosa, the pleats and the blood vessels. The patient must avoid even red or purple colour liquids such as prune and orange juice because these drinks can discolour the fluid at the colon mucosa level which will appear as blood and thus lead to incorrect bleeding lesions retrieval

for example. Also, the state of the vascular pattern will not be effectively evaluated. We will see later that the vascular pattern is an essential descriptor of the UC disease, and should be correctly evaluated in order to make correct disease severity assessment.

On the contrary of these steps to follow before doing a colonoscopy, the patient who will undergo a colonoscopy will be asked to increase the consumption of clear beverages including plain or carbonated water, tea or coffee without cream/milk, clear sports fluids, lemonade. If this diet was not sufficient enough to empty the colon, some laxatives can be proposed by the doctor in order to accelerate or ameliorate the elimination of possible residuals.

The colon preparation is evaluated using Boston Bowel Preparation Scale (BBPS) created by the Boston University in 2009 [Lai+09]. The BBPS scores the three principal segments of the colon each one from 0 to 3 using the Table 3.2. The three parts composing the colon are the right segment evolving the cecum and the ascending colon; the transverse segment, including the hepatic and the splenic flexures; and the left segment, counting for the descending colon, sigmoid and rectum (see Figure 3.3).

Description	Scoring
Unprepared colon segment with mucosa not seen due to solid stool that cannot be cleared	0
Portion of mucosa of the colon segment seen, other areas of the colon segment not well seen due to staining, residual stool and/or opaque liquid	1
Minor amount of residual staining, small fragments of stool and/or opaque liquid, mucosa of colon segment seen well	2
Entire mucosa of colon segment seen well with no residual staining, small fragments of stool or opaque liquid. The wording of the scale was finalized after incorporating feedback from three colleagues experienced in colonoscopy	3

Table 3.2: Computation of the BBPS index per segment [Lai+09]

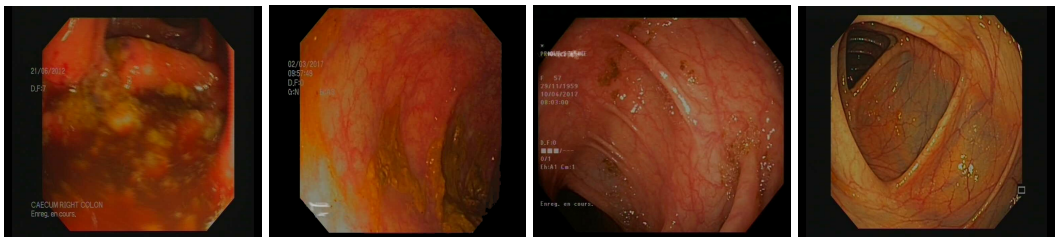


Figure 3.10: Examples of the use of BBPS score to evaluate bowel preparation on a given segment of the colon. From left to right, the classification of bowel preparation is respectively unsatisfactory or poor (BBPS=0), fair (BBPS=1), good (BBPS=2) and excellent bowel preparation (BBPS = 3)

It is suggested to repeat the colonoscopy examination if the total BBPS score is less than or equal to 3 or one of the three considered segments is rated less than 2, i.e. equal to 1. An ideal preparation corresponds to a score = 9. So if the colon isn't sufficiently cleaned, a new colonoscopy examination should be rescheduled, and new preparation indications to adapt which is very complicated to the patient at the health level and financial level. In fact, most of the preparation medications used before colonoscopy are not taken by the health security, and the colonoscopy examination itself is very expensive (Put $..)

3.2.3 During the colonoscopy

In general, the colonoscopy examination is almost a painless procedure because it is realised under patient sedation. Indeed to make the examination more comfortable and with less to no pain, some pain reliever and an intravenous sedative (i.e. introduced by the vein) will be given to the patient just a few times before the colonoscopy. This step is called sedation, and it is generally performed by an anaesthetist or a nurse anaesthetist before any medical operation, or surgical intervention depending on the doctor's recommendations.

There are different types of sedations, complete or general sedation and conscious sedation. The first one refers to general anaesthesia when using propofol medication so the patient will feel nothing during the examination because he will be sleeping through the procedure without remembering any medical activity upon waking up. After this type of sedation, the patient should not make an effort or drive because he will feel some somnolence and disturbance of his cognitive function for about four to twenty-four hours afterwards. So, he should have an accompanying person to drive him home back after the colonoscopy [Son16]. On the other hand, during the conscious sedation the patient stays awake, he can even respond to the doctor ..

The colonoscopy procedure will begin with the patient lying flat on a padded examination table. He will be positioned on the left side while his knees are

drawn up towards his chest to get a better angle to the colon. The gastroenterologist inserts an instrument called the endoscope (or simply saying scope) gently into the rectum. The endoscope as shown in Figure 3.8, is a long flexible thin tube of millimetre thickness. It is composed of a light, video camera, channels for air and water, and special attachments instruments to remove any possible polyps that appear during the exploration of the colon wall. In addition, the endoscope permits to doctor to collect some tissue samples, also called biopsies, to be analyzed at the laboratory and then have a precise diagnosis.

For thorough colon examination, the colon should be expanded to let the endoscope move easily. Hence, some air (carbon dioxide) is inflated to unfold tissues which may cause some abdominal cramps and uncomfortable bloated sensation. These disagreeable sensations can be reduced by taking slow and deep breaths. Another technique to facilitate the movement of the endoscope into the colon is to use "underwater" examination, i.e. by introducing water in the colon during the passage of the endoscope (see Figure 3.11).

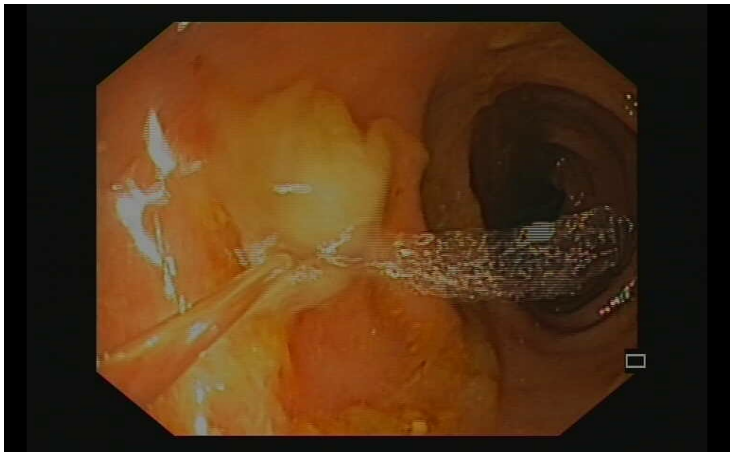
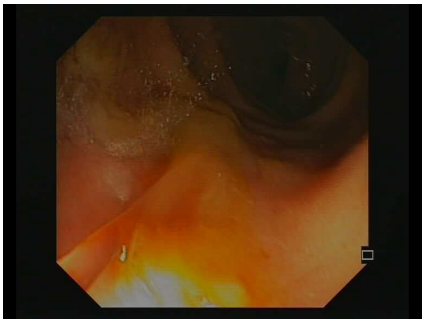


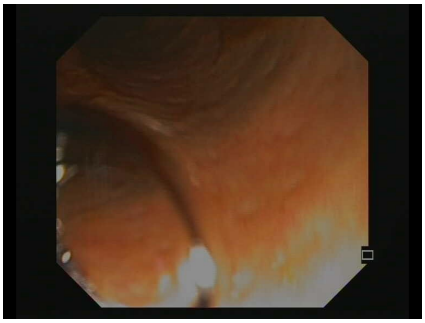
Figure 3.11: Introducing water by the endoscope during the colonoscopy

This technique can replace the air inflation or combine it with it, it depends on the visibility option during the colonoscopy. The decision is made by the gastroenterologist. The use of water to expand the colon during the colonoscopy leads to fewer pain patients feeling/experience and facilitates the endoscope passing. However, this technique implies an abrupt in the movements of the endoscope, due to forced movements caused by the introduced water and this induces artefacts. In fact, the rapid movement of the endoscope generates blurry images (see for instance images of Figure 3.12). As it can be shown, it can produce several problems of visibility of the colonic mucosa in the images. These artefacts will affect the evaluation of the UC activity because some parts of the mucosa even are blurry or completely obscured. On the other hand, these problems will induce/produce problems during the construction (learning phase) of

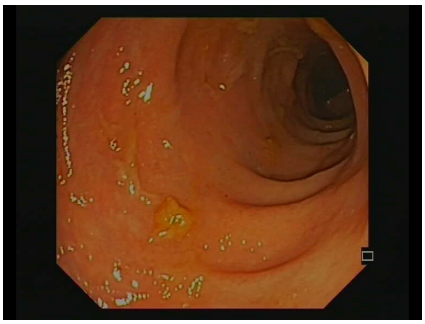
our proposed automatic detection algorithm in chapter 6 and its performance evaluation even.



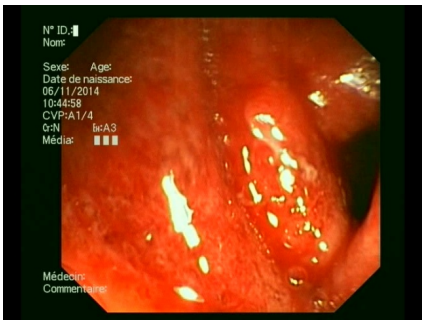
(a) Some parts of the colon wall are obliterated



(b) The scope line is visualized



(c) Some white spots on the mucosa



(d) Large spots totally hide the mucosa

Figure 3.12: Illustration of various images containing artifacts resulting of the use of water during the colonoscopy

The endoscope is then slowly advanced under direct vision and is moved around the various bends of the colon. Throughout the endoscope movements, the doctor may apply some pressure to let the endoscope move ahead of the entire colon. And sometimes, the patient will be asked to change position to have a better examination process in the case of conscious sedation.

Once, the physician realising the examination have reached the relation point between the colon and the small intestine, which is the cecum (see Figure 3.3), then he begins the slow removal/ withdrawal of the endoscope and the acquisition of the images begins and at the same time he examines carefully the lining of the colon.

The colonoscope has a tiny camera on the end of it, which is connected to a monitor as presented in Figure 3.13. This permits the transmission in real-time of the images taken by the camera of the endoscope to the screen of the monitor. This ensures that the doctor visualizes the colon wall as much as the camera goes

throughout the colon. The length of the colon is about five to six feet, which demand about 12 minutes to be passed by the enfig:InstrumentSnaredoscope and additive 12 minutes to remove it. If biopsies removal and/or polys extraction is performed, the time of colonoscopy procedure varies between 30 minutes to one hour to be completely terminated. The set of images is shown as a video on the monitor and saved like a colonoscopy video.

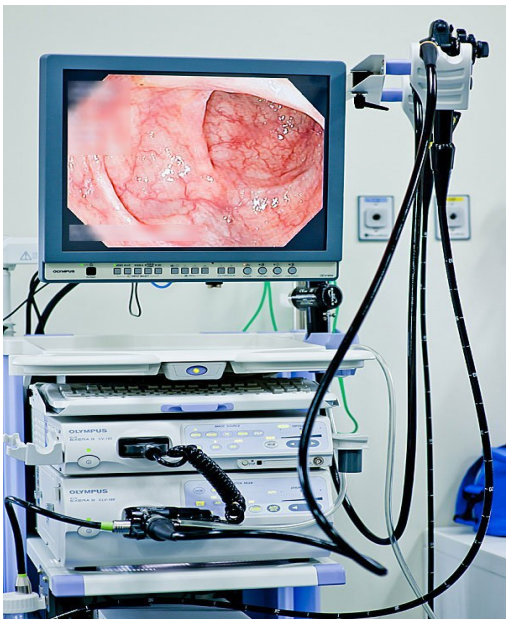


Figure 3.13: Monitor machine to which the endoscope in Figure 3.8 is related in order to trasmit the images of the colon

During the colonoscopy, if it is necessary, the doctor uses some medical equipment to remove malignant polyps due to colorectal cancer (CRC), or biopsies or tissue samples from the intestinal mucosa in order to confirm the diagnosis. To remove polyps, he uses for example biopsy forceps given in Figure 3.14. The tissue samples are removed and then sent to a laboratory to be analyzed. Another instrument is the snare forceps mostly used to polypectomies, i.e. to remove polyps(cf Figure 3.15). Another instrument is the brush to remove tissue sample mucosa...??? These instruments are of variable size depending on what the doctor sees in the colonoscopy as the size of the polyp for example. For more details about these instruments, we refer the reader to papers [Ura+14; Pig+15; Chi+18]. Biopsy samples should be taken from normal and abnormal tissue and compared to histological findings.

3.2.4 After the colonoscopy

When the colonoscopy is finished, the patient is invited to wait in a recovery room until he completely wakes up if he submitted to general anaesthesia or until the pain medications wore off if he had conscious anaesthesia. So the time

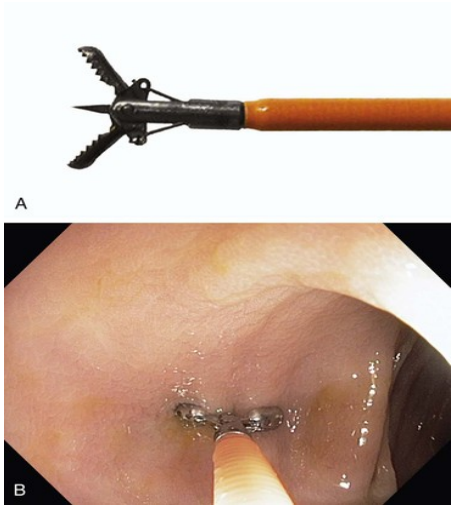


Figure 3.14: A. Biopsy forceps, B. Application of the forceps to collect biopsies during the colonoscopy

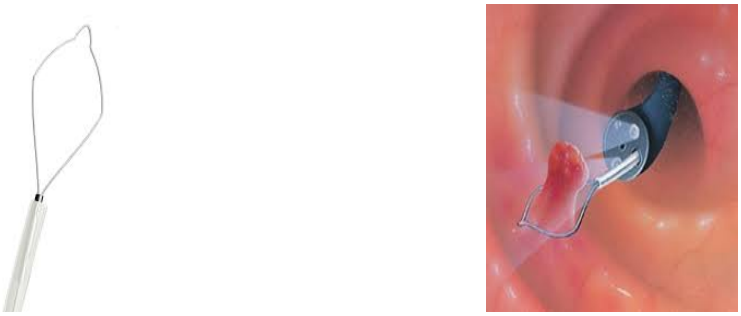


Figure 3.15: Snare forceps (left) and it is application during the colonoscopy (right)

that patient needs to recover depends on the sedation medications. In general, directly after the colonoscopy, the doctor writes a report about the examination, the biopsies and or polyp removal if was done, and a global analysis is addressed to the patient. Depending on the situation, the patient will be sent to a specific doctor for example for whom will be addressed the results of affected biopsies and or tissue samples or he will be asked to check a specific doctor to have a convenient treatment in the case of cancer detection for example.

In the case of the diagnosis of UC disease, the obtained colonoscopy video will be generally reviewed by the physician realising the examination to provide a numerical evaluation of the disease activity. The third key element for a successful colonoscopy, meaning the expertise of the physician is of great importance for this stage. The physician needs to control the speed of the endoscope to have clear images and withdrawal of the whole colon. Also, he needs to know the use of medical instruments such as forceps and snares.

Some problems with colonoscopy can be dangerous as infections due to the non-cleanliness of the endoscope and the used instruments used during the procedure. Thus, the physician realising the examination should be aware of this type of problem. In addition, some reports show perforation of the colon due to the unsuitable or incorrect movement of the endoscope. Consequently, this leads to an additive health problem for the patient.

In the following section, we will expand the different mucosal abnormalities that doctors may find in the endoscopy videos.

3.3 Endoscopic lesions

Endoscopy allows direct visual inspection of the intestinal mucosa and, therefore, helps assess the severity of inflammation, elucidate its location relative to the digestive tract, and estimate other possibilities of disease during a flare-up that can mimic an IBD.

Despite differences in the affected mucosal layers, Crohn's and ulcerative colitis diseases share many overlapping lesions visualized during an endoscopic exploration of the digestive tract. In Table 3.3, we summarise the initial mucosal appearance that characterizes UC and CD. The intestinal lesions range from bleeding, ulceration, erythema, mucosa's granularity to the damage of the vascular pattern visibility.

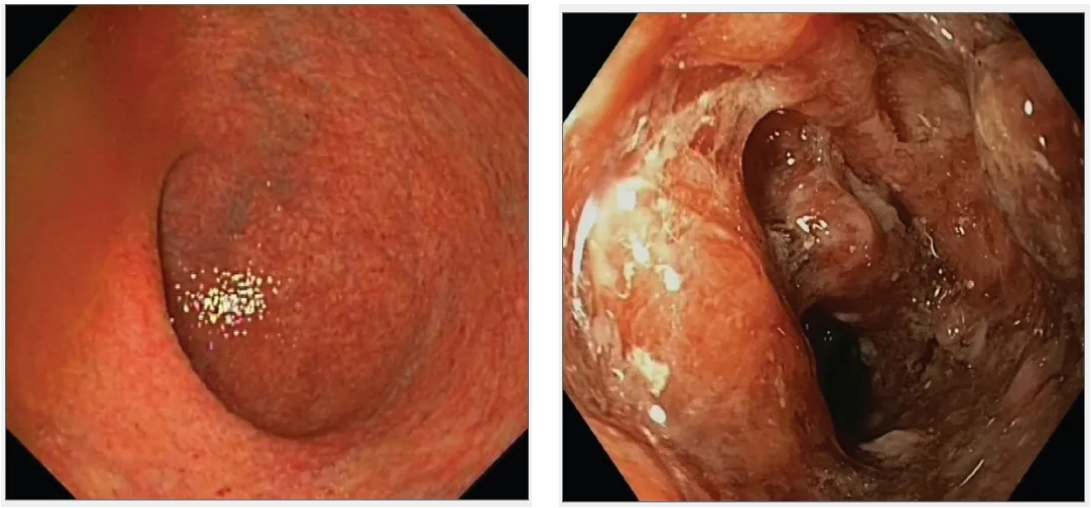
Features	UC	CD
Bleeding	Always	Almost always
Erythema	Always	Almost always
Loss of vascular pattern	Always	Rare
Granularity of the mucosa	Always	Rare
Cobblestone appearance	No	Yes
Pseudo polyps	Always	Always
Aphtous ulcers	Rare	Always
Deep ulcers	No	Always
Patchy inflammation	Infrequent	Always
Ileal ulcers	No	Always
Rectal involvement	Almost always	Frequent
Fistula	No	Frequent
Perianal lesions	No	Frequent
Stricture	Rare	Frequent

Table 3.3: Aspect of mucosal lesions found during an endoscopy for UC and CD (cf [RA12; LL16]).

In the case of CD (Figure 3.18), the ulcers lesions tend to be linear and deeper in comparison to UC's ulcers. They are usually associated with the cobblestone appearance of the ileum and strictures in the colon or ileum. Stenoses as well as

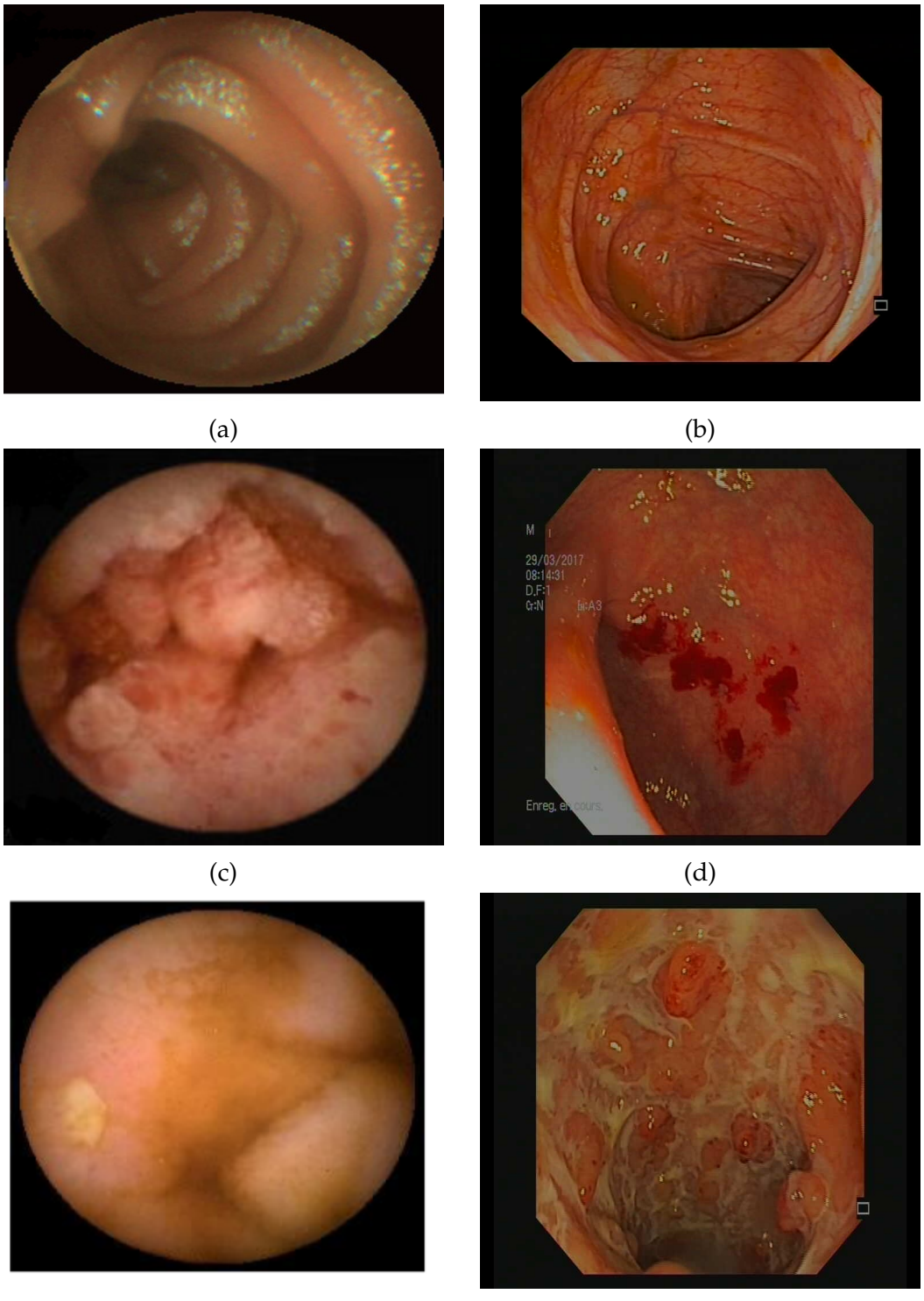
open internal or perianal fistulas suggest CD disorder and are rarely found for UC patients [RA12]. Pseudopolyps and adenomatous polys can be visualized for the two illnesses. The most commonly recognized lesions for CD and UC are bleeding and ulcers (cf Figure 3.16).

In the case of UC (Figure 3.17), the mucosa is granular and brittle. When the disease becomes severe, erosions lesions corresponding to superficial mucosal damage appear. Due to oedema, the visibility of the blood vessels can be degraded.



(a) (b)

Figure 3.17: Some endoscopic lesions in the case of UC disease. (a) Granular colonic mucosa with a faded vascular pattern, mild friability, and erythema. (b) Large ulcerations and spontaneous bleeding (copyright Photos)

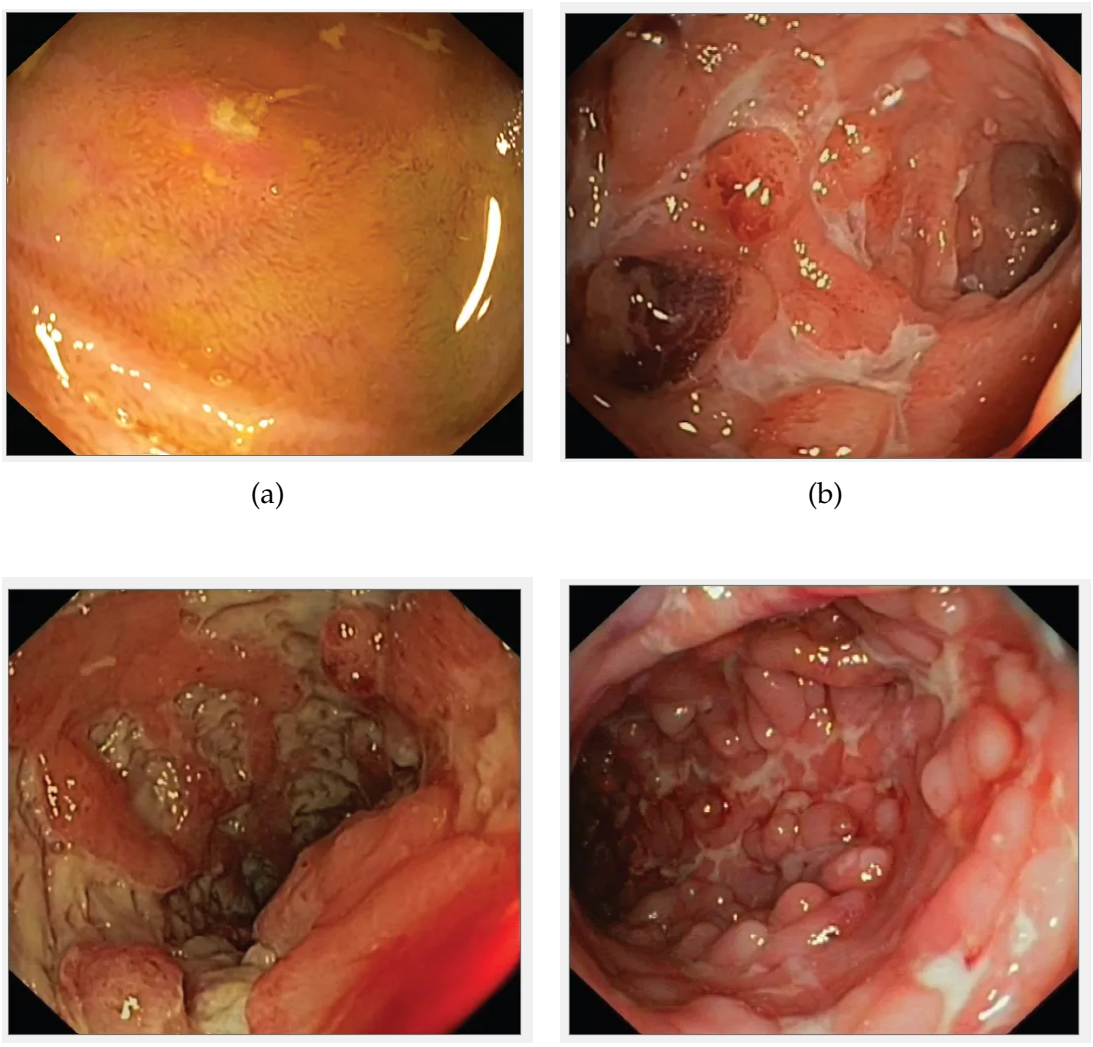


(a) (b)

(c) (d)

(e) (f)

Figure 3.16: Endoscopic findings in WCE and colonoscopy for CD (left column) and UC (right column) respectively. First row: healthy mucosa. Principal common lesions: Second row: bleeding lesions. Third row: ulceration lesions



(a) (b)

(c) (d)

Figure 3.18: Illustration of some sensitive endoscopic lesions that suggest Crohn's disease. (a) Aphthous ulceration. (b) Deep ulceration associated with hemorragic mucosa. (c) Deep ulceration associated with granular mucosa. (d) Cobblestone-like mucosa corresponding to mucosal pattern with raised nodules, resembling the paving of a "Roman" road (copyright Photos)

Further, the damage of the mucosa layer is usually measured with extra lesions during the diagnosis. In the case of suspicious CD, these lesions include the existence of frank erythema, frankly swollen mucosa, ulcerated and non-ulcerated stenosis (cf Table 3.3).

The severity of the lesions is usually used as a key parameter for the evaluation of the disease activity. However, an index is required to caracterize the disease state and then guide to a convenient treatment. At present, in clinical practice and clinical trials, the doctors use the endoscopy videos to provide a numerical evaluation about the severity of the principal descriptors of the disease. This

evaluation is called endoscopic score which plays a crucial role in determining the effectiveness of proposed treatments and mucosal healing. In chapter 4, we will describe the numerical assessment of the UC severity used in the medical routine practice. In addition, we will emphasize the endoscopic scores emerged nowadays for disease management and following up as well as treatment efficacy.

3.4 Conclusion

The visualisation of the intestinal mucosa state is crucial to the estimation of the disease activity, evaluation of the treatment efficacy and surveillance. In this chapter, we presented the endoscopy procedures used to investigate the digestive tract in the case of inflammatory bowel diseases. We distinguish two main methods: wireless capsule endoscopy (WCE) and colonoscopy.

In the case of WCE presented in section 3.1, the patient swallows a pill-size capsule including a camera and light (cf Figure 3.4) connected via radio frequency system to a recorder as indicated in Figure 3.5. Despite its ease of utilization and fewer preparation steps required compared to the colonoscopy, the WCE is quite difficult to be controllable once used and the localisation of the lesions regarding the digestive tract is not precisely reported.

In the case of colonoscopy discussed in section 3.2, a colonoscope, illuminated tubular instrument, is used to visualise the interior of the colon and the rectum (see Figure 3.8). Unlike the WCE procedure, some medical interventions can be made during a colonoscopy such as removal of malignant polyps (see Figure 3.15), extraction of biopsies (Figure 3.14) and stoping mucosa's bleeding when it is possible. In contrary to WCE, colonoscopy is an invasive procedure that may (rarely) cause mucosa perforation or infectious.

In the case of detection of small intestine abnormalities and/or suspicious conditions such as Crohn's disease, the gastroenterologist prefers the examination of type WCE because of its ability to take pictures of the whole digestive tract which cannot be reached by the colonoscope. On the other hand, for the examination of the large intestine like the case of ulcerative colitis disease, he uses the colonoscopy able to make precise visualisation of the colon and rectum.

These techniques produce an endoscopic video containing images (cf Figure 3.16) of the gut mucosa (see section 3.3). The video will be reviewed by the doctor to assess the disease severity through the endoscopic scores, a numerical grading. In the following chapter, we will discuss the list of endoscopic scores used in the grading of the UC activity.

Chapter 4

Endoscopic scores

Abstract

Endoscopy remains the key tool for the evaluation of the state of the mucosa, the adaptation of medical treatment and the surgical decision for the management of IBD. To standardize the estimation of disease activity among physicians, an endoscopic scoring system has been established. It is based on what the doctor sees when examining the digestive tract for mucosal appearance defects such as vascular pattern, mucosal erythema, friability, granularity, bleeding, and ulcers. Many ranking scores have been developed for UC and CD. This thesis focuses on the severity of UC, therefore we will highlight the most commonly used scores to measure UC activity in medical practice.

Although incompletely validated, the endoscopic MAYO and Ulcerative Colitis Endoscopic Index of Severity (UCEIS) scores are applied in medical practice. Nevertheless, these scores do not take into account the spatial extent of lesions, their number and their surface area throughout the colon.

The UC is a long-standing disease without a definitive cure. Depending on clinical and endoscopic measurements, disease activity range from quiescent to fulminant. As key of goal, treate-to-target (T2T) is a new concept for IBD's management evolved by the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) committee [PB+15]. The process involves identifying a defined objective that will be achieved through a certain treatment. This objective can be reached by regular monitoring and patient's consultation.

The therapies currently proposed for the management of IBD aim to maintain the healing of the mucosa from existing lesions. Mucosal healing is defined as the healing of all presented ulcerations in the colon and histological inflammation [LKM19]. Specifically, the mucosa is healed when ulcerations are healed in CD and friability as well as ulceration are absent on colonoscopy or sigmoidoscopy in UC [LPF20]. Mucosal healing has been shown to be correlated with decreased rates of surgeries, required hospitalizations, and corticosteroid-associated morbidity. Additionally, the target of therapy is to maintain clinical and endoscopic remission with the prevention of long-term complications such as colorectal cancer, irreversible colon damage, and colectomy. Ultimately, patient quality of life (QoL) is the endpoint of actual therapies for inflammatory bowel disease.

Endoscopy examination, presented in the previous chapter, can enable the assessment of the state of the mucosa and the severity of present lesions as well. It is a key indicator to estimate the disease state, its progression along the digestive tract as well as the efficacy of proposed treatments or therapeutic agents and general disease following up and controlling as much as possible.

Therefore, many measures/indexes of disease activity, called endoscopic scores, have been developed to provide a numerical assessment of disease severity using endoscopy videos. These indexes are used also to evaluate the effectiveness of given therapy. The endoscopic scores involve clinical symptoms, the main mucosal abnormalities (appearance) that the doctor can observe during the endoscopy examination such as friability, granularity, the state of the vascular pattern (Figure 4.2), the presence of bleeding (Figure 4.3), erosions, ulcers such as those shown in Figure 4.4.

At present, several endoscopic scoring indices have been introduced for clinical trials and medical practice for IBD diagnosis and treatment. For UC disease, Truelove and Witts [TW55] were the first to propose an endoscopic scoring system. Since then, many researchers have developed endoscopic scores such as Baron index, Endoscopic Activity Index, Mayo Endoscopic Score and more recently, Ulcerative Colitis Endoscopic Index of Severity and Ulcerative Colitis Colonoscopic Index of Severity. For CD, we can report the CD Endoscopic Index

of Severity [MM89], Simple Endoscopic Score for CD [Dap+04] and Rutgeerts Postoperative Endoscopic Index [GL+15] proposed to measure the endoscopic postoperative recurrence at the neo-terminal ileum. Since this thesis is dedicated to the mathematical modelling of the UC's severity, we will focus only on UC endoscopic scores.

In this chapter, we will present the endoscopic features considered during the computation of endoscopic scores. The disease severity class will be also presented. However, these scores do not include all mucosal abnormalities or do not associate precise evaluation for the measured parameters. These limitations will be discussed as well.

4.1 Truelove and Witts score

The first proposed endoscopic score of the UC disease is the Truelove and Witts Endoscopy Index in 1955 [TW55]. It considers the state of six clinical and biological elements which are stool number per day, nature of blood in stool, temperature, pulse, haemoglobin rate and erythrocyte sedimentation rate (ESR) [1]. The UC disease is then classified into three-level: mild, moderate and severe. The appropriate details are given in Table 4.1. The Truelove ad Witts Index does not present a numerical assessment of the disease which make its use impractical in clinical routine and obvious between doctors.

Element	Description		
Number of bloody stools per day	< 4	4-6	> 6
Frequency of the blood in stool	Rarely	Intermediate	Often
Fever (°C)	Normal	Intermediate	> 37.5
Hemoglobin (gram/deciLiter, g/dL)	Normal	Intermediate	< 10.5
Heart rate pulse (beats per minute, bpm)	< 90	≤ 90	> 90
ESR (millimeter/Liter, mm/L)	< 20	20-30	>30
Disease classification	Mild	Moderate	Severe

Table 4.1: Computation of the Truelove and Witts endoscopic score [Gaw+17]

4.2 Baron index

Next, in 1964 [BCLJ64a], the Baron index is developed considering descriptors such as bleeding, vascular pattern, friability. The detailed computation of the

[1] ESR is a blood test to detect inflammation which can be generated by a condition such as arthritis, vasculitis, or IBD. This test can be also adapted to monitor an existing disease

Baron score is given in Table 4.2. Despite its easy operation, and advantages of presenting a numerical scoring system compared to the Truelove and Witts endoscopic score, the Baron index is still imprecise and not sufficiently correlated to histological abnormalities.

Description	Score	Disease classification
Pale mucosa, Normal vascular pattern clearly visible throughout, No spontaneous bleeding and no bleeding to light touch	0	Normal
Glistening mucosa, erythema and edema, vascular pattern little obliteration	1	Mild
Granularity, bleeding to light touch, no spontaneous bleeding seen ahead of instrument on initial inspection	2	Moderate
Bleeding seen ahead of instrument on initial inspection and on light touching, mucopurulent exudate, occasional mucosal ulceration	3	Severe

Table 4.2: Computation of the Baron Index [Mar06]

4.3 Modified Baron

Feagen et al. in 2005 proposed a modified version of this score, called Modified Baron Index ranging from zero to four and taking into consideration the granularity of the mucosa, bleeding and ulcers [Fea+05]. The score is 0 if the mucosa is normal without the presence of granularity. This evaluation corresponds to endoscopic remission. When the mucosa represents some granularity with damage in vascular pattern, the disease severity score is 1. In the case of friable mucosa, the score is 2. The score of 3 corresponds to tiny ulcerations with spontaneous bleeding, and it is 4 in the case of gross ulcerations. The modified Baron index does not separate the evaluation of the superficial and deep ulcerations.

4.4 Powell-Tuck Index

Later in 1982, Powell-Tuck et al. proposed to combine sigmoidoscopy findings to physical signs such as the body temperature, the abdominal sensitivity to evaluate the UC disease activity (see Table 4.3). Although its effectiveness is to differentiate between the possible nature of mucosa bleeding, the Powell-Tuck Index does not include the evaluation of the appearance of ulcerations and the mucosal healing.

Type	Element	Description	Score
Symptoms	Bowel frequency	3-6	1
		>6	2
	Stool consistency	Formed	0
		Semi-formed	1
		Liquid	2
	Abdominal pain	Before/after bowel movements	1
		Prolonged	2
	Anorexia		1
	Nausea/vomiting		1
	General health	Normal	0
		Slightly impaired	1
		Activities restricted	2
		Unable to work	3
	Extracolonic manifestations	One/mild	1
		More than one/severe	2
Signs	Abdominal tenderness	Mild	1
		Marked	2
		Rebound	3
	Fever	< 37.1	0
		37.1 - 38	1
		>38	2
	Blood in stool	Trace	1
		More than trace	2
	Sigmoidoscopy	Non-hemorrhagic	0
		Friable	1
		Spontaneous bleed	2

Table 4.3: Computation of the Powell-Tuck Index [Wal+98]

4.5 Sutherland index

In the work of Sutherland et al. [Sut+87] in 1987, the assessment of the mucosal healing is included leading to a score of four levels ranging from 0 to 3 for each considered element (cf Table 4.4). The authors classified the mucosal state into normal, slight or moderate friability, exudation. The final score is the sum of each element's score. According to the computation of the Sutherland Index, the disease activity is then classified as follow:

- inactive (Index $\leq$ 2),
- mild (3 < Index < 5),
- moderate (6 < Index < 10),
- severe (11 < Index < 12).

Element	Description			
Diarreha frequency per day	Normal	1-2	3-4	>5
Rectal bleeding	None	Streaks	Obvious	Often
Mucosal appearance	Normal	Slight	Moderate	Exudation or spontaneous bleeding
Physician's global disease evaluation	Normal	Mild	Moderate	Severe
Score	0	1	2	3

Table 4.4: Computation of the Sutherland endoscopic Index [Sut+87]

The Sutherland index is subjective due to consideration of the physician global ranking/rating.

4.6 MAYO score

The MAYO score or the so-called Ulcerative Colitis Disease Activity Index (UC-DAI, [Rut+05]) mixes evaluation of clinical and endoscopic variables. It is composed of 4 categories of measurement (stool frequency, bleeding, endoscopic appearance and overall physician assessment) each one rated from 0 to 3. The final score varies between 0 and 12 (cf Table 4.5). In particular, the MAYO score integrates elements from the clinical state of the patient and the results of the endoscopic exam.

The final UC disease severity assessment via the MAYO score will be:

- inactive disease (Score $\leq$ 2 points)
- mild disease (3 points $\leq$ Score $\leq$ 5 points)
- moderate disease (6 points $\leq$ Score $\leq$ 10 points)
- severe disease (11 points $\leq$ Score $\leq$ 12 points)

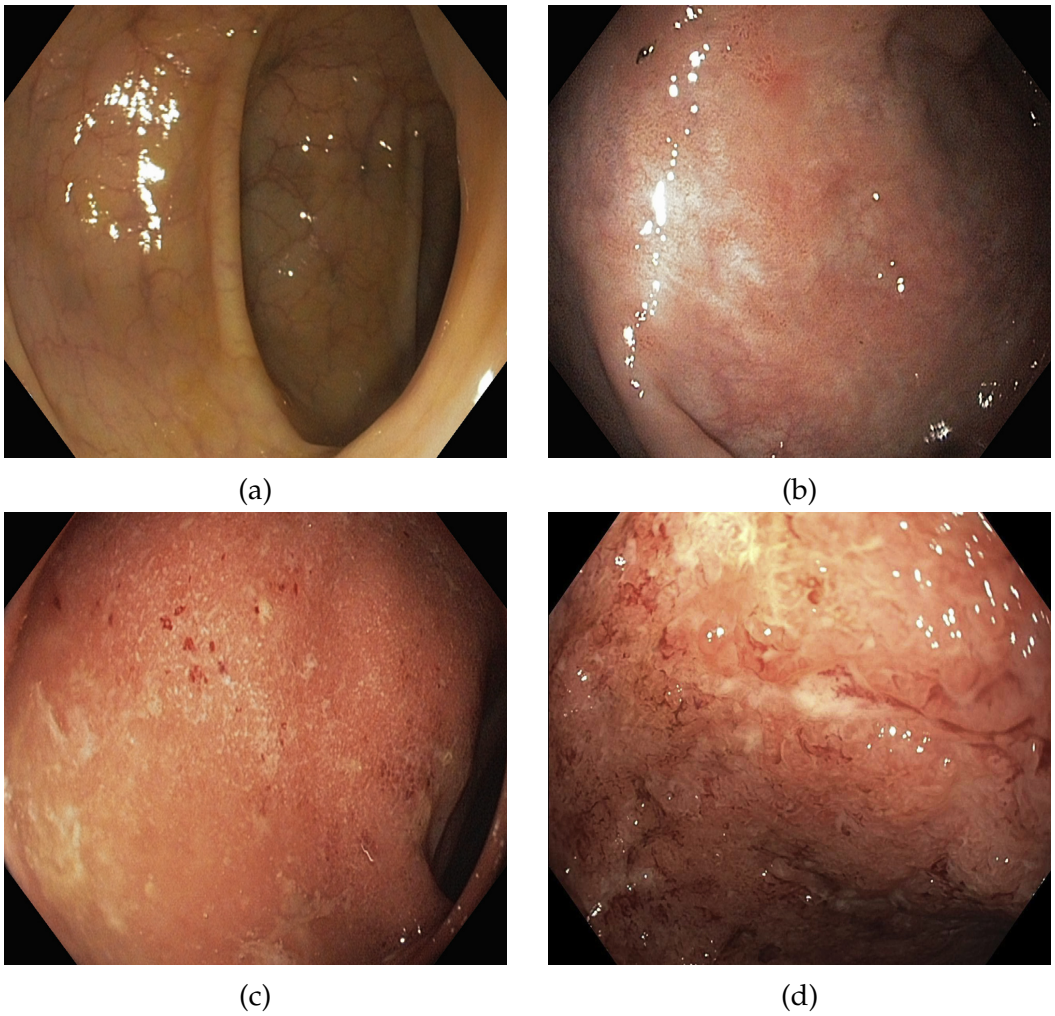
Descriptor	Description	Scoring
Stool frequency	Normal 1-2 stools/day more than normal 3-4 stools/day more than normal $>$ 4 stools/day more than normal	0 1 2 3
Rectal Bleeding	None Visible blood with stool less than half the time Visible blood with stool less than half the time or more Passing blood alone	0 1 2 3
Mucosal appearance at endoscopy	Normal or inactive disease Mild disease (erythema, decreased vascular pattern, mild friability) Moderate disease (marked erythema, absent vascular pattern, friability, erosions) Severe disease (spontaneous bleeding, ulceration)	0 1 2 3
Physician's rating of disease activity	Normal Mild Moderate Severe	0 1 2 3
Total		$0-12$

Table 4.5: Computation of the MAYO score [Lew+08]

Despite its simple use, the MAYO score (cf Table 4.5) presents many defects. There are difficulties in distinguishing between mild UC lesions and moderate UC lesions, more precisely the absence of distinction between superficial and deep ulceration, which causes opposite therapeutic consequences. Therefore, sometimes doctors prefer to use its simplified version, called the MAYO subscore (or Disease Activity Index, DAI) composed of four levels of measurement mainly depending on the presence/absence of bleeding, ulcers, erosions and the state of the vascular pattern (cf Table 4.6).

Disease activity	Endoscopic features	Scoring
Normal or inactive [Lew+08] disease	None	0
Mild	Erythema, decreased vascular pattern, mild friability	1
Moderate	Marked erythema, absent vascular pattern, friability, erosions	2
Severe	Spontaneous bleeding, ulceration	3

Table 4.6: Computation of the MAYO sub-score



(a) (b)

(c) (d)

Figure 4.1: MAYO subscore evaluation. (a) Score 0, normal mucosa appearance with absence of abnormalities. (b) Score 1, the mucosa presents erythema and friability. (c) Score 2, presence of erosions. (d) Score 3, the mucosa is bleeding and ulcerated [LPF20].

4.7 Rachmilewitz index

Later proposed, the index of Rachmilewitz [Rac89] is established during the colonoscopy examination and is more precise than the Baron index. It counts for more elements as the granite appearance and the fragility of the mucosa, vascular pattern, ulcerations and erosions. As the case of the Sutherland index, it varies between 0 and 12 (see Table 4.7), but is not yet validated prospectively. The Rachmilewitz index is also subjective and quite difficult to be used in clinical practice.

Element	Description	Score
Mucosal granularity	Absent	0
	Present	2
Vascular pattern	Normal	0
	Degraded	1
	Invisible	2
Fragility of the mucosa	None	0
	Bleeding on contact	1
	Spontaneous bleeding	4
Erosions and ulcerations	None	0
	No confluent Erosions	2
	Confluent erosions or deep ulcerations	4

Table 4.7: Computation of the Rachmilewitz Index [Mar06]

4.8 Endoscopic Activity Index

In 2010 [Nag+10], Naganuma developed a new endoscopic score called Endoscopic Activity Index (EAI) for evaluating six elements: size and depth of ulceration, redness, bleeding, mucosal oedema and mucous exudate (cf Table 4.8). The EAI score is useful in the case of acute ulcerative colitis patients, and it is correlated to clinical activity [Nag+10].

Feature	Description	Score
Size of ulcers	None	0
	Erosion/small ulcer	1
	Intermediate	2
	Wide-ranged mucosal defects	3
Depth of ulcers	None	0
	Shallow	1
	Intermediate	2
	Deep	3
Redness	None	0
	Mild	1
	Marked	2
Bleeding	None	0
	Contact bleeding	1
	Spontaneous	2
	Massive	3
Edema	None	0
	Mild	1
	Moderate	2
	Severe	3
Mucous exudate	None	0
	Mild	1
	Marked	2

Table 4.8: Computation of the EAI score [Nag+10]

4.9 Ulcerative Colitis Endoscopic Index Score (UCEIS)

In 2012 [Tra+12], Travis et al. proposed a simple and reliable endoscopic severity score UCEIS. Their study was performed by 30 investigators of different nationalities into two phases. They used a total of 680 sigmoidoscopy videos containing five videos for normal/healthy individuals, and five videos for severe UC hospitalized patients. For the first stage of the study, 10 investigators have evaluated 16 to 24 randomly selected videos according to their experience to attempt the presence and absence of 10 principal UC endoscopic descriptors presented in previous studies [BCLJ64b; Gom+86; Per+87; Orl+98; DLLA04].

These descriptors are:

- Vascular pattern
- Mucosal erythema
- Mucosal granularity
- Mucosal oedema
- Mucopus
- Bleeding
- Incidental friability
- Contact friability
- Erosions and ulcers
- Extent of erosions or ulcers

Descriptor from litterature	ponderated kappa intra-observer	ponderated kappa inter-observer
Erosions and ulcers	0.65	0.45
Vascular pattern	0.61	0.42
Bleeding	0.6	0.37
Extent of erosions or ulcers	0.60	0.42
Spontaneous friability	0.49	0.40
Mucopus	0.47	0.40
Mucosal granularity	0.45	0.34
Mucosal oedema	0.43	0.31
Mucosal erythema	0.43	0.35
Contact friability	0.33	0.34

Table 4.9: Intra and inter observers variations to define the 3 descriptors of the UC disease to generate the UCEIS score [Tra+12]

More details about the definition of these descriptors as well as their scoring are presented in reference [Tra+12].

For the second stage, the 30 investigators assess the overall UC activity in 25 different videos on a 0 to 100 visual analogue scale (VAS). Table 4.9 illustrates the numerical agreement between investigators for the selected descriptors based on the ponderated Cohen's kappa coefficient (k) [2].

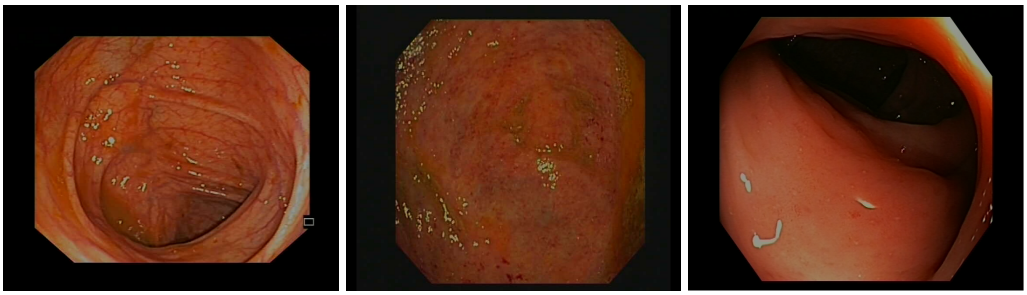
Descriptor	Description	Definition	Scoring
Vascular pattern	Normal	Normal vascular pattern with arborisation of capillaries clearly defined or with blurring or patchy loss of capillary margins	0
	Patchy obliteration	Patchy obliteration of vascular pattern	1
	Obliterated	Complete obliteration of vascular pattern	2
Bleeding	None	No visible blood	0
	Mucosal	Some spots of coagulated blood on the surface of the mucosa ahead of the scope that can be washed away	1
	Luminal mild	Some free liquid blood in the lumen	2
	Luminal moderate or severe	Frank blood in the lumen ahead of the endoscope or visible oozing from the mucosa after washing intraluminal blood, or visible oozing from a hemorrhagic mucosa	3
Erosions and ulcers	None	No visible erosions or ulcers	0
	Erosions	Tiny defects ($\leq$ 5 mm) in the mucosa of a white or yellow color with a flat edge	1
	Superficial ulcer	Larger defects ($>$ 5 mm) in the mucosa, which are discrete fibrin-coverted ulcers when compared with erosions but remain superficial	2
	Deep ulcer	Deeper excavated defects in the mucosa with a slightly raised edge	3
Total			$0-8$

Table 4.10: Computation of the UCEIS score [Tra+13]

The final UCEIS score rates the most severe lesion of the three principal descriptors type vascular pattern, bleeding and erosions/ulcerations chosen according to the maximum values of the ponderated Cohen's kappa coefficient (k).

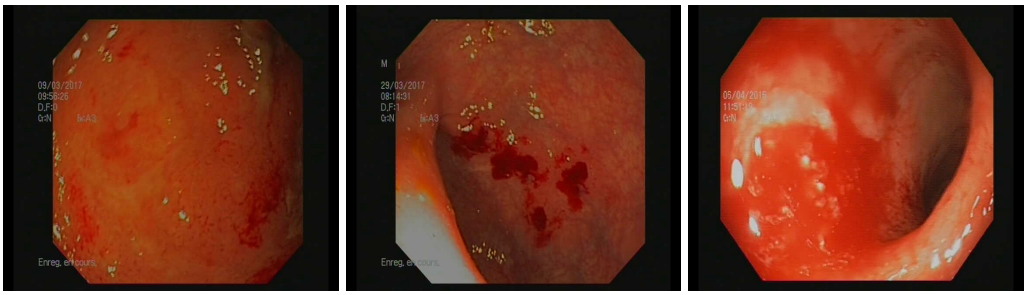
In Table 4.10, we present the computation of the final score, UCEIS. Bleeding, erosions and ulcers lesions are ranked between 0 and 3, while the vascular pattern is ranked between 0 and 2. The sum of all sub-rates is the final score (0-8). The more the score is high the more the disease is active or severe.

[2] kappa is a statistical variable that measures the agreement between two raters. $k = 1$ means complete agreement between investigators



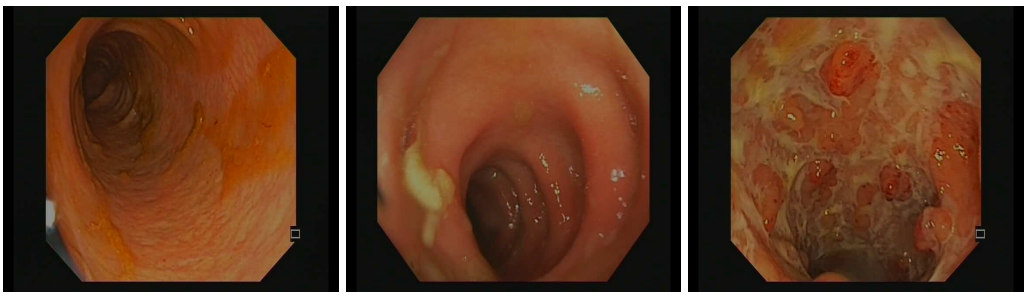
(a) (b) (c)

Figure 4.2: UCEIS sub-score for the descriptor vascular pattern (from left to right): Score = 0, Score = 1, and Score =3.



(a) (b) (c)

Figure 4.3: Some bleeding lesions findings in a colonoscopy video. (a) Spontaneous bleeding scored 1, (b) Moderate bleeding patchs scored 2 and (c) Severe bleeding scored 3



(a) (b) (c)

Figure 4.4: Examples of ulcerations and erosions lesions in a colonoscopy video. (a) Erosions corresponding to superficial mucosal damaging are generally scored 1 (b) Ulcers caracterized by a necrosis at the level of the mucosa thickness are scored 2, (c) Deep erosions and ulcers affecting the whole colon lining are scored 3

Although the UCEIS does not present a computation of mucosal healing, it is a simple and practical method to measure the severity of the UC disease. As for

the previous scores, UCEIS rates the most severe lesion of type bleed, ulcer, and vascular pattern and does not take into consideration the spatial extent of the lesions in the colon.

4.10 UC Colonoscopic Index of Severity (UCCIS)

A new UC Colonoscopic Index of Severity, denoted by UCCIS, was introduced by Samuel et al. in 2013 [Sam+13]. It counts for five mucosal abnormalities: the vascular pattern, bleeding, ulcers, mucosal friability and granularity. The authors assess the presence and severity of each lesion on the five colon principal segments (rectum, sigmoid, descending, transverse, and cecum/ ascending).

As supplement disease evaluation, they decide to add evaluation of two elements: the segmental assessment of endoscopic severity denoted by SAES, and the global assessment of endoscopic severity denoted by GAES for the considered colon segments in addition to a 10 cm visual analogue scale (from normal to severe).

The grade zero of the parameters SAES and GAES corresponds to a normal vascular pattern with the absence of bleeding and ulcers. Grade one ties in the case of the presence of erythema and patch/total obliteration of the vascular pattern. Grade two is given for the presence of friability with bleeding, granularity and erosions. The highest level four corresponds to spontaneous bleeding and/or gross ulcerations.

The authors reported a high-level inter-observer agreement. However, the final numerical score is not clearly defined which make its utilization difficult for medical issues.

4.11 Discussion

Endoscopic grading systems are suitable to outline the mucosal appearance and translate it to a numerical value. They are useful to the assessment of the disease activity which improves the following-up of the disease and the efficacy of proposed therapies as well.

The Truelove and Witts Endoscopy Index [TW55] presented in Table 4.1 classifies the severity of the disease by counting for six clinical and biological elements which are stool number per day, nature of blood in stool, temperature, pulse, haemoglobin rate and erythrocyte sedimentation rate (ESR). It is quite difficult to be applied in medical routine practice.

The Baron index [BCLJ64a] considers other types of descriptors such as bleeding, vascular pattern, friability (cf Table 4.2) but it is still inaccurate and not correlated with histological abnormalities.

The modified Baron Index [Fea+05] is limited to four parameters and does not separate the evaluation of the superficial and deep ulcerations.

These scores and the Rachmilewitz Index [Rac89] (cf Table 4.7) as well miss a precise evaluation about the state/severity of bleeding and ulcerations lesions.

The scores used nowadays in clinical practice, UCEIS and MAYO sub-score are based on 3 (see Table 4.10) and 6 (cf Table 4.6) parameters respectively depending on the most relevant lesions of the UC disease.

In Table 4.11, we summarised the endoscopic variables of the mucosa's abnormalities reported in the endoscopic scores. They involve bleeding, ulcerations as well as mucosal friability, granularity and blood vessel visibility.

Score	Bleeding	Erosion/ Ulcer	Exudation	Friability	Granularity	Edema	Vascular pattern
Truelove & Witts					✓		
Baron index	✓			✓			✓
Modified Baron	✓	✓		✓	✓		✓
Powell-Tuck	✓						
Sutherland	✓		✓	✓			
Mayo	✓	✓		✓			✓
Rachmilewitz		✓		✓			✓
EAI	✓	✓	✓			✓	
UCEIS	✓	✓					✓
UCCIS	✓	✓		✓	✓		✓

Table 4.11: Common descriptors of the UC used for the computation of the endoscopic scores

However, mucosal remission is not yet reported when calculating scores. Consequently, this leads to a lack of evaluation of the effectiveness of the treatment. Most of the scores cited in the chapter evaluate the state of the mucosa without reporting the colonic segment concerned. The UCCIS is the single score taking into account all colonic segments. It has been shown to be correlated between raters. On the other hand, the raters inter-reliability for the scores Baron, EAI, Mayo sub-score, UCEIS and UCCIS ranged from 0.44 to 0.97, while the intra-reliability was in the range 0.41-0.81 [Vas+18]. In addition, UC endoscopic scores, Rachmilewitz and UCCIS exhibit correlation with objective inflammation biomarkers such as albumin, blood leukocytes, C-reactive protein and haemoglobin [Vas+18]. Furthermore, there is no score that has been fully

validated i.e. they should measure the result that it is intended to assess.

Additionally, it is well-known that the severity assessment is variable between experts, so colonoscopy videos are commonly read by third-party experts in a procedure called central reading. A central reading system is often used to mean that the interpretation of imaging findings is not, or not only, done by local endoscopists of the hospital but instead is supervised, amended or adjudicated by at least one off-site/external reader who is neither aware of nor influenced by knowledge of the patient characteristics or treatment arm assignment. This helps to ensure having a more accurate and unbiased grading of disease activity [Pan+16].

To date, the Mayo Endoscopic Subscore (cf section 4.6) and UCEIS (cf section 4.9) are mainly applied to clinical trials and medical practices. Indeed, the MAYO score integrates elements from the clinical state of the patient and the results of the endoscopic exam. In comparison to the UCEIS score, it can be relevant to only focus on the endoscopic sub-score instead of the complete MAYO score.

Although practical and simple, and like all the UC proposed endoscopic scores, MAYO and UCEIS scoring systems rely on the gastroenterologist personal interpretation when reviewing the colonoscopy video and target the evaluation of the most severe lesions. They miss the assessment of the spatial extent of the lesions along the colon, their amount as well as the relapse and remitting course of the disease. Later, the lack of such type of information may lead to an incomplete evaluation of the disease severity, and consequently the monitoring of the disease course. The extent of disease is relevant to disease classification and prognosis because of the correlation between extensive ulcerative colitis and the high risk of complications such as colorectal cancer and colectomy.

In this thesis, we decide to consider the spatial extent of the endoscopic lesions to model the disease severity. We believe by this way that the severity assessment of the disease activity will be more precise than considering the most severe lesion like the case of the UCEIS and MAYO (complete or sub) scores. Taking into account the spatial extent of lesions in the colon based on the colonoscopy video was the core of the works presented hereafter in this manuscript:

- In chapter 7, we visualize the spatial distribution of the UC lesions along the colon. The obtained results lead to precise evaluation of the inflammation state of the colon and a fine distinction between patients with the same endoscopic scores. We showed that in actual medical practice, the doctor rates the disease of two patients similarly although they have a different mucous impairment due to the UC disease.

- In chapter 8, we proposed a linear model taking into consideration the lesion's count at each segment of the colon and demonstrating high agreement with Mayo endoscopic subscore and UCEIS using the mean square error criteria.
- In chapter 9, we tried to bring some responses to a set of medical hypotheses concerning the abundance of the lesions in the colon. In addition, we studied the correlation between the count of lesions as well as the endoscopic scoring variability between the gastroenterologists.
- Finally, in chapter 10, we estimate the lesions' invasion time by using the spatial extent of the lesions through a convenient partial differential equation.

4.12 Conclusion

Endoscopic numerical assessment of disease activity is steadily used in medical practice. It permits the diagnosis, prognosis, monitoring of disease course and determination of response to treatment.

Truelove and witts were the first to propose the evaluation of the mucosa appearance through rigid sigmoidoscopy (cf section 4.1). Since then, numerous endoscopic scores were developed such as Baron index in section 4.2, Rachmilewitz in section 4.7, Powell-Tuck score in section 4.4, Mayo in section 4.6, UCEIS in section 4.9 and many others. They count for the principal descriptors of UC disease: bleeding, erosions, ulceration, mucosal friability and granularity, also the vascular pattern (more details are given in Table 4.11). Their computation takes into account only the most severe lesions without considering their count, area or extent along the colon. Most of the proposed scores are subjective and don't consider the mucosal remission and the disease course as well.

Currently, the doctors use the Mayo endoscopic subscore and the UCEIS score in clinical trials and medical practice (see section 4.11). Like all the other scores, they are incomplete to measure the disease activity due to lack of parameters such as C-reactive protein, faecal calprotectin and the lesions spatial extent along the colon.

Motivated by these reasons, we will therefore take into account the spatial extent of lesions to visualize their spatial distribution in chapter 7, to mathematically model the UC severity in chapter 8, to respond to some medical questions in chapter 9 and estimate the lesions' invasion time in chapter 10.

Chapter 5

Vatic database

Abstract

we will bring details about the database used in all this thesis project. It is composed of colonoscopy videos for patients affected by UC disease. And, we will expand the annotation process applied by our collaborators gastroenterologists to indicate in each video image the abormal pixels corresponding to UC lesions. At the end of the chapter, we introduce the data that we take up from the annotations in order to represent the extent and/or severity of the lesions along the colon. These extracted informations will be used in all the following chapters of this manuscript.

5.1 Database: Vatic

Due to our collaboration with gastroenterologists working on UC, all the work presented in this thesis concerns UC patients rather than CD. However, as the two IBDs share some visual similarities, some of presented works (chapters..??) could be extended for some CD problems. We consider a database of 37 (complete) colonoscopy videos corresponding to patients affected by UC disease in different stage and forms, for both genders, woman and man. John Chaussard, have developed a tool/an interface based on Vatic software [VPR13], proposed in 2013 for car tracking, to allow doctors to mark UC abnormalities in each image of the available videos (Figure 5.1). From Figure 5.1, we can see different number of lesions (bleeding and ulcers) which are delineated by rectangles of different sizes depending on their shapes area on the mucosa.

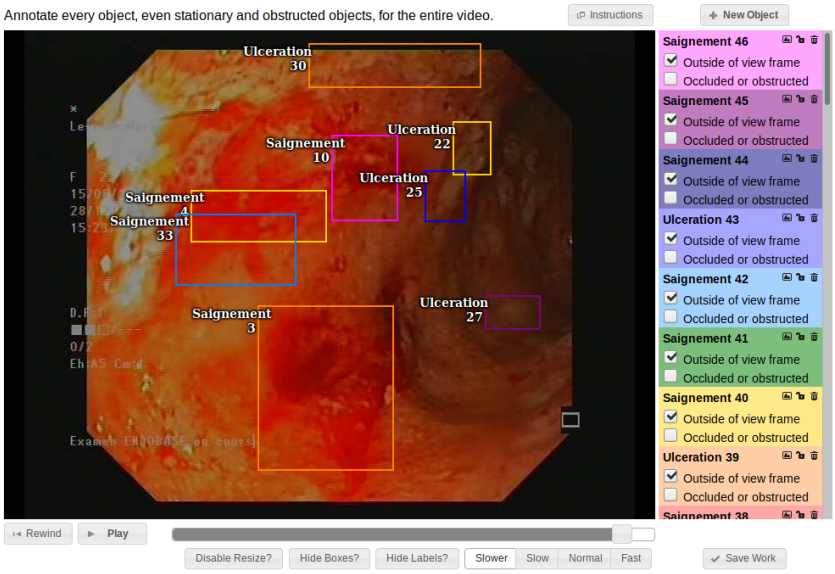


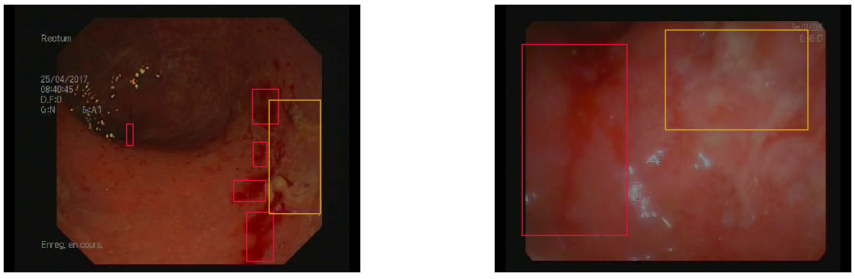
Figure 5.1: Annotation of a colonoscopy image by gastroenterologist using Vatic software

	UCEIS			MAYO		
Video number	XT	CS	YB	XT	CS	YB
1	8	8	7	3	3	3
2	7	6	8	3	3	3
3	2	2	3	1	1	2
5	1	1	1	1	1	1
6	3	2	1	1	1	1
7	1	4	0	1	2	0
9	5	5	4	3	3	3
10	5	5	5	3	3	2
13	5	6	3	3	3	2
15	5	3	2	3	1	1
16	7	6	3	3	3	2
17	5	4	3	3	3	2
21	7	7	4	3	3	2

Table 5.1: Endoscopic scores (UCEIS and MAYO sub-score) for 13 patients from Vatic database evaluated by Drs XT, CS and YB

5.2 Data analysis/generation

In Figure 5.1, we have many number of lesions for each descriptor (bleeding and ulcerations). It corresponds to a new added annotation on the frame and not new type of descriptor and not other form or nature (luminal, mucosal, superficial..). So, to facilitate the reading process of annotations we decide to mark all bleed annotations by red rectangles and ulcer annotations by yellow rectangles as in Figure 5.2. The video timer is then related to the colon curvilinear abscissa and the annotations are represented either by their amount in a given frame (**count of lesions**, cf section A.3) or their area comparing to the total image frame size (**percentage of lesions**, cf section A.4).



(a) (b)

Figure 5.2: Annotated frames

Vatic software permit us to transform the video into a text document to analyze in order to retrieve the doctors's annotations and then use them in our work.

We launch the following command lines in a Linux terminal to save text format of a given annotated video:

Code chunk 1: shell

```
ssh comptepersonnel@laga-vatic
Mot de passe: ***
cd vatic
cd vatic
turkic dump video.avi -o video.txt
exit
scp comptepersonnel@laga-vatic:vatic/vatic/video.txt ./
Mot de passe: ***
```

Code chunk 2: shell (part 2)

```
head "video1.txt"
```

Interpret with `shell`

```
0 475 129 598 406 0 0 0 0 "Ulceration"
0 487 148 592 364 1 0 0 1 "Ulceration"
0 500 167 586 323 2 0 0 1 "Ulceration"
0 513 187 581 282 3 0 0 0 "Ulceration"
0 502 171 593 328 4 0 0 1 "Ulceration"
0 491 155 605 374 5 0 0 1 "Ulceration"
0 481 139 618 421 6 0 0 0 "Ulceration"
0 478 140 615 427 7 0 0 1 "Ulceration"
0 475 142 613 433 8 0 0 1 "Ulceration"
0 473 144 611 439 9 0 0 0 "Ulceration"
```

With the following signification of each column [VPR13]:

- Column 1: Track lesion identity
- Column 2: xmin the top left x-coordinate of the annotation box
- Column 3: ymin the top left y-coordinate of the annotation box
- Column 4: xmax the bottom right x-coordinate of the annotation box
- Column 5: ymax the bottom right y-coordinate of the annotation box
- Column 6: frame number of the presented annotation
- Column 7: lost. If 1, the annotation is outside of the view screen
- Column 8: occluded. If 1, the annotation is occluded.
- Column 9: invisibility of lesion
- Column 10: lesion label (bleeding or ulceration) enclosed in quotation marks

Code chunk 3: Data-code.py

```
# python librairies
import matplotlib as mpl
mpl.use('PDF')
import matplotlib.pyplot as plt
import numpy as np
import cv2, io, string
from math import sqrt,exp
import pickle
```

Code chunk 4: Data-code.py (part 2)

```
# Data base informations
num_video= list of video number in Vatic database
fichier_image_video= list of folders containing images of each video
```

Code chunk 5: Data-code.py (part 3)

```
#computation of the amount of lesions per frame
def Count_Lesions(num_video):
    fichier_txt="Folder of videos in format text/video%d.txt"%(num_video)
    fichier = io.open(fichier_txt, 'r')
    Lignes = fichier.readlines()
    Maxannot = Lignes[len(Lignes) - 1]
    lMaxannot = Maxannot.split()
    Nfmax = int(lMaxannot[5])
    NannotS_frame,NannotU_frame=np.zeros(Nfmax), np.zeros(Nfmax)
    for frame in range(0, Nfmax):
        for ligne in Lignes:
            l = ligne.split()
            if (int(l[7]) + int(l[6])) == 0 and int(l[5]) == frame:
                if '"Saignement"' in l:
                    NannotS_frame[frame]+=1
                elif '"Ulceration"' in l:
                    NannotU_frame[frame]+=1

    return NannotS_frame,NannotU_frame
```

Code chunk 6: Data-code.py (part 4)

```
# FIV: fichier_image_video
# computation of percent of abnormal pixels per frame
def Data_perc_annotations(num_video,FIV):
    fichier_video="Folder of images of video/%s"%(FIV)
    fichier_txt="Folder of videos in format text/video%d.txt"%(num_video)
    fichier = io.open(fichier_txt, 'r')
    Lignes = fichier.readlines()
    Maxannot = Lignes[len(Lignes) - 1]
    lMaxannot = Maxannot.split()
    Nfmax = int(lMaxannot[5])
    Masque = "%s/Test-Fig-Masque.png" % (fichier_video)
    M = cv2.imread(Masque)
    M = cv2.erode(M[:, :, 1], 255 * np.ones((5, 5), np.uint8))
    M1 = cv2.imread(Masque)
    M1 = cv2.erode(M1[:, :, 1], 255 * np.ones((5, 5), np.uint8))
    n_masque = np.sum(M == 255)
    per_bleed, per_ulcer=np.zeros(Nfmax),np.zeros(Nfmax)
    for frame in range(0, Nfmax):
        for ligne in Lignes:
            l = ligne.split()
            if (int(l[7]) + int(l[6])) == 0 and int(l[5]) == frame:
                if '"Saignement"' in l:
                    M[int(l[2]):int(l[4]) + 1, int(l[1]):int(l[3]) + 1] = 100
                elif '"Ulceration"' in l:
                    M[int(l[2]):int(l[4]) + 1, int(l[1]):int(l[3]) + 1] = 200

        M[M1 == 0] = 0
        per_bleed[frame]=float(np.sum(M==100))/n_masque
        per_ulcer[frame]=float(np.sum(M==200))/n_masque

    return per_bleed,per_ulcer, Nfmax
```

Code chunk 7: Data-code.py (part 5)

```
# application to Vatic database
y_b_count, y_u_count=[[]]*len(num_video),[[]]*len(num_video)
y_b_percent, y_u_percent=[[]]*len(num_video),[[]]*len(num_video)
for index in range(len(num_video)):
    y_b_c,y_u_c= Count_Lesions(num_video[index])
    y_b_count[index].append(y_b_c)
    y_u_count[index].append(y_u_c)
    y_b_p,y_u_p= Data_perc_annotations(num_video[index],fichier_image_video)
    y_b_percent[index].append(y_b_p)
    y_u_percent[index].append(y_u_p)

# Saving results
with open('Folder-DataVatic-Count.pickle','wb') as f:
        pickle.dump([y_b_count,y_u_count],f)

with open('Folder-DataVatic-Percent.pickle','wb') as f:
        pickle.dump([y_b_percent,y_u_percent],f)
```

Code chunk 8: Data-code.py (part 6)

```
# Use of saved results
with open('Folder-DataVatic-Count.pickle','rb') as f:
        [y_b_count,y_u_count] = pickle.load(f)

with open('Folder-DataVatic-Percent.pickle','rb') as f:
        [y_b_percent,y_u_percent]= pickle.load(f)

# bleeding
Data_Perc_b = [ [individualArray] for individualArray in y_b_percent]
Data_Perc_b=Data_Perc_b[0][0]
# ulcer
Data_Perc_u = [ [individualArray] for individualArray in y_u_percent]
Data_Perc_u=Data_Perc_u[0][0]
```

If we need to represent the disease by the amount of lesions (bleed, ulcers) in a given abscissa of the colon, the video frame at Figure 5.2a will correspond to the couple (5,1) and the Figure 5.2b to the couple (1,1). Otherwise, in the case of use of the area of annotations compared to the total image size, the first image will represented by the couple (72%, 10%) and the second one by the couple (41%, 21%).

Code chunk 9: Data-code.py (part 7)

```
print Data_Count_b[8][850],Data_Count_u[8][850]
print Data_Perc_b[8][850],Data_Perc_u[8][850]

print Data_Count_b[22][364],Data_Count_u[22][364]
print Data_Perc_b[22][364],Data_Perc_u[22][364]

print(len(Data_Count_b[0]))
```

Interpret with `python2`

```
5.0 1.0
0.7298445877498579 0.10278104274101606
1.0 1.0
0.4124588616710282 0.21212362642413693
812
```

We think that the two types of extracted data (count of annotations, percent of abnormal pixels) can bring several informations about the inflammation of the colon. This will be discussed progressively in the next chapters.

Chapter 6

Automatic detection of Ulcerative Colitis lesions in colonoscopy videos with limited quality annotations

Abstract

The severity grading using the UCEIS and MAYO-sub scores is based on the subjective interpretation of the doctor and does not take into account the size of the lesions, their number and their distribution. Automatic lesion detection methods can enable fine-grained assessment of lesion severity but require a training stage based on manual annotation which is a laborious and time-consuming task. Most of the current methods use generic datasets that are biased towards capsule endoscopy and are not adapted to the locally available hardware. In this chapter, we propose an algorithm to automatically detect the ulcerative colitis lesions from colonoscopy videos obtained during our collaboration with the Gastroenterology group at the Bichat and Beaujon hospitals.

During the learning phase, the doctors used rectangles to delineate bleeding and ulcers. This annotation process leads to many errors due to many mislabeled pixels, especially in the corners of the rectangles. This affects later the evaluation of the models' performance and the ability to find correct models. As a solution, we propose to evaluate the model sensitivity on the annotation level considering that a model correctly identifies a lesion if it agrees with the expert on a subset of the annotation, and count the detected annotations weighted by their area or their count. The specificity is kept calculated at the pixel level. On a data set of 10 colonoscopy videos, we explore the set of linear classifiers in suitable colorspaces and propose an efficient sampling scheme that rejects trivial models.

We find that all sampled models have a non-zero true negative rate. Despite the limited quality of the annotations, we correctly identify the lesions (93% specificity / 89% for bleeding and 57% specificity / 83% sensitivity for ulcers). The performance of the detectors was computed reliably. We evaluate sensitivity and specificity on 20 random

subsets containing 10% of the images and obtain similar performance for the same patient for all models.

However, the inter-patient performance is variable and the best models fail on some patients (sensitivity below 20% in some cases). This suggests that the variability in the appearance of the lesions between patients should be considered before conducting an automatic process.

To evaluate the state of the mucosa, the gastroenterologist has to review the video obtained after the investigation of the digestive tract using the WCE (cf section 3.1) or the colonoscopy procedure (cf section 3.2). The wireless capsule camera takes 50000 to 60000 pictures with the rate of 2 to 4 frames per second (fps).

On the other hand, it has been recommended that the time that an endoscopist takes to withdraw the endoscope should be in an average of at least six to ten minutes in normal colons, in the absence of biopsies and polypectomies [Rex07]. The obtained video may contain 10800 to 180000 frames. In the two types of medical examination, the doctor has to review a huge amount of images to detect mucosa abnormalities. In general, the abnormalities are present in only a few images. This is a tedious task that requires a lot of time and a high level of concentration of the doctor to detect by naked eye the endoscopic abnormalities.

In addition, the severity assessment of the ulcerative colitis disease using the endoscopic scores such as UCEIS and MAYO sub-score (see **??** and **??**) is subjective because it relies on the gastroenterologist personal interpretation. Thus, it is not reproducible neither inter-doctor nor intra-doctors.

Consequently, automatic lesion detection techniques can bring significant benefits by improving the reproducibility of severity assessment while decreasing the physician burden. We use a database from 37 colonoscopy videos obtained from the hospitals Bichat and Beaujon in Paris. Despite the method proposed in this chapter will be applied to abnormalities found in colonoscopy videos, it can also be applied to the same type of lesions that may be present in the case of WCE videos because of similar appearance.

The automatic detection technique requires a phase of abnormalities annotation. Manual and precise delimitation of the lesions in the colonoscopy videos is a laborious mission requiring considerable effort and time from doctors because of the large number of images and the complicated forms of the lesions as one can see in the image.

Thus, to facilitate the annotation process, we suggest to doctors to delineate the bleeding and ulcers lesions of the colonoscopy videos using rectangles. Indeed, during the colonoscopy procedure, the doctor aims to visualise the colonic mucosa and control the endoscope directions, which may lead to visualising a given lesion on successive frames. To help the doctors, to annotate the lesion only on its first appearance/once a time, we proposed to use an interface inspired by the Vatic software [VPR13], published in 2013 for video annotation for car tracking.

6.1 Lesions annotation process

The principle of the Vatic software's annotation process is that the doctor plays the video, draws bounding boxes around objects of interest, namely the lesions in our case, and tracks/follows each lesion throughout its appearance in the video. Each object can have multiple attributes that further describe its actions. Doctors can adjust reading speed too, search throughout the timeline, and mark lesions occluded or off-screen. Since scenes quickly become cluttered, the doctor making the annotations can lock the lesions to prevent accidental changes to their paths.

The main advantage of this software is that the doctor annotates all of the lesions in the first frame of the colonoscopy video, then he advances to the next frame, updates all of the annotations, and repeats this process for the entire video. Consequently, the doctor watches the video only once since when a frame is labelled, he will never need to return to it.

In Figure 6.1, one can see there are two objects to track their paths in the video: bleeding and ulcers. For the appearance of a new lesion, the doctor clicks on the button "New object" on the right, then he draws a box at the desired area of the figure and adjusts the size of the box later to contain all the lesions.

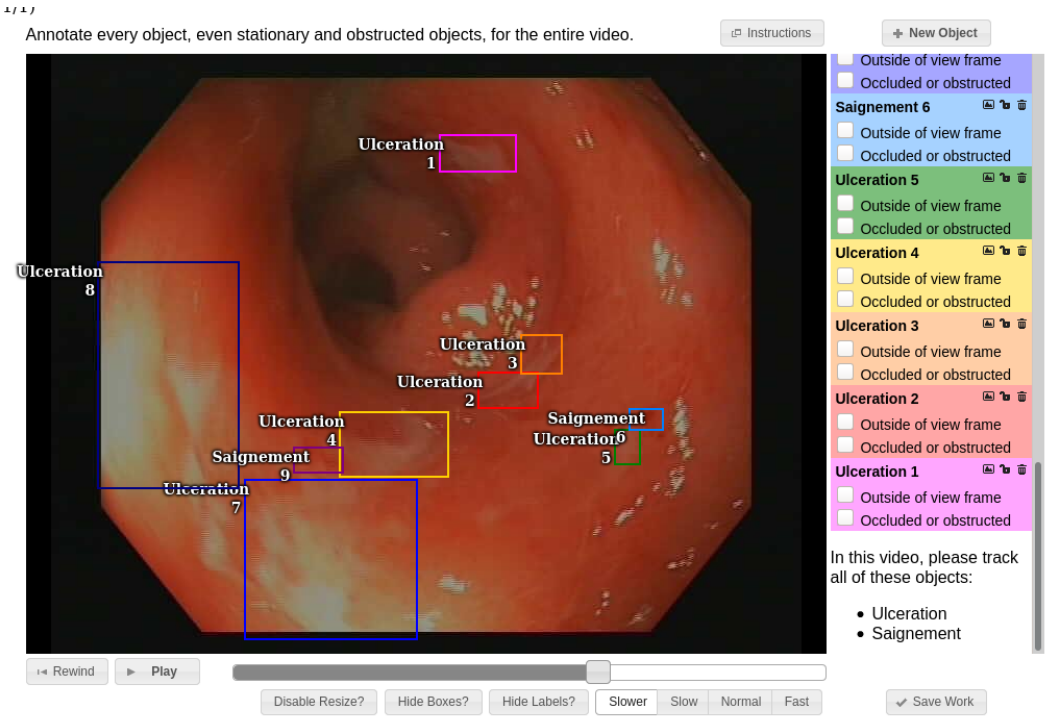


Figure 6.1: Annotation of a colonoscopy image by gastroenterologist using Vatic software

Although the annotation process can be practised to obtain a considerable database, it leads to many errors. The lesions are in general of arbitrary shapes (see Figure 6.2 for ulcers and Figure 6.1 for bleeding) which make their delineation with boxes like made in our work imprecise. Many normal pixels have the label abnormal once they are inside a given box.

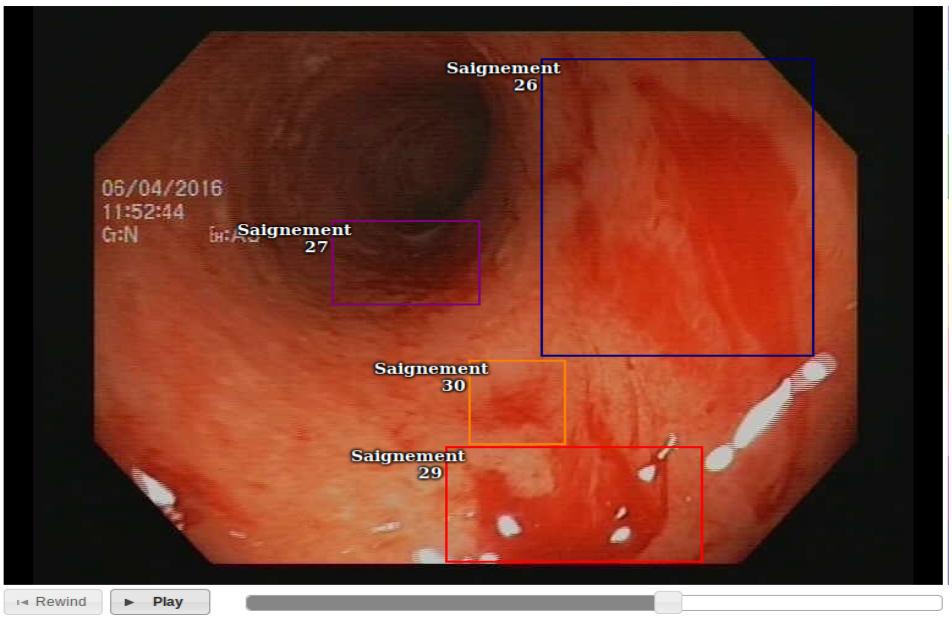


Figure 6.2: Annotation of bleeding using Vatic software

Therefore, the performance of the detector will not be computed accurately. To deal with this problem, we have proposed taking these pixels into account when calculating the detector's capacity to correctly identify lesion pixels. We then proposed to modify the sensitivity criteria from pixel level to annotation level and keep the specificity unchanged i.e. it will be kept computed on the pixel level. In the next section, we will bring the details about the proposed modification of abnormal pixels detection rate.

6.2 Statistical criteria to the performance evaluation

During the application of our algorithm, there are four distinct cases of abnormality detection outcomes that may occur [SW11] (see Figure 6.3):

- true positives (TP), i.e. the number of abnormal pixels correctly identified as abnormal
- true negatives (TN), i.e. the number of normal pixels correctly identified as normal
- false positives (FP), i.e. the number of normal pixels identified as abnormal

- false negatives (FN), i.e. the number of abnormal pixels identified as normal

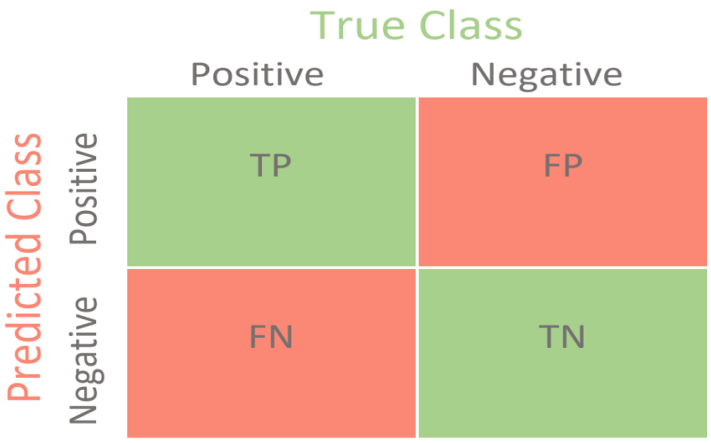


Figure 6.3: Confusion matrix to compute the lesions classification probabilities. True class designs the labels by the doctors whereas predicted class is related to the detector classification on the pixels level

In Figure 6.3, the positive rate, denoted by "P", indicates the count of positive pixels i.e. those labelled as lesion, whereas the negative rate, denoted by "N", is the count of pixels labelled as normal, namely they are the pixels out of the annotations in our case for example.

In the textbook statistics, many classification performance criteria can be extracted from the confusion matrix such as:

- Specificity, also called the true negative rate, measures the proportion of normal pixels correctly identified with the detector and it is computed by the formula:

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{6.1}$$

- Sensitivity, also called the true positive rate, measures the correctness of abnormal pixels detection and it is given by the following formula:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6.2}$$

- Precision, the positive predicted value given by

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6.3}$$

- Accuracy

$$\text{Accuracy} = \frac{(TP + TN)}{P + N} \tag{6.4}$$

- F_1 score,

$$F_1 \text{ score} = \frac{2TP}{2TP + FP + FN} \tag{6.5}$$

- Matthews correlation coefficient (MCC),

$$\text{MCC} = \frac{TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{6.6}$$

The methods of automatic detection of endoscopic lesions were emerged many years ago and begin challenging with machine learning evolution in medical applications. For the detection of the endoscopic abnormalities in wireless capsule endoscopy or colonoscopy videos, computer analysis techniques try to solve two kinds of problems. First, given a set of pixels, a region of interest (ROI) or a complete frame, binary classification algorithms are used to select a label between "lesion" (also called abnormal) or "not lesion" (also called normal). Second, segmentation algorithms are used to find regions with the same label such as bleeding or ulcers regions. The evaluation of performance is generally made by the criteria of the section 6.2.

In the following, we will review the methods actually proposed to detect bleeding and ulcers found in wireless capsule endoscopy or colonoscopy videos. Then we will exhibit our method to detect the lesions in our database.

6.3 State of the art of endoscopic lesions detection methods

Many.....

6.3.1 State of the art for bleeding automatic detection methods

Color information is the primary indicator used by specialists to recognize mucosal lesions from the surrounding normal mucosa. In particular, bleeding lesions usually show like dark red areas with diverse shapes and sizes. Most classification methods attempt to select a suitable color space, with maximum contrast between bleeding and non bleeding regions.

As bleeding pixels are red, it is natural to consider detection and classification in the RGB colorspace, or direct transformations of RGB ([FMG11; Fu+14;

Gho+14a; Gho+18; KFR18; Gho+14b; SBW14]).

In 2011, Fu et al. [FMG11] propose to work on the ratios (R/G,R/B,R/G+B+R) for each pixel. They train a 3-layer perceptron with 2 nonlinear outputs and 20 neurons in the hidden layers on 1000 bleeding pixels and 1000 non-bleeding pixels. The pixels in each video frame are labeled by the perceptron, then morphological erosion is applied to regularize the region. The authors report a sensitivity and accuracy of 94%, while the specificity was up to 95% for classification of frames presenting bleeding regions.

Later in 2014 [Fu+14], the authors extend their approach by working with superpixel regions, and an support vector machines (SVM) classifier trained on 60,000 pixels. To construct superpixel regions, the authors use the algorithm in [Ach+12a], with color similarity calculated by the Euclidean distance in the CIElab space, and spatial distance between pixels. For final decision/classification, from each region, three color ratios (R/G,R/B,R/R+G+B, the same ratios of their previous work) are fed into an SVM classifier. The proposed method provides 99% sensitivity, 94% specificity and 95% of accuracy.

In the same year, Ghosh et al [Gho+14a] proposed to use the R/G histogram, which is less dependent to illumination, then apply k-Nearest Neighbors (KNN). They extracted several statistical parameters like mean, mode, variance, kurtosis, maxima, minima, skewness, then tested the performance of several combinations of the computed parameters with the KNN classifier using 200 WCE images extracted from a publicly available database [Pub]. The authors report that the combination of only three parameters, i.e. {median, variance, kurtosis} is sufficient to identify bleeding frames with an accuracy of 98.5%.

In [Gho+18], Ghosh et al. extend their work on the R/G color histogram. First, every image is decomposed into 7×7 pixels blocks and 6 statistical parameters are extracted (mean, median, max, min, value of the central pixel and its deviation from the mean of the block) from each block. The blocks are then separated into two classes with a KNN algorithm using annotations from the training data set. Each image is summarised by three numbers : the averages of the components of each centroid, and the difference between the two averages. To detect bleeding in colonoscopy video frames, an SVM classifier is trained on a set of 2350 WCE 3-number summaries with 10 cross validation. The authors report an overall area under the curve (AUC) value up to 0.989, while the specificity and accuracy were upper than 98%.

Some bleeding detection algorithms work on the histogram bin levels instead of the pixel intensity values [Gho+14a; KFR18].

Kundu et al. [KFR18] first compute regions of interest (ROI) defined by the color ratios $R/B > m$ and $R/G > n$. The parameters $m = 2.8$ and $n = 2$ are chosen according to the maximal accuracy of pixel detection compared to ground truth provided for 65 images, with pixel level mean accuracy (88.54%) and the lowest standard deviation (3.64%). The ROIs are then classified with a KNN classifier using the 64 histogram bins in the green channel. The method performance was of 97.86% for accuracy, 95.20% for sensitivity and 98.32% for specificity.

Other methods work on color histograms in the original RGB space, with more complex classification methods ([Gho+14b; SBW14; GFW18]).

In [Gho+14b], the authors propose to combine the RGB values into a single number with bit concatenation, and work on the resulting histogram. Every pixel is represented by 3L bits where L is the number of bits beginning from the Most Significant Bit plane (MSB). SVM classification on the 128 bins of the resulting histogram provides an accuracy of 94.5%, sensitivity of 93% and specificity of 94.88% for bleeding detection.

In [SBW14], the authors propose to classify regions into bleeding or non-bleeding based on 5 statistical features (mean, standard deviation, entropy, skew and energy) computed on each of the three RGB color channels. When applied to a WCE image, the frame is first segmented into large regions by reducing the color palette, then a trained two-layer neural network provides the classification label for each region. All 2^{15} combinations of features were evaluated, and the authors obtain 93% of accuracy, 96% for sensitivity and specificity 90% for the best combination, i.e. mean, energy and entropy in the red channel, mean in the green channel, and energy in the blue channel. The authors also propose a semi-automatic annotation algorithm based on region growing from manual seeds to obtain a large training data set.

In [GFW18], the authors propose to classify bleeding images with KNN applied to specific histogram bins of the RGB color space. After decomposing images into blocks such as in [Gho+18], they compute vector of features such as mean, median, maximum and minimum. To compute the classifier input (parameter/feature), they use the color histogram of 2^{3L} bins for each block, where L is the number of bits chosen for the color space. Due to the fact that blue color value superior to 128 is negligible in WCE images, they only consider blue color range (0 to 127). After, they apply the principal component analysis (PCA). On the other hand, to localise the bleeding zone, they add to features the intensity of central pixel for each block. The pixel are then labeled bleed/non-bleed according to their block to which it belongs. After, they apply morphological dilation followed by an erosion to delineate the suspected bleeding zones. According to reported results, with 16 color histogram bins in RGB, they classify correctly

bleeding frames with an overall performance of 99.15%, 99.47% and 97.85% in term of accuracy, specificity, and sensitivity respectively. They localise bleeding zones with a precision of 95.33% in 70 tested images.

Furthermore/More recently (2018), researchers of approach called CHOBS presented in [GFW18] proposed to extract statistical parameters from a block of pixels in the RGB color histogram and then classify the images with the KNN algorithm. They segment each image into pixel blocks of 7×7 pixels to extract local statistical features. Bin frequencies of the color histogram are used to be the global feature. They proceed to reduce the features size by first taking only bin probability that contains specified range of blue color (chosen to be 0-127), resulting in 50% dimension reduction without missing relevant bleeding area's information and then to apply the principal component analysis (PCA). The final block feature for frame detection contains a list of block's mean in all the color plane. On the other hand, to localise bleeding zone, they construct five dimensional feature containing mean, median, min, max and the block pixel intensity value. The pixel are then labeled bleed/non-bleed according to their block to witch it belongs. After, they apply morphological dilation followed by an erosion to delineate the suspected bleeding zones. The authors reported that with 14% overlapping authorisation between two successive blocks and 16 color histogram bins (feature dimension), they classify correctly bleeding frames with an overall performance of 99.15%, 99.47% and 97.85% in terms of accuracy, specificity, and sensitivity respectively.

As the RGB space may not be suited to bleeding classification, some methods have considered other colorspaces such as YIQ ,HSV and HSI ([Gho+15; Dee+18; Pan+11]).

In the work of [Gho+15], the authors use the SVM classifier with statistical features computed from extracted ROIs in a projection of the Luma In-phase Quadrature (YIQ) color space. The ROI was chosen according to a threshold value Q = 0. In the composite coordinate plane (Y.I)/Q, they compute (mean, median, skewness, minima) from each pixel of the supposed ROI. In comparison to 50 original WCE bleeding images labeling, their algorithm provides an overlap of 96.89%. And, testing the proposed approach on a set of 1000 images, the authors report an accuracy of 93.9% and sensitivity of 93.5% outperforming the method presented in [Gho+14a].

Deeba et al. [Dee+18] merge two SVM classifiers built respectively from features extracted from RGB and HSV color histograms. Due to the fact that acute bleeding occurs in several consecutive frames, they construct supposed bleeding areas using region growing proposed by [SBW14] with automatic seed selection during the training phase. From each region (ROI), they compute five

statistical features (mean, standard deviation, entropy, skew, energy). Each SVM is trained by alternating between SVM parameter optimization, five-fold cross-validation and feature selection. The RGB SVM classifier obtains (accuracy=49.5%, specificity=32.4% and precision=33.4%), while the HSV classifier provides (accuracy =93.9%, specificity=94.9% and precision=85.7%). For each ROI, the largest distance from the SVM decision boundary is used to select the most reliable classifier. This combination of two color spaces improves sensitivity to 92% and specificity to 95%. Also, the method presents better accuracy (94.5%) but slightly higher computation complexity (0.2 seconds per frame) compared to the method of Sainju et al. [SBW14] (0.11 seconds per frame and 71.4% accuracy).

In [Pan+11], the authors used RGB and Hue-saturation-intensity (HSI) pixel intensities to construct a six dimensional feature vector. They train a three layered probabilistic neural network on a total of 768 530 pixels and report a sensitivity of 93.1% and specificity of 85.8%.

Recently, Pogorelov et al. [Pog+19] proposed to consider image texture besides color. They used RGB color features (R/G, R/B, R/R+G+B ...) and 22 texture parameters (entropy, contrast, dissimilarity, ...) extracted from the grey-level co-occurrence matrix (GLCM) used for input for pixel-wise classification. Among the variety of machine-learning approaches tested (Random Forest (RT), Random Tree (RT), ZeroR ...), the output result for pixels detection was the highest for the SVM classifier with specificity (95.55%), accuracy (97.77%), sensitivity (97.76%), F1 score (97.80%) and ROC (99.7%). Images with more than 280 bleeding pixels are classified bleeding frames like [Sum+17a]. The proposed method outperforms that presented in [Fu+14] in term of accuracy (97.6% against 95%). According to their results, the color ratio coefficient Red over Green (R/G) showed the highest capability of detecting bleeding with a Matthews Coefficient Correlation (MCC) of 0.832.

6.3.2 State of the art for ulcer automatic detection methods

Ulcerative lesions show as white spots distributed on the gut wall, which explains why most methods focus on detecting bright pixels in various color spaces [Sum+15; Sum+17b; GDS16; Yua+15; KG15; YWT+14; Fan+18].

Multiple color spaces (RGB, YCbCr, CMYK or HSV) and their sub-spaces have shown successful gastrointestinal ulcer detection ability. Instead of using only the canals of a unique color space, some algorithms of ulcer detection include the use of combined color channels in order to improve their accuracy/performance.

In [Sum+15], the authors use a statistical analysis of ulcer pixel values against foreground pixel values in RGB and CIElab (Lab) spaces. For each color band, they compute the distribution of ulcer and non-ulcer pixels and determine the overlapping rate. Next, they train SVM on a set of 110 images (50 ulcerated/60 normal) of 576×576 pixel resolution and test a variety of combination bands information. They conclude that the feature set of dimension three, (L, a and G) gives best detection performance (sensitivity = 88.09% and specificity=90.88%).

Later, in the paper [Sum+17b], Suman et al. propose to extend their work of [Sum+15] by introducing multiple color spaces information/pixel intensities (RGB, HSV, YCbCr, CMYK, YUV, CIElab, XYZ). In order to find the best color canal for lesions detection they tested the 22 canals and find that Cr (red chrominance) and labA give the minimum overlapping rate between ulcerated and normal pixels. Learning/Training the SVM algorithm on 8000 endoscopic images (divided into 5000 ulcer/3000 normal), the best canals combination was of dimension three counting for the bands Cr from YCbCr, Y from CMYK and B from RGB reporting high performance ability (accuracy=97.89%, sensitivity=96% and specificity=95%) exceeding the use of HSV space alone like [KG15].

As mentioned above, ulcer pixels exhibit shadows of white or yellowish colors that are difficult to distinguish due to illumination variation during colonoscopy. In [GDS16], the authors propose a KNN classifier based on RGB histogram bins. They use a data set of 220 endoscopic image (110 ulcers/110 normal) to compute the optimal number of bins in the range $[8, 256]$. With 32 bins for each color plane, their method exhibits an accuracy of 87.27%, a sensitivity of 88.64% and specificity of 85.75%.

Ulcers are characterised by rough surface compared to the surrounding mucosa which makes which is why texture features may be suited for ulcer detection. Some studies [Yua+15; KG15; YWT+14] propose texture descriptors in order to localise ulcers.

In [Yua+15], the authors propose to combine color and texture features extracted from super-pixel regions with an SVM classifier. They compute the superpixels by SLIC algorithm ([Ach+12b]). They use the color channels S from HSV and M from CMYK due to their ability to highlight the ulcerative regions. They also use texture features obtained from Leung-Malik (LM) filters [LM01]. On a data set of 340 images extracted from 20 WCE videos, the authors report an accuracy of 92.65%, sensitivity of 94.12% and specificity of 91.18%.

In [KG15], the authors use textural features as input to an SVM classifier. They propose to extract texture information from color spaces RGB, HSV and YCbCr and compute statistical moments (mean, standard deviation, skewness) of the Contourlet transform and Log Gabor filter. With a data set of 137 images, they

show that the HSV and YCbCr color spaces provide the highest accuracy (96.16%) and specificity (95.83%) respectively.

Yeh et al. [YWT+14] explore different choices of features, feature reduction methods, classification methods to obtain the best overall classifier. Regions containing abnormal pixels are identified as connex regions via a color coherence threshold in RGB and HSV spaces, and textural parameters are obtained from the Grey Level Co-occurrence Matrix. Different combinations of saturation value, coherence threshold, number of features, feature selection algorithm and classification algorithm are compared. The authors report that the different choices of classification method and coherence threshold lead to similar performance. The best classification is obtained with decision trees. The best sensitivity (93.6%) for bleeding detection were obtained with the combination of saturation value 65, coherence threshold 5, 20 features and feature selection method ReliefF. For ulcers, best combination was of saturation value 70, coherence threshold 10, number of parameters 40 and feature selection method ReliefF.

The methods reported previousely require complicated steps for the learning stage which make their application/utilization very difficult for the detection of the UC lesions. On the other hand, many reported methods are based on the manual delineation of the lesions which is not the case in our database as mentioned in section 6.1.

In this chapter, we aim to detect the lesions with less complication in the detector to be practised in the medical practice. We want to evaluate the proposed performance criteria of Equation 6.9 and Equation 6.10 with the limited quality of annotations presented in our customized database, such a subject is present in other medical issues like colorectal cancer, another type of lesions...

Therefore, for ease of interpretation, we decide to use the colour intensities of the pixels in suitable colour spaces. To be continued

In the videos, the bleeding shows as red patches compared with the surrounding regions (cf Figure 6.2), whereas ulceration regions show as pinkish white colour area spreading on the gut wall (cf Figure 6.1).

Following [Gho+14a; GFW18; Gho+18], we consider the problem of labeling pixels into bleeding/non-bleeding based on the pixel value in RGB colorspace. As the previous authors have shown that the R/G ratio is relevant (it leads to 11% overlap between normal /abnormal pixels in [Gho+18]), we consider the set of linear classifiers in the (R,G) subspace.

The colour space YCbCr shows effectiveness for delimiting the ulcer lesions as it is shown on the example of endoscopy images in Figure 6.6. The Y canal (second column) shows the contours of the spatial distribution of ulcers whereas the canal Cr (third column) exhibits high-intensity values on the normal pixels and dark colour shadow on ulcers patches, which approve the idea of research for a linear relationship between these two colour bands to localise the ulcer lesions.

6.4 Proposed method

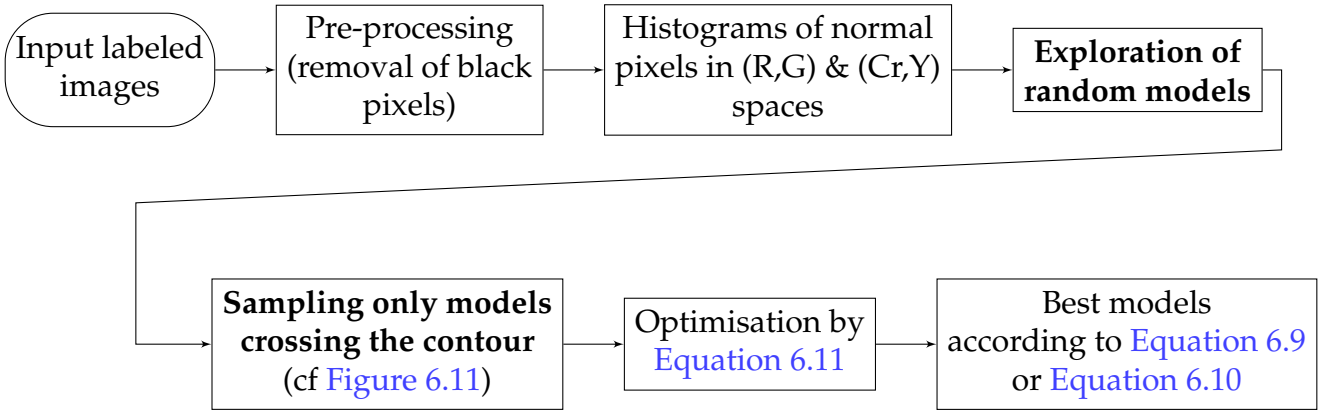


Figure 6.4: Flowchart of steps used for the computation of the bleeding and ulcers detectors in (R,G) and (Cr,Y) spaces respectively

6.4.1 Learning dataset

For the learning stage, we decide to use a database of 5 colonoscopy videos as detailed in the Table 6.1. The database contains both bleeding (1629 frames) and ulcer (1760 frames) annotations, for a total of 4349 frames.

	Bleeding	Ulcer	Total frame video
video 1	671	554	812
video 2	224	378	378
video 3	254	86	1116
video 4	140	204	910
video 5	340	538	1133
Total	1629	1760	4349

Table 6.1: Dataset used in the proposed approach

6.4.2 Pre-processing

Due to the camera's field of view, only an octagonal portion of the image is actually recorded in the endoscopic video, and the outer portions are set to

black. Additionally, some textual information is embedded in the video frames and should be removed prior to bleeding or ulcer detection. Both correspond to pixels with similar colours throughout the video. Consequently, we detect them as pixels with small grey-level variance and grow the detected region with morphological erosion to obtain a mask for each video as shown in figure 6.5.

Morphological erosion operation is used to remove the noisy spots of the image (or to discard isolated regions (very small)). Suppose A is the image after pixel classification and B is a structuring element. The erosion of A by B, denoted $A \ominus B$, is defined as follows:

$$A \ominus B = \{z|(B)_z \subseteq A\} \tag{6.7}$$

The erosion of A by B is the set of all points z such that B, translated by z, is contained in A. In our work, we use a $B = 5 \times 5$ structuring element with full of values 255.

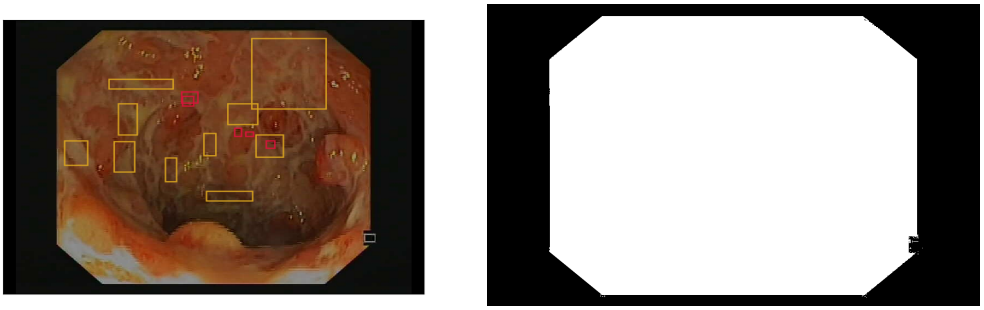


Figure 6.5: Endoscopic frames (left) and corresponding mask (right) used to remove pixels that do not correspond to the colonic wall

As the lighting shines on wet spots on the gut wall, some patches appear very bright but do not correspond to ulcers (see Figure 6.1). We remove them before ulcer detection the Y component, i.e. $\mathbb{1}_{\{Y>c\}}$. As these were not annotated, we chose the value $c = 150$ by visual inspection.

6.4.3 Detectors definition and exploration

6.4.3.1 Definition of the detectors

In the videos, the bleeding shows as red patches compared with the surrounding regions (cf Figure 6.2), whereas ulceration regions show as pinkish white colour area spreading on the gut wall (cf Figure 6.1).

Following [Gho+14a; GFW18; Gho+18], we consider the problem of labeling pixels into bleeding/non-bleeding based on the pixel value in RGB colorspace. As the previous authors have shown that the R/G ratio is relevant (it leads to

11% overlap between normal /abnormal pixels in [Gho+18]), we consider the set of linear classifiers in the (R,G) subspace, i.e. $\{aR + b > G \text{ for } (a, b) \in \mathbb{R}^2\}$.

The colour space YCbCr shows effectiveness for delimiting the ulcer lesions as it is shown on the example of endoscopy images in Figure 6.6. The Y canal (second column) shows the contours of the spatial distribution of ulcers whereas the canal Cr (third column) exhibits high-intensity values on the normal pixels and dark colour shadow on ulcers patches, which approve the idea of research for a linear relationship between these two colour bands to localise the ulcer lesions.

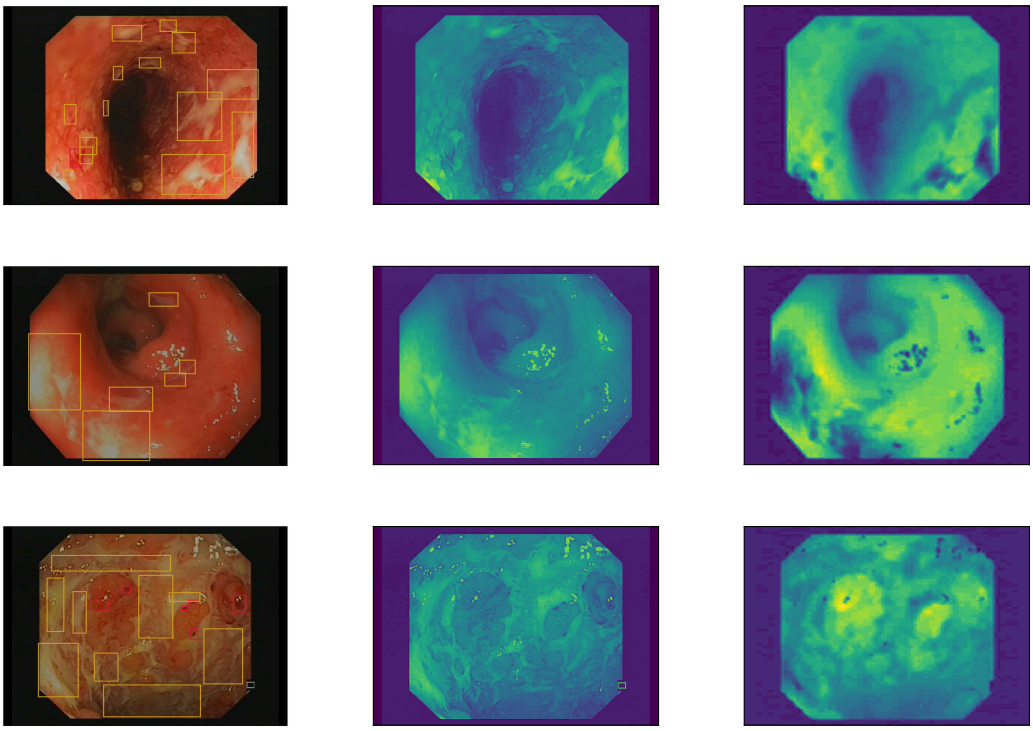


Figure 6.6: Examples of endoscopy images in our data base, From the left: original images, transformation in the Y and Cr respectively using Equation 6.8

As consequence, following the works of [Sum+17b; KG15], we consider the set of linear classifiers in the (Y, Cr) subspace of the YCbCr colorspace i.e. $\{aCr + b < Y \text{ for } (a, b) \in \mathbb{R}^2\}$.
As the videos are encoded in RGB, we apply the following transformation [Kum+] for ulcer detection:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.279 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.8}$$

6.4.3.2 Random sampling of the detectors

Prb1:

The first step for model research is to explore all the colour spaces (R,G) for bleeding and (Cr,Y) for ulcers. To count all the linear models in the space $[0:255] \times [0:255]$ requires a lot of time to run, then we decide to sample only one hundred models for each type of lesion as one can see in the following figures:

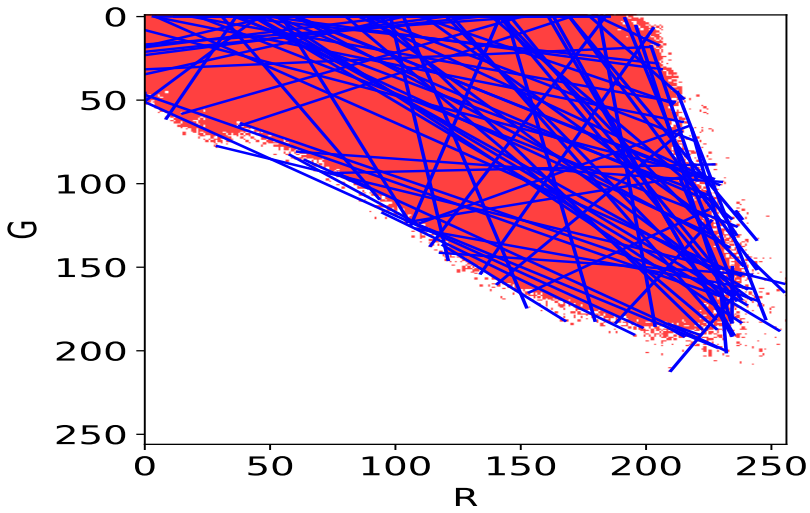


Figure 6.7: Random sampling of linear models in the color space (R,G) without histogram restriction to detect bleeding

According to Figure 6.9 and Figure 6.9, if a line i.e. model does not cross the histogram than any of normal pixel will be identified by the detector and then for this latter, the true negative rate will be zero. For example, the model ..

Prb2

If we plot the tested linear models and the histograms of all the pixels of the learning database, namely the histograms the normal and bleeding pixels in the (R,G) space and the histograms of the normal pixels and ulcer pixels in the (Cr,Y) color space, we obtain the figures
In fact, the blue lines correspond to models crossing the colour histograms of normal and abnormal pixels. The red models correspond to models that are disjointed with the histograms. These latter classify all the pixels, independently of their labelling class, into one class, normal or abnormal. Their performance will be decreased because of zero true positive rates or zero true negative rates.

This is undesirable because the overlap between the colour distributions of the bleeding class and non-bleeding class is unavoidable and some amount of

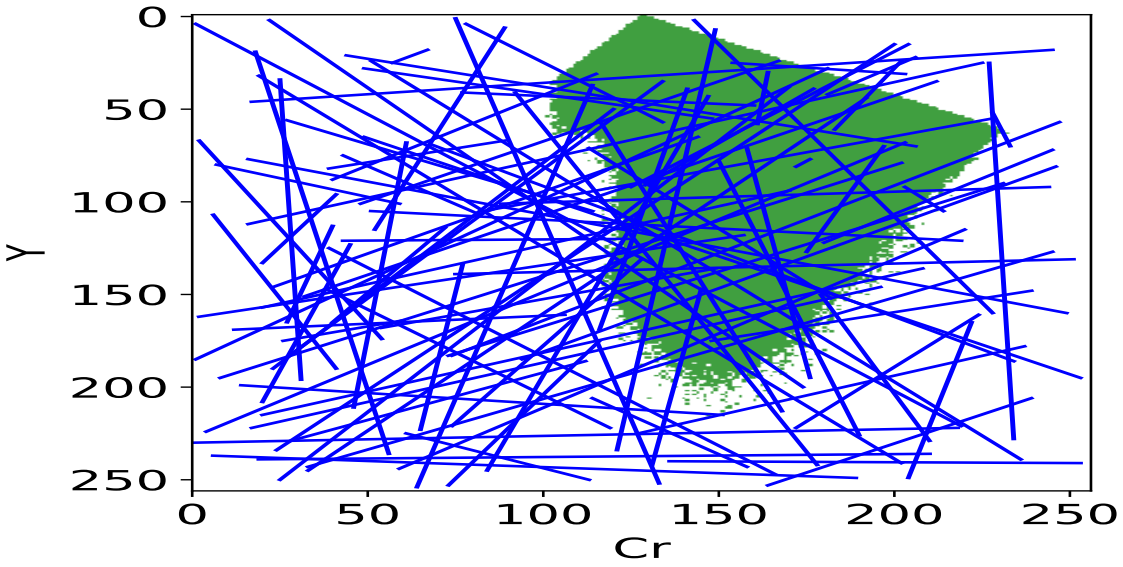


Figure 6.8: Random sampling of linear models in the color space (Cr,Y) without histogram restriction to detect ulcers

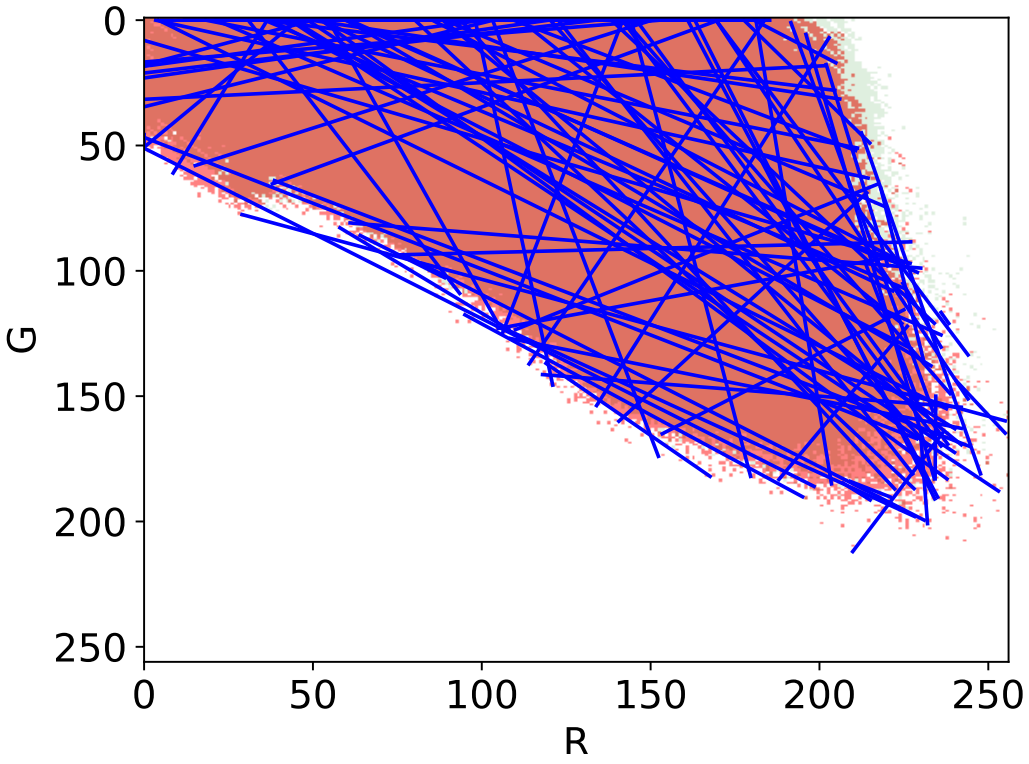


Figure 6.9: Without histo + bleeding

compromise is necessary. Consequently, we should restrict the optimisation space to the set of linear classifiers that go through the interior of the histogram.

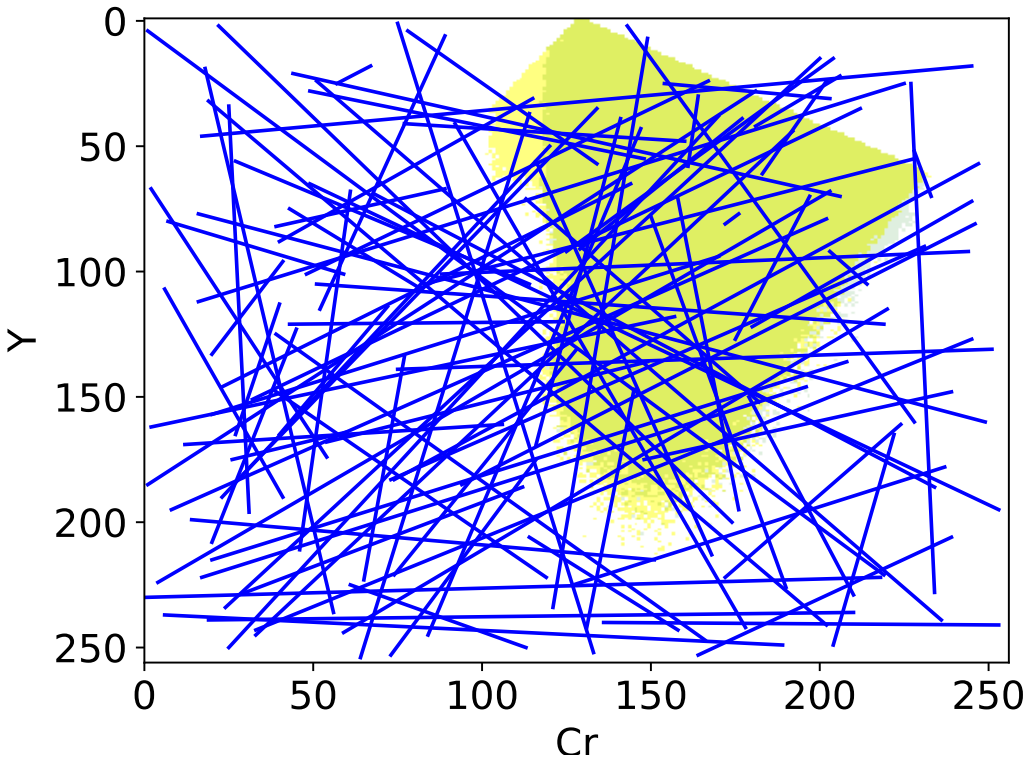


Figure 6.10: Without histo + ulcers

Additionally, if a line goes through the interior, it must cross the boundary of the set. We can avoid sampling redundant linear classifiers and speed up the computation time by focusing on the contour of the histogram instead of its interior. To sample the set of lines, we will thus draw two points in the contour of the (R,G) and (Cr,Y) histograms (represented in Figure 6.11) and consider the line that goes through these two points as the associated linear classifier.

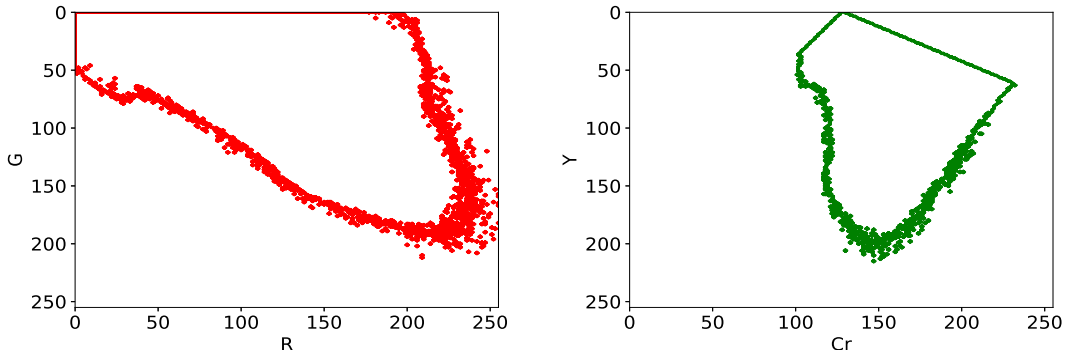


Figure 6.11: Contour of the (R,G) histogram of normal pixels (left) and (Cr,Y) histogram (right)

6.4.4 Modified Sensitivity

Due to the annotations errors in our database as discussed before, we expect over-inflated levels of FP and FN based on the database annotations. This will hide the correct classifier, and decrease the confidence in our results. Hence, evaluating specificity and sensitivity rely on our ability to correctly identify classification errors, that is to say on reliable pixel annotations by gastroenterologists.

To counter these problems, we propose to modify the definition of sensitivity in order to take the labeling issues into account. More precisely, we want pixels that were annotated by the gastroenterologist but not detected to count as true positives when they are near pixels detected by our algorithm.

Consequently we will count as true positives all pixels belonging to a gastroenterologist annotation as soon as one pixel is detected inside the annotation. Due to the fact that all pixels inside an annotation are either true positives or false negatives, this scheme corresponds to counting "detected annotations" instead of "detected pixels" in the definition of sensitivity. More precisely, an annotation is considered as detected if at least one pixel inside is detected. Alternatively, we can count "detected annotations" weighted by their surface or count.

Thus, we can denote two notions of sensitivity on annotation level:

$$\text{Sensitivity}^A = \frac{\text{Area of detected annotations}}{\text{Total area of annotations}} \tag{6.9}$$

$$\text{Sensitivity}^N = \frac{\text{Number of detected annotations}}{\text{Total number of annotations}} \tag{6.10}$$

The specificity criteria of Equation 6.2 was not modified, namely it is computed on pixels level, because we expect missing annotations to represent a small number of pixels relative to the number on non-annotated pixels $\sum TN + \sum FP$.

In the next section, we will present the optimisation procedure applied to find the best linear detectors.

6.4.5 Optimisation to find the best detector

Detector performance can be graphically represented with the aid of sensitivity vs (1-specificity) plot. As we are only interested in single detectors, each detector's performance is represented as a point in the Receiver Operating Characteristic (ROC) space, as opposed to a curve, which represents a family of detectors indexed by a threshold parameter. As we have proposed two forms of the modified sensitivity, we will present two figures, one with the criteria

"Sensitivity^A", and one with the criteria "Sensitivity^N".

The ideal classifier corresponds to the upper left corner or coordinate (0,1), with no false negatives (100% for sensitivity) and no false positives (100% specificity). Other good models are a compromise between sensitivity and specificity and are close to (0,1).

In this manuscript, we choose to select the classifier that is farthest from the line of no-discrimination, i.e. which has the largest euclidean distance from the diagonal in ROC space.

Consequently, our algorithm searches the optimal model $\hat{m}$ such that maximises the Youden index [You50] denoted by J and given by the following formula:

$$\begin{aligned} \hat{m} &= \underset{m}{\operatorname{argmax}}\ d_{\mathrm{ROC}}(\{y = x\}, m) \\ &= \underset{m}{\operatorname{argmax}}\ J(m) \\ &= \underset{m}{\operatorname{argmax}}\ \left(S(m) + \mathrm{Specificity}(m) - 1\right) \end{aligned} \tag{6.11}$$

where $S(m)$ denotes the sensitivity of classifier m (Sensitivity^A or Sensitivity^N).

As we detect the bleeding and ulcers separately, the optimisation process of Equation 6.11 will be performed twice, for bleeding and ulcers.

In the next section, we will present the results of the application of our algorithm on a set of 10 colonoscopy videos (5 used for the learning stage presented in Table 6.1 and 5 new videos).

6.5 Results of lesions's detection

6.5.1 Best bleeding detectors

We made random research for possible optimal linear models crossing the contours of the (R,G) histogram of normal pixels (figure 6.11) by choosing a sample of size 100 as explained previousely.

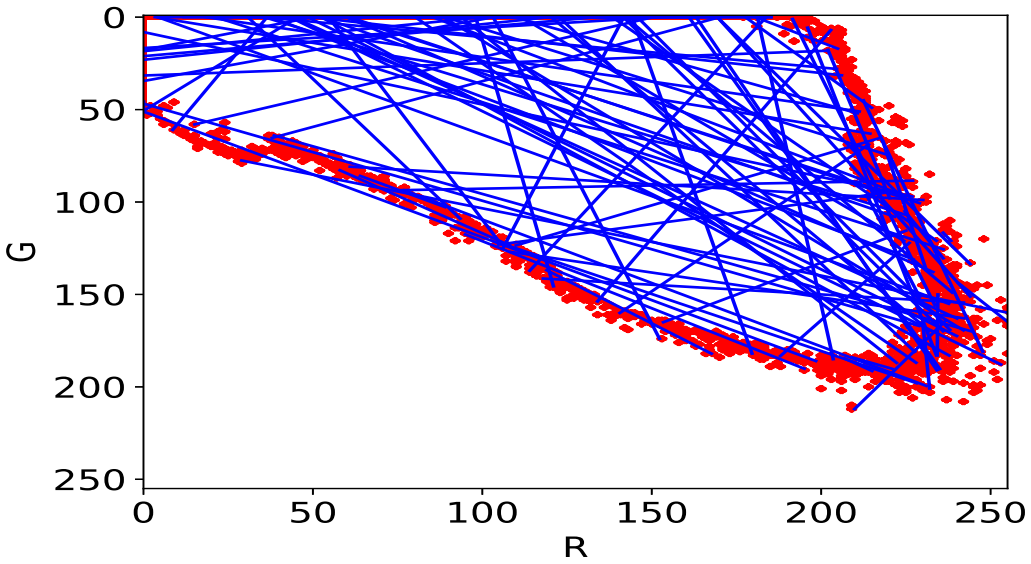


Figure 6.12: 100 linear classifiers are sampled by drawing two points from the contour of the (R,G) histogram of normal pixels

In the ROC space, we plot the S vs (1-specificity), S can be Sensitivity^A or Sensitivity^N, plots to visudalise the performance of the detectors:

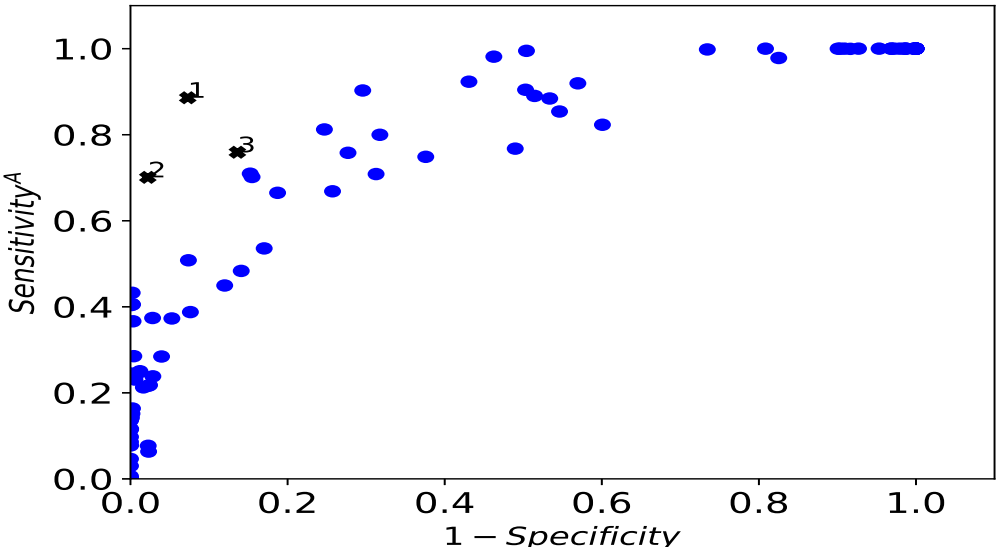


Figure 6.13: ROC space for bleeding by Sensitivity^A

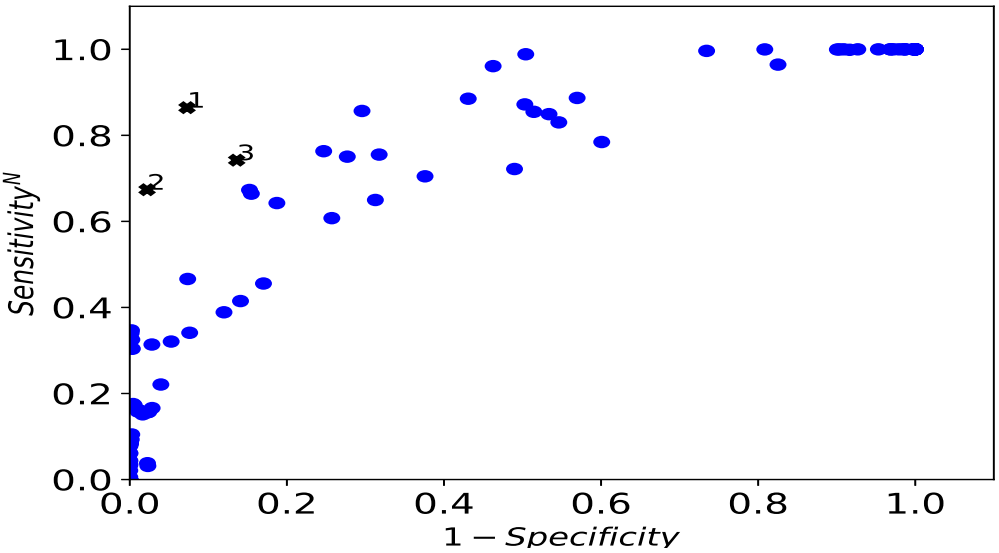


Figure 6.14: ROC space for bleeding by Sensitivity^N

Best models for bleeding	Specificity	Sensitivity^A	Sensitivity^N	Sensitivity
$G < 0.298R - 1.03$	92.71%	88.58%	86.44%	9.56%
$G < 0.264R - 4.837$	97.79%	70.08%	67.37%	4.16%
$G < -0.066R + 31.58$	86.39%	75.92%	74.23%	13.79%

Table 6.2: Performance of the best linear models for bleeding detection. Good performance is obtained based on Sensitivity^A or Sensitivity^N, but standard sensitivity is low due to annotation errors

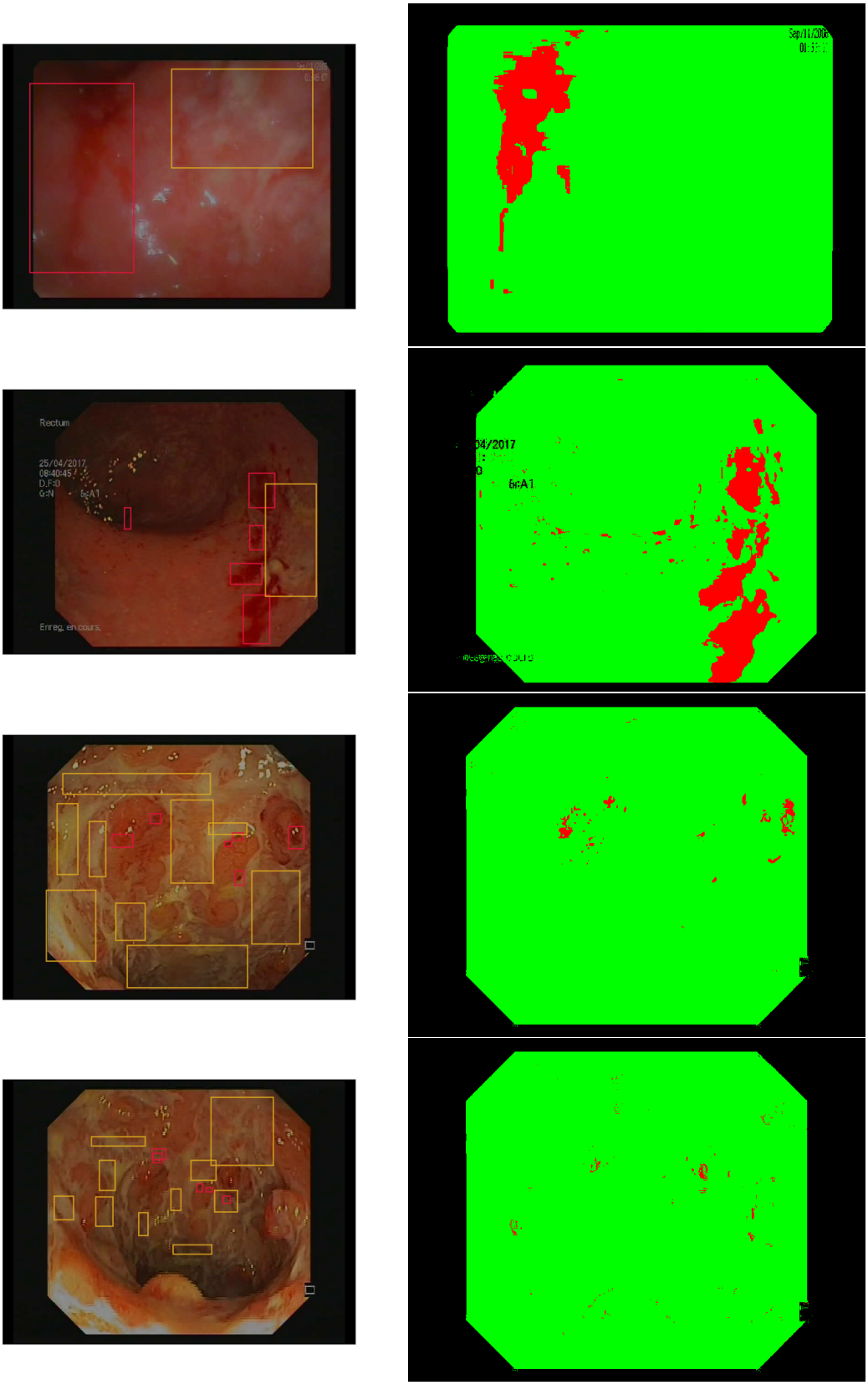


Figure 6.15: Annotated frame (left) and corresponding bleeding detection with the best linear models of Table 6.2

6.5.2 Best ulcer detectors

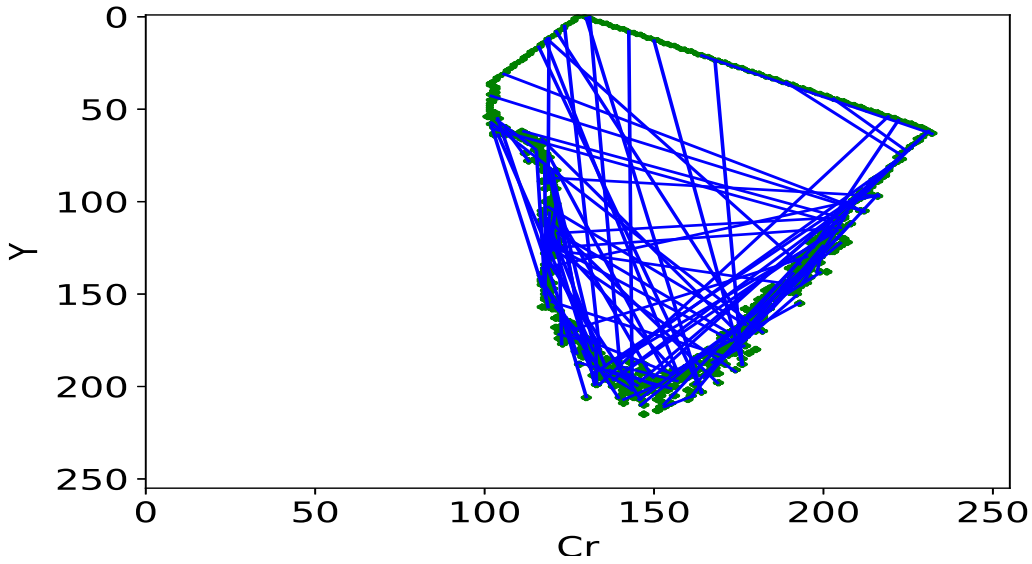


Figure 6.16: 100 linear classifiers are sampled by drawing two points from the contour of the (Cr,Y) histogram of normal pixels

In the ROC space, we plot the S vs (1-specificity), S can be Sensitivity^A or Sensitivity^N, plots to visudalise the performance of the detectors:

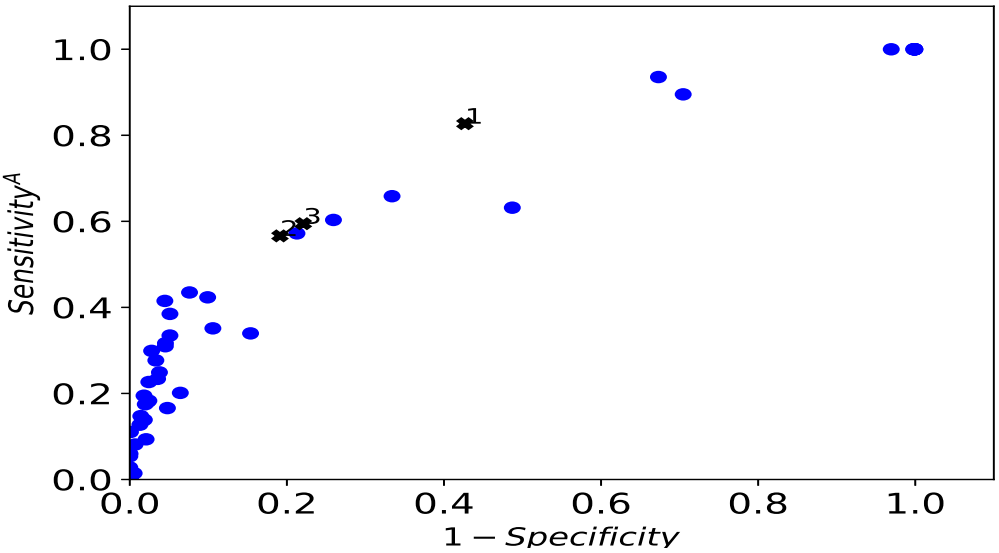


Figure 6.17: ROC space for ulcers by Sensitivity^A

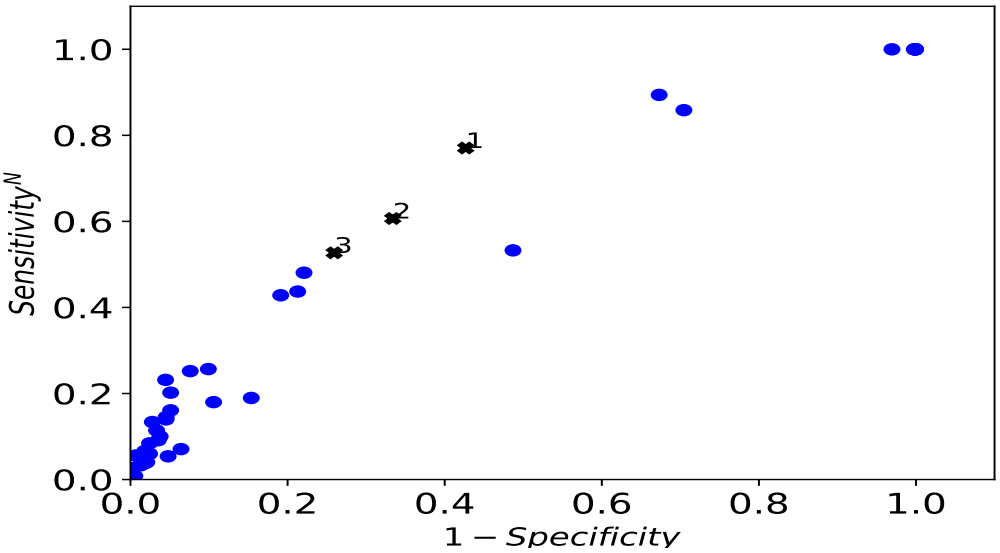


Figure 6.18: ROC space for ulcers by Sensitivity^N

As shown on Fig. **??**, the models achieve good performance results in ROC space, i.e. specificity and Sensitivity^A or specificity and Sensitivity^N. Fig. 6.19 shows that there is good visual agreement between the colors of detected lesions and the expert annotations. The best linear model can focus on the relevant areas, and select candidate ROIs that were not annotated. From figure **??**, we can deduce that unlike ulcers, the best models for bleeding detection are the same for Sensitivity^A and for Sensitivity^N.

Best models for ulcers	Specificity	Sensitivity^A	Sensitivity
$Y > 0.698Cr - 42.799$	57.33%	82.71%	38.85%
$Y > 0.505Cr + 8.816$	80.88%	56.62%	14.27%
$Y > 0.499Cr + 6.318$	77.91%	59.46%	17.17%
Best models for ulcers	Specificity	Sensitivity^N	Sensitivity
$Y > 0.698Cr - 42.799$	57.33%	77.06%	38.85%
$Y > 0.569Cr - 15.151$	66.61%	60.68%	28.69%
$Y > 1.633Cr - 181.444$	74.07%	52.7%	21.87%

Table 6.3: Performance of the best linear models for ulcer detection. Good performance is obtained based on Sensitivity^A or Sensitivity^N, but standard sensitivity is low due to annotation errors

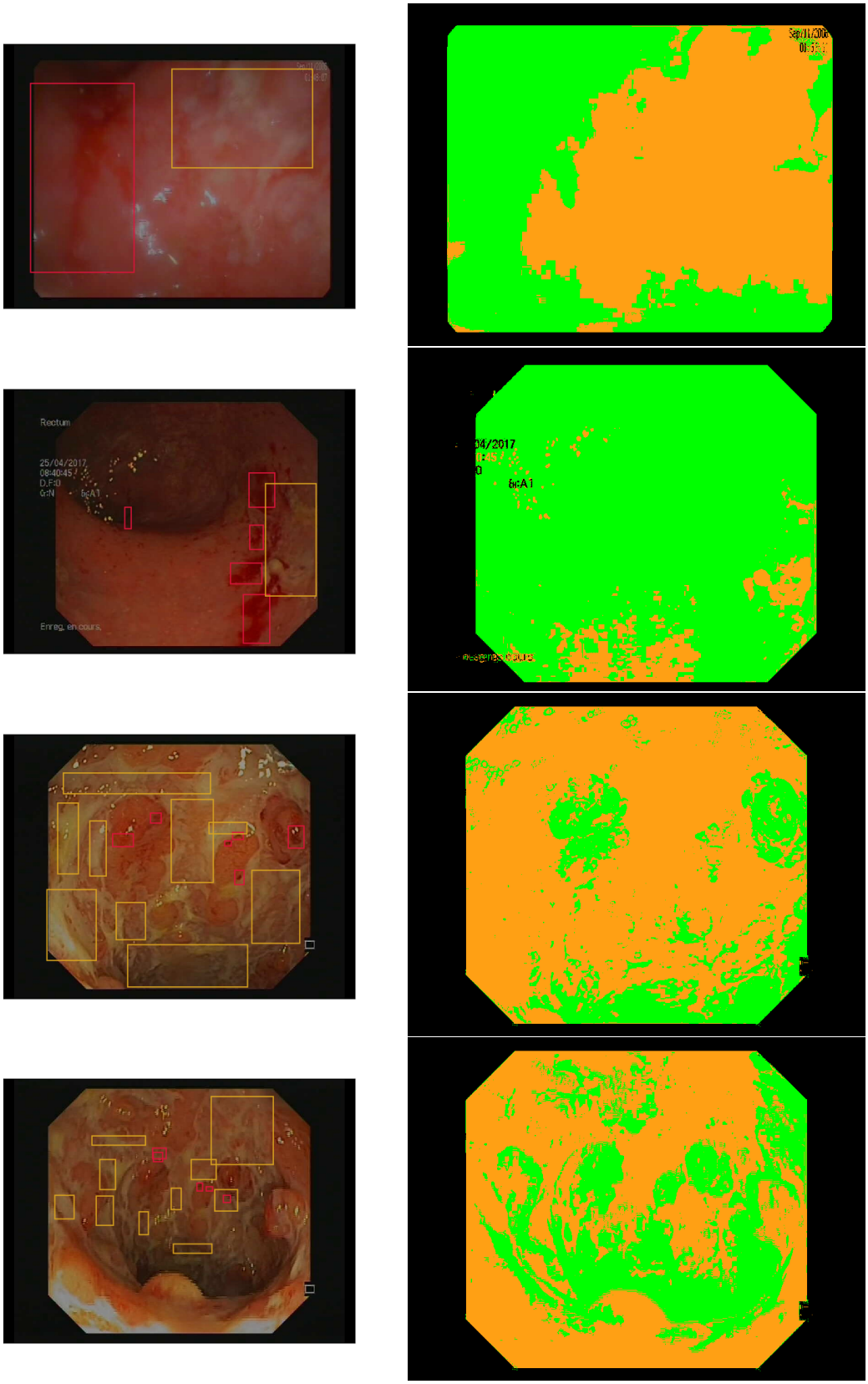


Figure 6.19: Annotated frame (left) and corresponding ulcer detection with the best linear models of Table 6.3

6.5.3 Variability of the performance of best detectors

In this section, we show the results of evaluating specificity and Sensitivity^A and Sensitivity^N on a random subset of frames in each video of a data base of 10 videos (5 training and 5 testing). We tested the performance in different cases of number/portion of frames or videos.

6.5.3.1 For 10 videos, 10% of frames and 20 random turn as paper

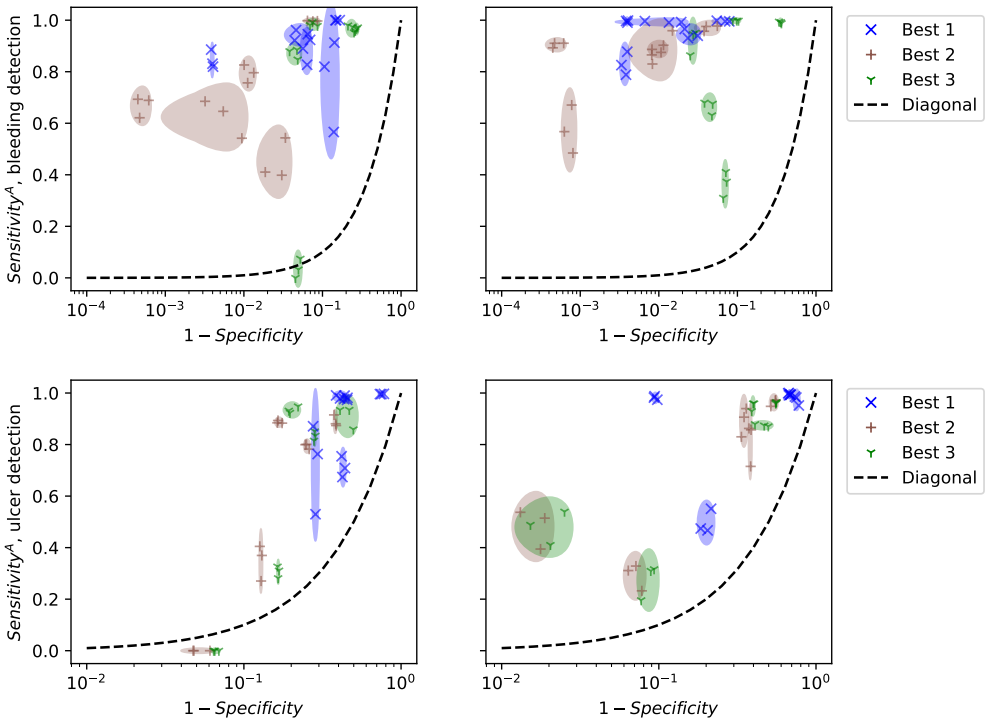


Figure 6.20: Performance of the 3 best linear models depending on patient, 5 training videos (left) and 5 test patients (right)

Fig. 6.20 shows the performance of the 3 best models for the patients in the training database (left) and 5 new patients (right). For each patient, we estimated specificity and Sensitivity^A on 20 random subsets containing 10% of the frames. Only three points are drawn, but the size of the ellipses are computed from the standard deviations of the 20 subsets. Fig. 6.20 shows that specificity and Sensitivity^A are estimated precisely, even on a fraction of the frames. This suggests that computational time can be reduced by using only a small subset of the video. However, the performance varies a lot between patients, even inside the training set. This means that the selected models are not universal and that specific models should be trained for each patient. This observation was not reported in previous works because the datasets used contain frames that are not organized "by patient". Consequently, significant methodological advances are

necessary to make colonoscopy videos comparable, in order to apply trained models to new patients.

6.5.3.2 Case of 80% of the 5 videos firstly used – Optional

In this section, we will make a randomised choice of data set, more precisely we will choose 80% of the totals amount of frames from the data set used for bleeding detection. Same procedure is applied for ulcers. Then, we will evaluate the performance of the three calculated best models in the section just below.

6.6 Conclusion

In this chapter, we presented a technique to automatically detect the bleeding and ulcer lesions found in colonoscopy videos. The lesions were annotated by rectangles using the Vatic software (cf section 6.1).

Consequently, the database contained many mislabeled pixels due to the complicated shape of a given annotation, not corresponding to rectangles in almost all of the cases. Accordingly, the detector's ability to correctly classify the pixels will not be computed precisely. Therefore, we decide to modify the standard sensitivity performance by a sensitivity that measures the detector's performance according to its ability to correctly detect the lesions instead of the pixels. We thus define two forms of sensitivity, the first one, Sensitivity^N counts the detected lesions weighted by their number, the second one, Sensitivity^A counts the detected lesions weighted by their area (see subsection 6.4.4).

According to the state of the art of automatic detection methods presented in section 6.3, we find that the colour parameter is effective to detect bleeding and ulcers lesions. We explored the set of linear classifiers in (R,G) and (Cr,Y) for bleeding and ulcer detection respectively. We also proposed in subsubsection 6.4.3.2 an efficient sampling scheme that samples only models with non-zero true negative rates.

The bleeding lesions were detected with 93% specificity, 86% Sensitivity^N and 89% Sensitivity^A whereas the performance for ulcer detection were 57% specificity, 77% Sensitivity^N and 83% Sensitivity^A.

We tested the variability of the performance of the best models on 10 colonoscopy videos in subsection 6.5.3 by using 20 subsets each containing 10% of the total frames. The performance is computed precisely for all the models on all the tested subsets (see Figure 6.20).

On the other hand, the best models can fail for some patients as one can see in Figure 6.20 for the two best models, with Sensitivity^A less than 20%). This

suggests that the models depend on the characteristics of the patient. There are many parameters including the mucosa type of the patient and its colour, the equipment used during the colonoscopy examination... Therefore, the appearance of the mucosa lesions should be corrected before proceeding with an automatic detection algorithm.

Chapter 7

Geometrical map of Ulcerative Colitis lesions

Abstract

The standard endoscopic scores such as UCEIS and MAYO-sub score do not take into consideration the spatial extent of the lesions in the colon.

This chapter is devoted to the localisation of the lesions regarding the colon. Therefore, we will use a colon's scheme to represent the bleeding and ulcer lesions found in a colonoscopy video.

In this work, we decide to represent the colon by its curvilinear abscissa. Assuming that the withdrawal speed of the colonoscope is constant, the frames are linearly related/distributed in the video. The inflammation is taken either from the annotations of the doctor or the results of the detectors computed in chapter 6.

We find that the proposed lesion scheme allows doctors to clearly see the distribution of lesions along the colon. In application to the Vatic database, it leads to a fine distinction between patients with the same score.

The proposed method is a first attempt at figuring all the colon lesions found in a colonoscopy video and differentiating between patients of the same class of disease. However, it needs several clinical validation and evaluation steps for the reproducibility

7.1 Motivation

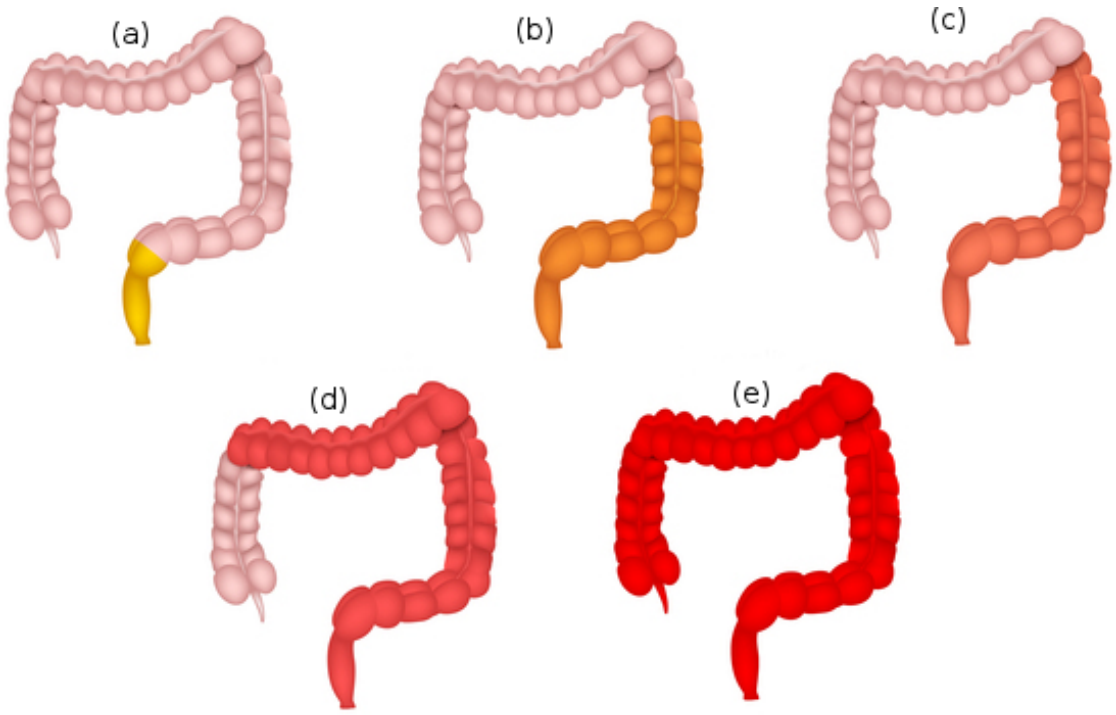


Figure 7.1: Main types of Ulcerative Colitis. (For details about the disease classes, we refer the reader to review section 1.2)

We have stated in previous chapters that Ulcerative Colitis disease manifests by inflammation beginning by the rectum and spreading continuously into the entire colon. Based on the affected segments of the colon, UC is classified into five classes ranging from proctitis, when the rectum is the only organ affected by inflammation, to pancolitis, when the entire colon is inflamed.

In Figure 7.1, we can see different stop-progressing points of the inflammation phenomena caused by the UC. It is the point of the large intestine at which the inflammation, inexplicably, stops. In Figure 7.1, the point is the limit of the coloured parts (yellow for the case (a), orange for the cases (b,c) and red for the cases (d,e)). From a medical perspective, the colon point at which the propagation of the disease stops progressing is until now inexplicable, and cannot be correctly identified by the video obtained after a colonoscopy exam.

The evaluation of the disease state using endoscopic scores such as UCEIS score and MAYO subscore (cf chapter 4) does not take into account the size of the lesions, their number and spatial distribution in the colon. Indeed, these scores only take into account the most severe lesion, regardless of the overall spatial extension of all the lesions on the mucosa. However, it has been recognized that the disease extension has a significant implication for the prognosis due to the association between the extensive UC and colectomy (cf chapter 2) and

colorectal cancer [D'h+07]. The risk of colon cancer is higher than normal if UC disease affects one part of the colon, and much higher for pancolitis patients.

To manage and treat large intestine conditions such as inflammatory bowel diseases, diverticulitis, colorectal cancer, bowel obstruction and uncontrolled bleeding, the doctors use a colectomy. The principle of this surgical operation is to remove the affected parts of the colon, which will be totally damaged by diseases in some cases. We have discussed the different types of colectomy (cf Figure 2.8). Consequently, to proceed with a colectomy operation, the doctor needs to know the affected parts of the colon to be able to decide the necessity or not and the type of special medications and surgical intervention as well.

In this chapter, we are interested in localising the inflammation of the colon, using colonoscopy videos. We can then provide to the doctors the global state of the lesions in the colon based on their spatial distribution.

The section 7.2 presents the proposed method to represent the localisation of the lesions in the colon. First of all, we discuss the colon parametrization considered in this work (cf subsection 7.2.1). This parametrization will be used in the following to establish a severity score (chapter 8), study the lesion distribution models (chapter 9), and the presence of propagation fronts (chapter 10). Second, to link the frame and the colon scheme, we assume some hypotheses that will be discussed in subsection 7.2.2. We compare the lesions spatial distribution between patients having equal endoscopic scores (UCEIS and MAYO sub-score) in the section 7.3. The application of the method to the Vatic database was made using the annotations of the doctors (cf subsection 7.3.1) or the detection with the linear models (cf subsection 7.3.2).

7.2 Methodology

7.2.1 Colon parametrization

We decide to represent the information about lesions in the colonoscopy video on the Figure 7.2 able to represent the colon anatomy needed in our work. It can represent all the parts of the colon. The cecum is represented by point A which corresponds to the curvilinear abscissa zero (beginning of the colon), the point G corresponds to the curvilinear abscissa 1 representing the end of the rectum (also the end of the colon). The remaining points (B to F) are chosen to link elbows angles of the large intestine.

These assumptions mean that if we choose another image to represent the colon, the number of points changes and thus their abscissas in respect to the colon. However, the result of the method will not be affected because it will be

adapted automatically to compute the frames abscissas.

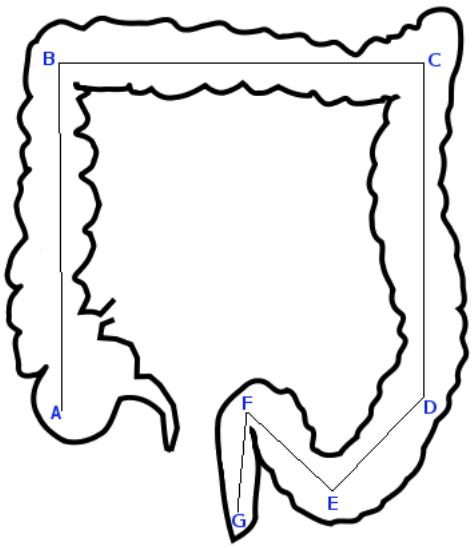


Figure 7.2: Colon curvilinear abscissa. The point A represents the cecum, the points B, C and E represent the turning angles of the colon, while G is placed at the endpoint of the colon (end of the rectum)

We integrate the coordinates of the curvilinear abscissa A to G of Figure 7.2 in a python code using the following chunks:

Code chunk 10: Localisation-lesions.py

```
# import of python librairies
import numpy as np
import cv2, io,pickle
from math import sqrt
```

Code chunk 11: Localisation-lesions.py (part 2)

```
# coordinates
A = (51.5, 306.5)
B = (49.5, 46.6)
C = (322.5, 46.5)
D = (322.5, 296.5)
E = (254.5, 366.5)
F = (189.5, 308.5)
G = (183.5, 382.4)
```

We use the parameter $s \in [0,1]$ to design the location on the colon and chosen to pass approximately by the centre of the colon. The coordinate of points A to G will be used to compute the coordinate of the frame regarding the colon. The coordinates of the points A to G to the curvilinear abscissa are given by the following chunk:

Code chunk 12: Localisation-lesions.py (part 3)

```
print "The ccordinates of the points A to G on the curvilinear abscissa are: \n"
print "absc_A= "+str(absc_A)
print "absc_B= "+str(absc_B)
print "absc_C= "+str(absc_C)
print "absc_D= "+str(absc_D)
print "absc_E= "+str(absc_E)
print "absc_F= "+str(absc_F)
print "absc_G= "+str(absc_G)
```

Interpret with `python2`

```
The ccordinates of the points A to G on the curvilinear abscissa are:

absc_A= 0
absc_B= 0.249489814901
absc_C= 0.511547173369
absc_D= 0.751526423266
absc_E= 0.845205667036
absc_F= 0.928828708498
absc_G= 1
```

In chapter 3, we have made a detailed review of the colonoscopy exam for the exploration of the intestinal wall. Roughly speaking, the endoscope as shown in Figure 3.8, is a long flexible tube that can be inserted into the colon, to arrive at the cecum. The endoscope is then withdrawn slowly from the rectum and the video images are recorded. During this work, we assume that the withdrawal speed of the endoscope is constant so that the curvilinear abscissa is linearly related to the frame number.

The algorithm used to produce the lesions cartography is formed of two main steps:

1. Computation of the two-dimensional coordinates of each frame of the colonoscopy video regarding the chosen colon image (Figure 7.2). The frame will be represented by point P.
2. Trace/Representation of the lesions data of the frame at the point P on the colon of Figure 7.2.

It will be normally a set of in-between steps to perform the results. We will explain them alternatively in the next sections. But first, we will expand the hypotheses that we considered along with this work and some drawbacks.

7.2.2 Some hypotheses

Some hypotheses were assumed during this work. We can list them as follows:

1. The colon is a two-dimensional organ
2. The speed of the endoscope removal during the colonoscopy exam is constant
3. The frame is related linearly to the colon curvilinear abscissa

4. turning angles of the colon [1] were considered placed like the other colon points

The colon is generally a three-dimensional tube/cylindrical like an object, for which the length and the diameter differ between humans. The colonoscope takes a picture alternatively with its path in the colon, this means that the picture of the colon/mucosa is taken into one direction and there are no available pictures for the other directions. As a consequence, in this thesis, we consider the colon a one-dimensional object (hypothesis 1) represented by its curvilinear abscissa.

The hypothesis 2 does not reflect exactly the medical exam. It is difficult for the gastroenterologist to retain a constant speed while maintaining the endoscope. All along with the exam, he needs to monitor the endoscope to supervise the visualization of the colon mucosa. However, sometimes he will stop moving the endoscope to take some shots for suspicious abnormalities at the level of the intestinal wall. Furthermore, to confirm his diagnosis about some previous suppositions, he proceeds to a biopsies sample removal. This leads to some break-times or brutal movements in the video progress. Hence, there will be a lot of images artefacts/errors in the video, because these images do not correspond to new locations or spots of the colon, but they are mostly the colon parts that were previously shown in the video. Therefore, when we have considered that the frames are linearly distributed according to the video's time, we have indirectly neglected these artefacts and errors.

The hypothesis 3 of the linearity relation between the frame number and the colon curvilinear abscissa is very simple. The frame number should be adapted to the colon length and the withdrawal speed of the colonoscope.

The total length of the colon varies depending on a variety of parameters such as gender, height age. Many reports have been published, but no one has yet explained the difference in length. We will pick a paper published in 2019 to extract the information about the colon parts length. According to the publication of Gaur et al. [Gau+19], the large intestine measures about 123 cm to 152 cm ($\approx$ 5 feet) in length and it is divided into the following parts:

- the Cecum for about 6 to 9 cm
- the ascending colon measuring 20 cm to 25 cm long
- the transverse colon measuring 40 cm to 46 cm long
- the descending colon of 10 cm to 15 cm long
- the "S" shaped sigmoid colon which is approximately 35 cm to 45 cm long
- the rectum of 12 cm.

We compare the measurements provided by [Gau+19] to the colon dimensions considered in this work. According to Gaur [Gau+19], the segment composed of the cecum and the ascending colon represent approximately 21% or 22% of

[1] The turning angles of the colon are the right colic (hepatic), left colic (splenic) and sigmoid flexures

the colon length. According to what we used as figure anatomy of the colon (cf Figure 7.2), the cecum and the ascending colon represented by the segment AB correspond to $\approx$ 25% of the total colon length. The same problem occurs in the transverse colon, which counts for 30 to 33% of the colon ([Gau+19]), is represented by the segment BC, only takes 26% of the colon length. The difference in representing the colon by the segments formed by A to G is bigger for the descending segment which measures 0.06% to 0.08% of the colon in real life, but it counts for 24% of the colon length in our work. The sigmoid colon, measuring about 28% or 29% of the colon as indicated in the paper [Gau+19], represents only 18% of the colon in Figure 7.2, corresponding to the segment DF. These percentages mean that the time video is decomposed is far from the exact anatomy of the colon. So the linear relation as assumed by the hypothesis 3 is not the best choice but in this thesis, we didn't get in-depth for another mathematical relation.

There are three principal angles in the colon anatomy, which are called turning angles. The turning angles are generally located by the gastroenterologist during the exam but they are hardly detected by a computer science algorithm. They are the hepatic angle which is placed where the colon turns below the liver from the transverse colon, the splenic angle (as it is next to the spleen) is the tight bend between the transverse colon and the descending colon, and the sigmoid angle which connects the descending colon to the sigmoid colon.

Adopting the hypothesis 3 will badly affect the localization of these angles in the collected colonoscopy videos from the database Vatic. To counter this problem, we have proposed to our collaborators' gastroenterologists to locate the position of points B, C and E shown in Figure 7.2 by making a shot during the colonoscopy exam. Unfortunately, the obtained figures for the angles are either very blurred to the rapid movements of the gastroenterologist during the exam or very similar and very difficult to distinguish between them using image processing algorithms. Hence, assuming the hypotheses 3 and 4 implies that locations of the turning angles are not accurate, but we can do better for now. We present in Figure 7.3, the images of the three splenic angles of the colon as captured by the gastroenterologist during the colonoscopy exam.

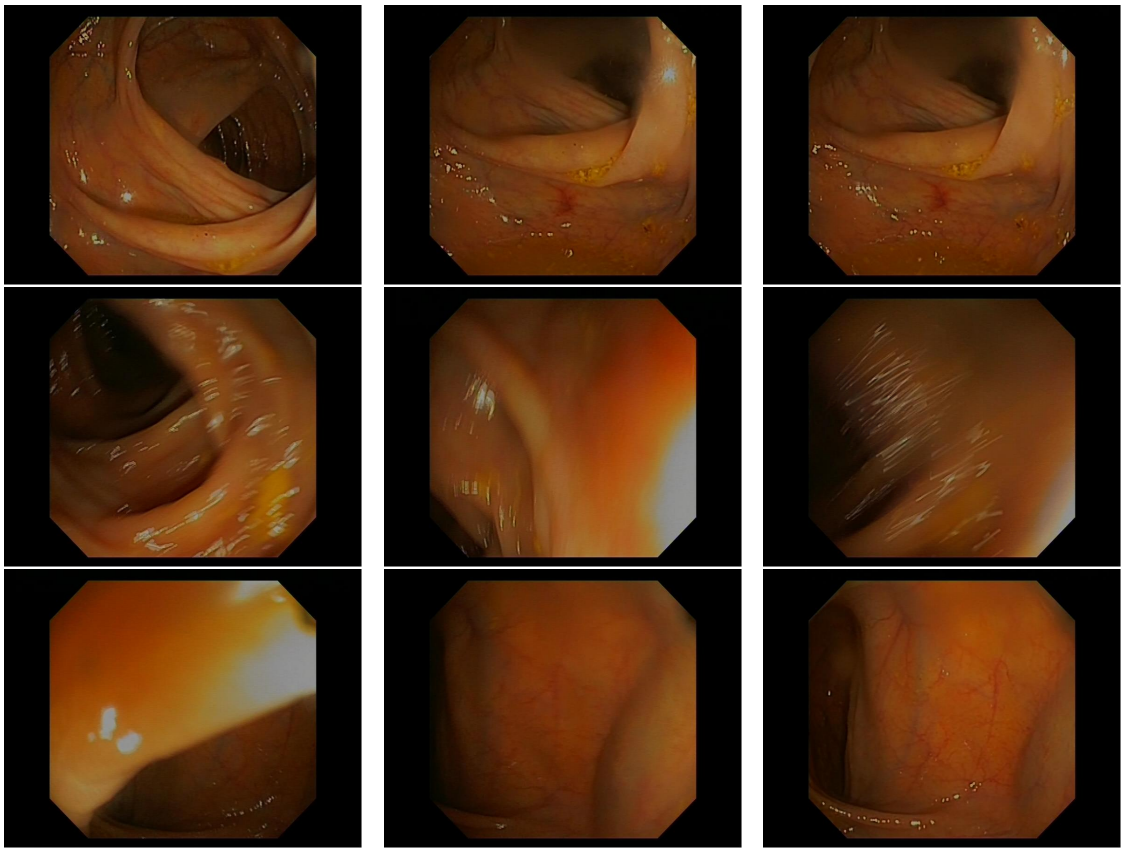


Figure 7.3: Shots of the colon turning angles

After the discussion made to present the effects of some assumptions/hypotheses considered in this work, we will give in the next section, details about the steps to do to get the visualization of the UC lesions in the colon.

7.2.3 Generation of the lesions map

7.2.3.1 Step 1: Computation of the frame point

To generate the map of the lesions, the first step to be ordered to find the position of the video frame according to the Figure 7.2. We will use the function `coord_frame` which takes as input the frame number and total video frames number. The output is the frame coordinates represented by the point P to be placed on the Figure 7.2.

Code chunk 13: Localisation-lesions.py (part 4)

```
# Computation of the frame coordinates
def coord_frame(num_frame, nframe):
    absc_frame = float(num_frame) / nframe
    if absc_frame >= absc_A and absc_frame <= absc_B:
        P = (A[0] + AB[0] * (absc_frame - absc_A) / (absc_B - absc_A),
             A[1] + AB[1] * (absc_frame - absc_A) / (absc_B - absc_A))
    if absc_frame >= absc_B and absc_frame <= absc_C:
        P = (B[0] + BC[0] * (absc_frame - absc_B) / (absc_C - absc_B),
             B[1] + BC[1] * (absc_frame - absc_B) / (absc_C - absc_B))
    if absc_frame >= absc_C and absc_frame <= absc_D:
        P = (C[0] + CD[0] * (absc_frame - absc_C) / (absc_D - absc_C),
             C[1] + CD[1] * (absc_frame - absc_C) / (absc_D - absc_C))
    if absc_frame >= absc_D and absc_frame <= absc_E:
        P = (D[0] + DE[0] * (absc_frame - absc_D) / (absc_E - absc_D),
             D[1] + DE[1] * (absc_frame - absc_D) / (absc_E - absc_D))
    if absc_frame >= absc_E and absc_frame <= absc_F:
        P = (E[0] + EF[0] * (absc_frame - absc_E) / (absc_F - absc_E),
             E[1] + EF[1] * (absc_frame - absc_E) / (absc_F - absc_E))
    if absc_frame >= absc_F and absc_frame <= absc_G:
        P = (F[0] + FG[0] * (absc_frame - absc_F) / (absc_G - absc_F),
             F[1] + FG[1] * (absc_frame - absc_F) / (absc_G - absc_F))
    return P
```

Interpret with `python2`

```
```

Let take examples:

Code chunk 14: Localisation-lesions.py (part 5)

```
coord_frame(45,200)
coord_frame(100,200)
```

Interpret with `python2`

```
(49.69631915564096, 72.11167427554193)
(310.47065517169597, 46.50440635341696)
```

7.2.3.2 Step 2: Representation of the data

The second step is to choose the data to be represented on the colon scheme. In this part of the work, We decide to represent the annotations provided by the doctors. The frame is consequently represented by the total surface of annotations regarding the image size.

We treated separately the bleeding and ulcers. The bleeding lesions will be shown in red color (cf function `Fct_local_bleeding`) while the ulcer lesions will be represented in green color (cf function `Fct_local_ulcer`). The form considered is the orthogonal lines to the colon curvilinear abscissa (cf function `vect_orthogonal`).

Code chunk 15: Localisation-lesions.py (part 6)

```
def vect_orthogonal(M,N):
  if N[0]!=M[0]:
    m=( (N[1])-(M[1]) )/( (N[0]) - (M[0]) )
    xo,yo=-m,1
  else:
    xo,yo=1,0

  n=sqrt(xo**2+yo**2)
  xo,yo=xo/n,1/n
  return xo,yo
```

Code chunk 16: Localisation-lesions.py (part 7)

```
def Fct_local_bleeding(img, abs_curv_P, P,rP):
    RR=(0, 0, 255)
    if rP!=0.0:# there is annotation in the frame
        rP=rP*100.0 # to the clear the visibility on the scheme
        if abs_curv_P >= absc_A and abs_curv_P <= absc_B:
            xo, yo = vect_orthogonal(A, B)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]-xo*rP),\
                           int(P[1]-yo*rP)),RR)
        if abs_curv_P >= absc_B and abs_curv_P <= absc_C:
            xo, yo = vect_orthogonal(B, C)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                           int(P[1]+yo*rP)),RR)
        if abs_curv_P >= absc_C and abs_curv_P <= absc_D:
            xo, yo = vect_orthogonal(C, D)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]-xo*rP),\
                           int(P[1]-yo*rP/100)),RR)
        if abs_curv_P >= absc_D and abs_curv_P <= absc_E:
            xo, yo = vect_orthogonal(D, E)
            cv2.line(img,(int(P[0]), int(P[1])),(int(P[0]-xo*rP),\
                           int(P[1]-yo*rP)),RR)
        if abs_curv_P >= absc_E and abs_curv_P <= absc_F:
            xo, yo = vect_orthogonal(E, F)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]-xo*rP),\
                           int(P[1]-yo*rP)), RR)
        if abs_curv_P >= absc_F and abs_curv_P <= absc_G:
            xo, yo = vect_orthogonal(F, G)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]-xo*rP),\
                           int(P[1]-yo*rP)),RR)
```

Code chunk 17: Localisation-lesions.py (part 8)

```
def Fct_local_ulcer(img, abs_curv_P, P,rP):
    GG=(39, 204, 6)# color green
    if rP!=0.0:
        rP=rP*100.0
        if abs_curv_P >= absc_A and abs_curv_P <= absc_B:
            xo, yo = vect_orthogonal(A, B)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                          int(P[1]+yo*rP)),GG)
        if abs_curv_P >= absc_B and abs_curv_P <= absc_C:
            xo, yo = vect_orthogonal(B, C)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]-xo*rP),\
                          int(P[1]-yo*rP)),GG)
        if abs_curv_P >= absc_C and abs_curv_P <= absc_D:
            xo, yo = vect_orthogonal(C, D)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                          int(P[1]+yo*rP/10)),GG)
        if abs_curv_P >= absc_D and abs_curv_P <= absc_E:
            xo, yo = vect_orthogonal(D, E)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                          int(P[1]+yo*rP)),GG)
        if abs_curv_P >= absc_E and abs_curv_P <= absc_F:
            xo, yo = vect_orthogonal(E, F)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                          int(P[1]+yo*rP)),GG)
        if abs_curv_P >= absc_F and abs_curv_P <= absc_G:
            xo, yo = vect_orthogonal(F, G)
            cv2.line(img,(int(P[0]),int(P[1])),(int(P[0]+xo*rP),\
                          int(P[1]+yo*rP)),GG)
```

We will use the data generated in chapter 5 about the surface of annotated lesions in every frame of the video.

Code chunk 18: Localisation-lesions.py (part 9)

```
# Use of saved results
with open('Folder-DataVatic-Percent.pickle','rb') as f:
        [y_b_percent,y_u_percent]= pickle.load(f)

# bleeding
Data_Perc_b = [ [individualArray] for individualArray in y_b_percent]
Data_Perc_b=Data_Perc_b[0][0]
# ulcer
Data_Perc_u = [ [individualArray] for individualArray in y_u_percent]
Data_Perc_u=Data_Perc_u[0][0]
```

To generate the lesions map for a given video using saved data, we launIn this sech the following code lines:

Code chunk 19: Localisation-lesions.py (part 10)

```
def Marque_Local_annotation_savdata(num_video,fichier_image_video,img_save,Db,Du):
    img = cv2.imread('/users/alali/Images-colon/Imagecolonbase.png')
    Nfmax = len(Db)
    for frame in range(0, Nfmax):
        P = coord_frame(float(frame), Nfmax)
        abs_curv_P = float(frame) / Nfmax
        Fct_local_ulcer(img, abs_curv_P, P, Du[frame])
        Fct_local_bleeding(img, abs_curv_P, P, Db[frame])

    cv2.imwrite(img_save, img)
```

Code chunk 20: Localisation-lesions.py (part 11)

```
# num_video = list of video numbers
for v in range(len(num_video)):
    img_save="/Folder-MapsDataVatic-Perc/DrAnnotation-Perc-vid-%s.png" %(num_video[v])
    Marque_Local_annotation_savdata(num_video[v],fichier_image_video[v],img_save,
    Data_Perc_b[v],Data_Perc_u[v])
```

In the following section, we will discuss the resulted lesions maps for patients affected by UC disease.

7.3 Results and discussion

In this section, we will present the maps of the spatial extent of bleeding and ulcer lesions for the colonoscopy videos from our database Vatic. We will compare the disease state for patients graduated/evaluated equally by the gastroenterologist (Dr. CS) according to the UCEIS score or the MAYO sub-score. In addition, we will compare the mucosa state for two UC patients having inflammation/lesions in the whole colon. The proposed technique gives further and more precise information than classic endoscopic scores about the state of the patient.

7.3.1 Maps of the doctor's annotations

We implemented the method evoked in section 7.2 to a set of 13 from the videos of the Vatic database for which we have the endoscopic scores, UCEIS and MAYO sub-score. First, we will compare only the maps for patients having the same endoscopic scores.

7.3.1.1 Patients with equal UCEIS score

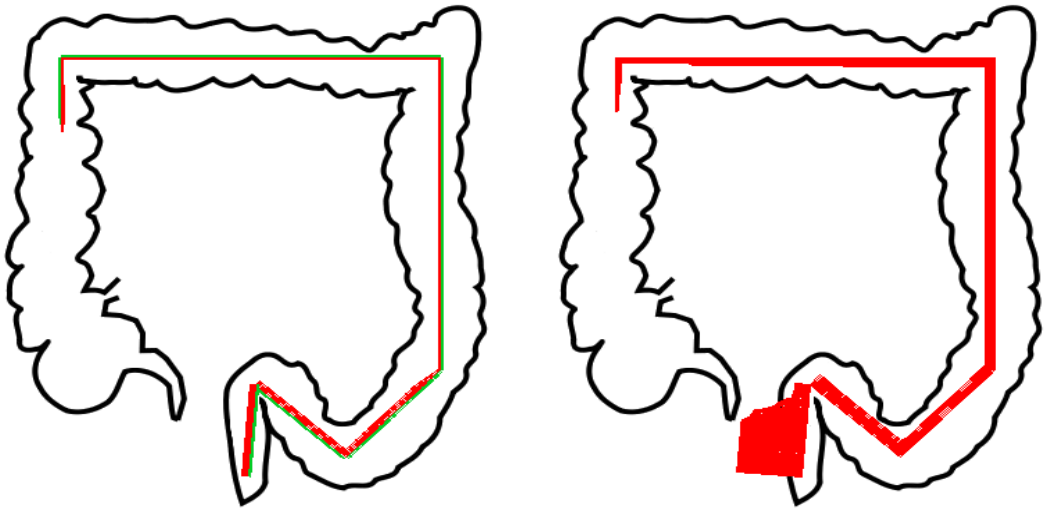


Figure 7.4: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with UCEIS = 2

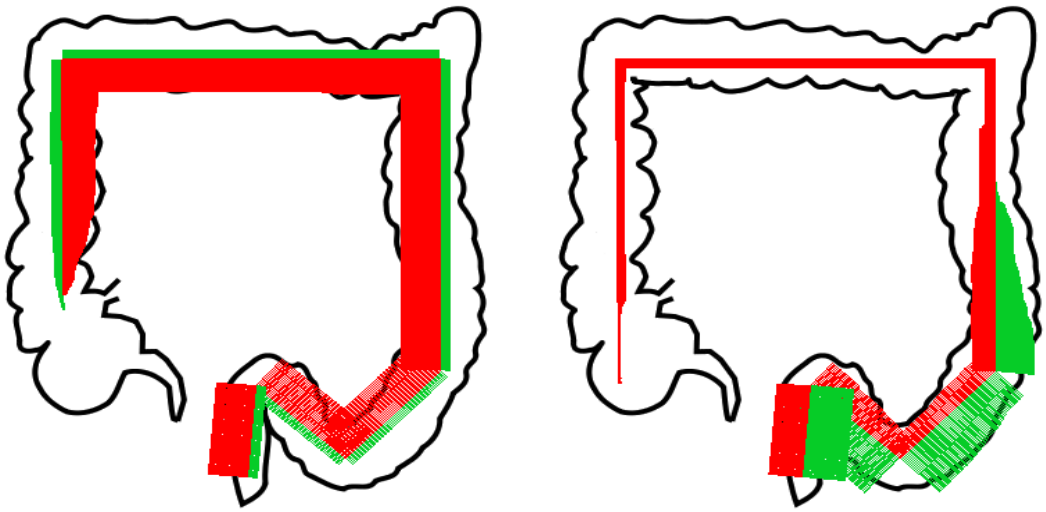


Figure 7.5: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with UCEIS = 4

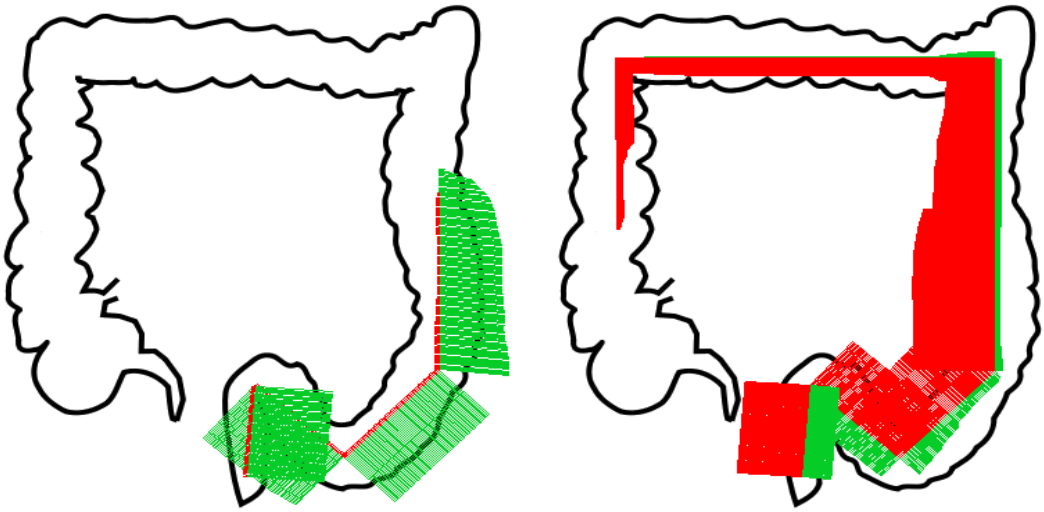


Figure 7.6: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with UCEIS = 5

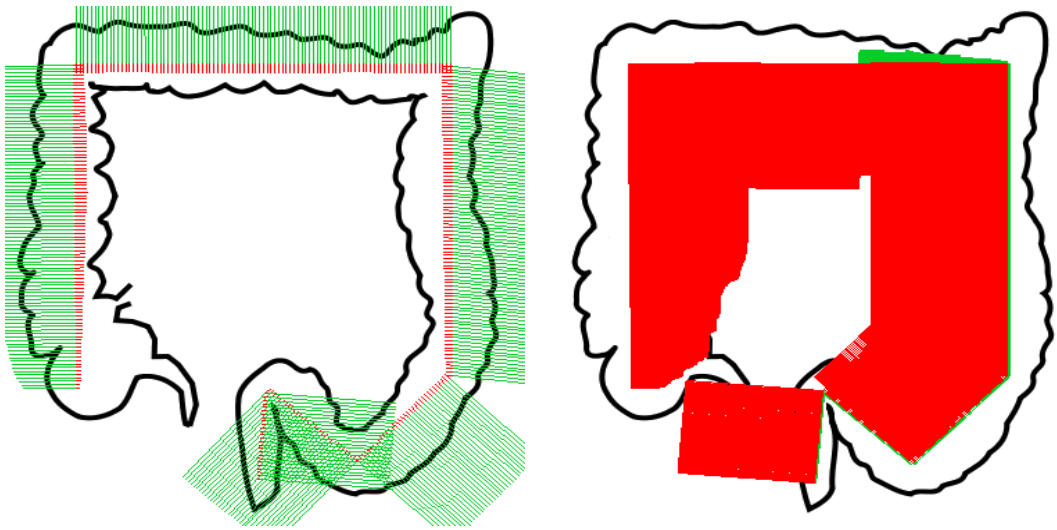


Figure 7.7: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with UCEIS = 6

The proposed method generates a new image of the behaviour of the disease. We remark relative differences about the spatial distribution and the severity of the lesions between patients having the same UCEIS score.

In Figure 7.4, the patients have approximately inflammation in the same colon segments, but the patient corresponding to the right image has the rectum bleeds more. The patients of Figure 7.5 to Figure 7.7, have differences in affected areas and type of lesions.

In Figure 7.6, the patient (at left) has a lot of ulcers from the rectum to the middle of the ascending colon, while the second patient (at right) presents

bleeding lesions on almost all the colon to the middle of the descending colon, the ulcerations are mostly concentrated at the rectum and sigmoid colon.

The disease behaviours for patients of Figure 7.7 are very dissimilar. The patient at left has ulcerated intestinal mucosa while the other one (at right) has the bleeding colon.

These inflammation appearances suggest different treatment strategies and different medications and hence different surgical interventions when needed.

7.3.1.2 Patients with equal MAYO sub-score

In this section, we will show the map of the lesions distribution in the colon for patients in the vatic database having the same MAYO sub-score

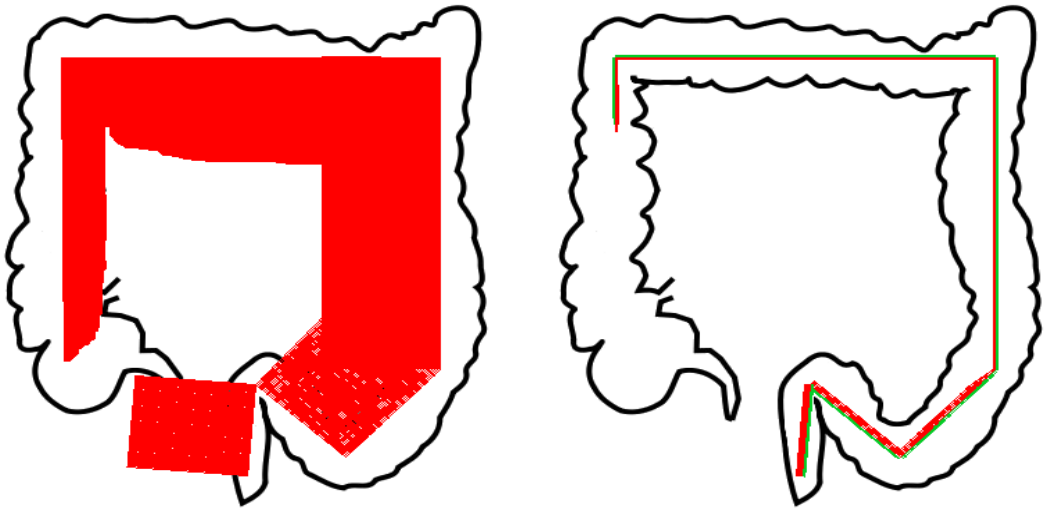


Figure 7.8: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with MAYO sub-score = 1

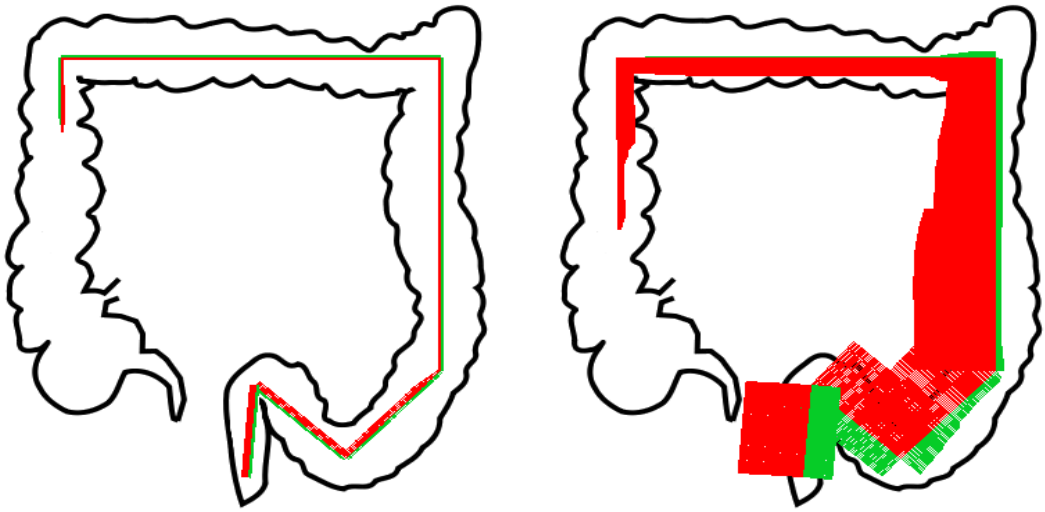


Figure 7.9: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with MAYO sub-score = 2

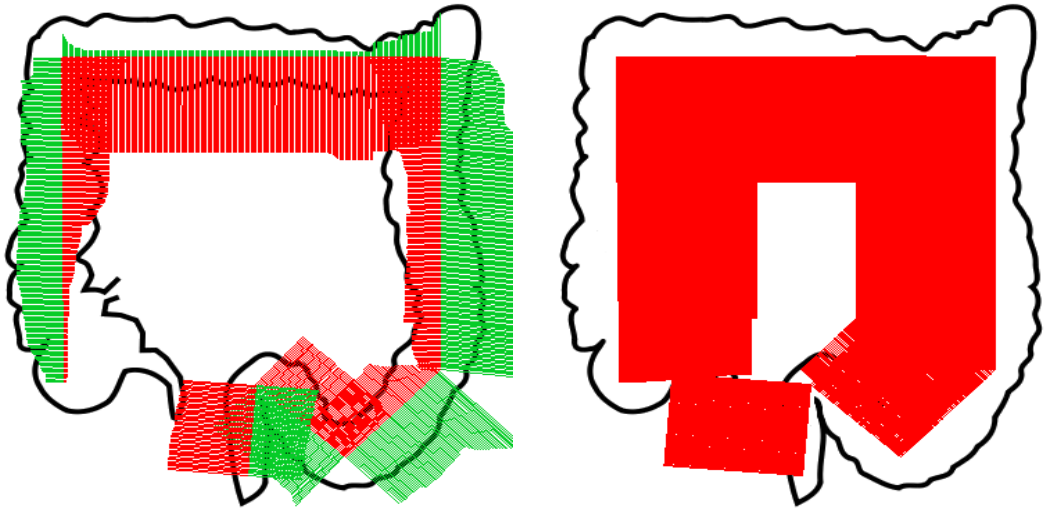


Figure 7.10: Distribution of the bleeding (in red) and ulcer (in green) lesions for patients with MAYO sub-score = 3

The same remarks are seen for patients having the same MAYO sub-score. The bleeding lesions are much more severe and affect more colon segments for the patient at the left than the patient at the right in Figure 7.8, for a MAYO sub-score equal to 1. We presented in chapter 3 the table of computation of the MAYO sub-score (cf Table 4.6), the score 1 indicates a partial loss of the vascular pattern visibility and friability. The bleeding state was not reported. Despite that we did not take these elements into consideration in the map presented here, we remark that the patient needs urgent intervention due to decreased state of the colon. So the score, as proposed, does not represent completely the severity of the disease. For patients of Figure 7.10, in spite of their classification

to the pancolitis UC class, the ulcer lesions are active for the patient at the left and absent for the patient at the right.

These observations lead to the proposition of stratification of patients in a given class of score, and the necessity of the integration of the spatial information to the UC diagnosis procedure.

In the next section, we will compute the lesions maps using the results of the algorithm of automatic detection presented in chapter 6. After, we will compare the results with those simulated according to the doctor's annotations as presented previously.

7.3.2 Maps of the linear models detection

7.3.2.1 Strategy of computation

We will use the optimal linear models computed in chapter 6 for bleeding ($\hat{m}_b$) and ulcer ($\hat{m}_u$) detection subject to the maximum of the criteria Sensitivity^A (cf Equation 6.9). We recall the parameters of the models:

- $\hat{m}_b : G < 0.298R - 1.03$
- $\hat{m}_u : Y > 0.698Cr - 42.799$

For each frame, we compute the count of the classified pixels (into bleeding/ulcers) regarding the total size of the frame. These pixels represent the detected abnormal areas using the models $\hat{m}_b$ and $\hat{m}_u$ for bleeding and ulcer respectively.

7.3.2.2 Results

In this section, we present the maps of the detected lesions to compare it with annotated lesions for the same patient.

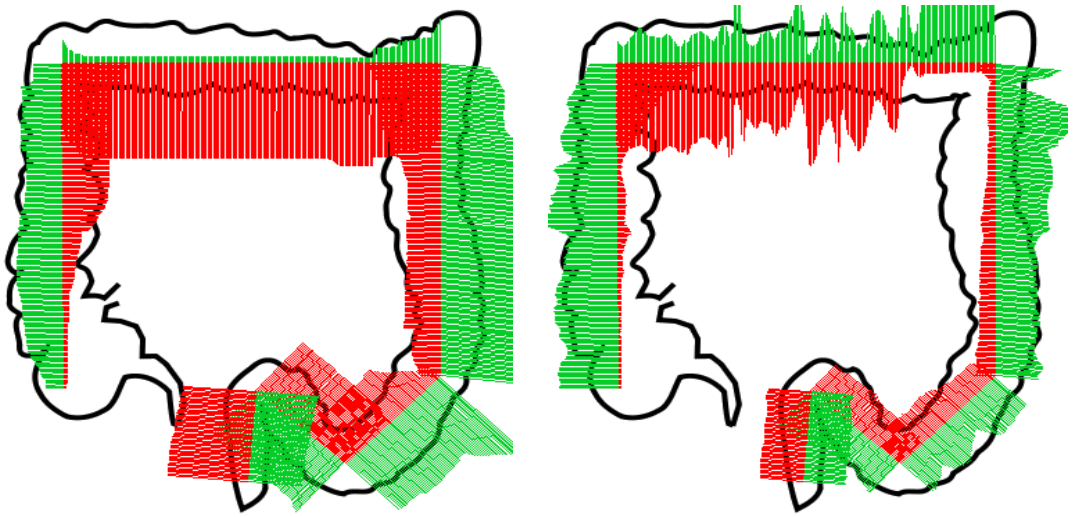


Figure 7.11: Distribution of the UC lesions for the same patient. Medical annotations on the left, and automatic detection (algorithm of chapter 6) on the right.

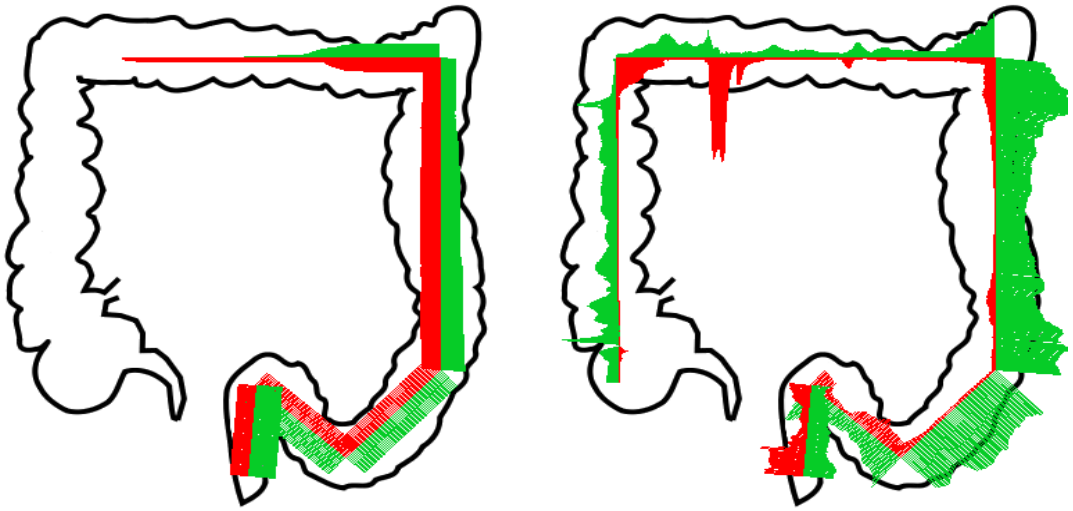


Figure 7.12: Distribution of the UC lesions for the same patient which UCEIS score is equal to 6. Medical annotations on the left, and automatic detection on the right.

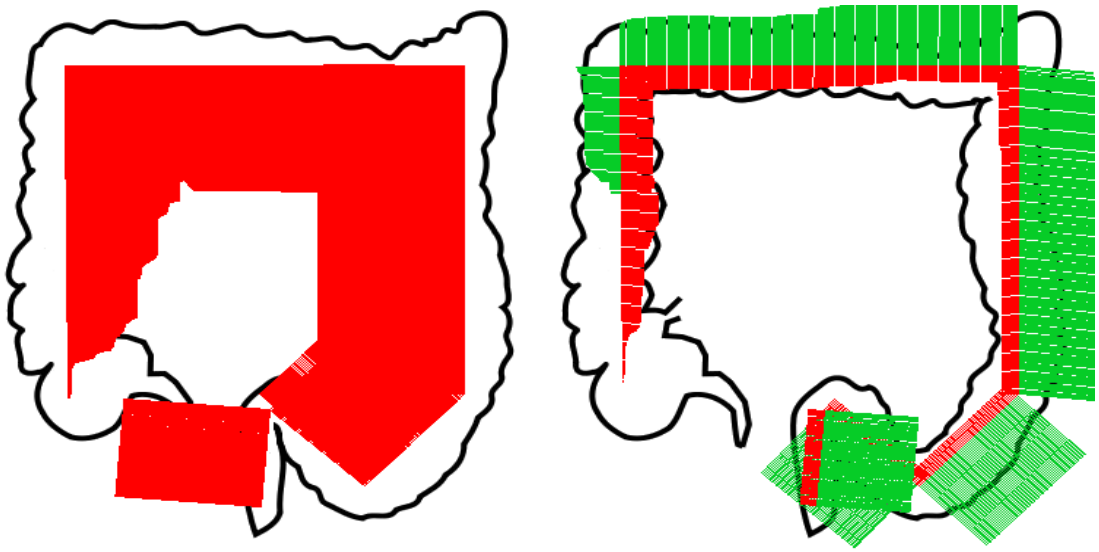
From Figure 7.11 and Figure 7.12, we conclude that the proposed algorithm and criteria Sensitivity^A in chapter 6 are almost effective to locate the abnormal pixels in a given colonoscopy video. However, the performance values reported for ulcer detection (cf Table 6.3) are remarkable in the case of a patient of Figure 7.12 when normal pixels were incorrectly classified as ulcers on the descending colon. These errors correspond to camera light reflection on the mucosa.

7.4 Limitations

Despite the effectiveness of the geometrical map that we presented in this work, it counts some defaults that we will explore in this section and propose some possible alternatives/solutions:

7.4.1 Video processing

The annotations of the doctors were not cleaned. In fact, they contain a lot of mislabeling errors as we discussed in chapter 6. We can remark that due to the appearance of a given lesion on a succession of frames, sometimes the length of the representative lines (red and green) was the same for some length on the colon as we can see for example in Figure 7.13a for bleeding lesions and in Figure 7.13b for ulcerations lesions.



(a) (b)

Figure 7.13: Distribution of the annotated bleeding (in red) and ulcer (in green) lesions for patients affected by UC disease. The colon parts from the descending colon to the rectum are affected identically by bleeding (on the left) and ulcer (on the right)

This demands the necessity of a video processing stage before computing the map in order to avoid the counting of the same lesion multiple times. A solution could be to put this information just at the frame where it was appearing for the first time. Another video processing is to study the frames taken multiple times in the video as in the cases where the doctor encounters some difficulties to move the endoscope during the colonoscopy.

7.4.2 Position of the turning angles

The use of orthogonal lines in the proposed method at the turning angles, especially at points B, C, D should be more precise to represent clearly the degree of the disease at these locations, we do not know if at these points the orthogonal type lines are the best choice. The orthogonal lines of the Figure 7.14 are crossing together implying an unclear view of the degree of the lesion at the locations of the turning angles/elbows.

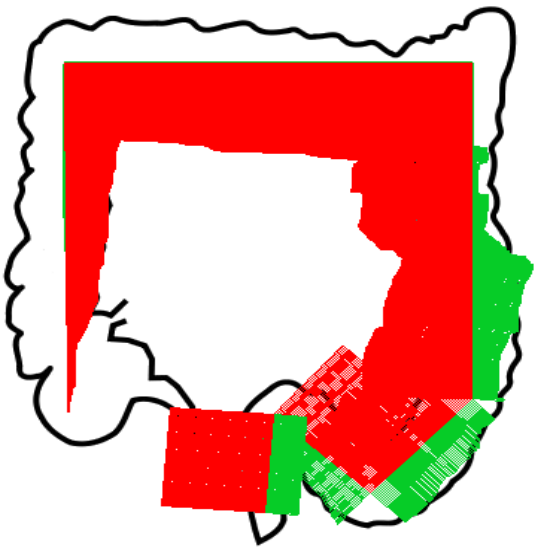


Figure 7.14: At the turning angle of curvilinear abscissa D, there is an overlapping between the red lines representing the bleeding at the surrounded parts leading to ambigous view of the state of the colon at this location

The same remark for the set of points E to G. In Figure 7.15, the problem of overlapping lines occur at the points E, F and G. The bleeding and ulcer surface are not represented in a good manner making their visibility quite difficult and imprecise.

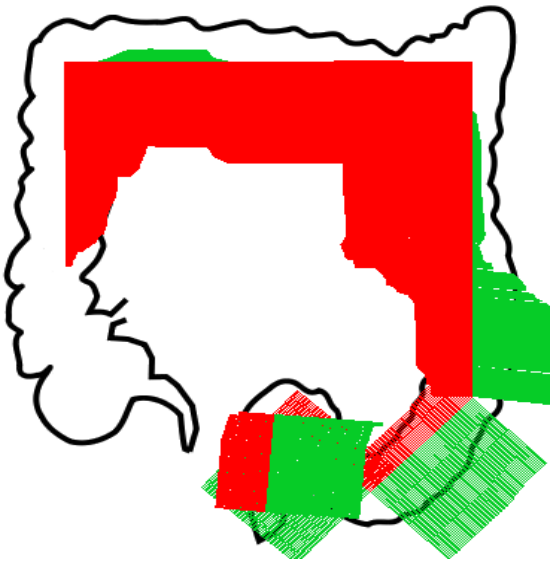


Figure 7.15: At the sigmoid colon, the lesions are not represented clearly. The bleeding on the segment connecting E to F are hidden by ulcers of the segment F to G.

7.4.3 Lack of clinical validation

This work requires also a clinical validation to be applied on new cases of UC patients. We have to study its reproducibility for the same patient. It has been demanded to the doctors, to make twice the colonoscopy exam for the patient but it is not yet realized because it is quite harmfull for the patient. Another procedure was to pass the endoscope double time (at the end of the colonoscopy) for the rectum, but the videos are not yet done.

7.4.4 Need of convenient performance criteria

To evaluate the capacity of the proposed method, we need to use some distance mesurements or performance criteria in order to finely/rigorously compare the spatial extent of the annotations/annotated lesions to the extent of the detected lesions.

7.5 Conclusion

In this chapter, the objective was to represent the lesions found in the colonoscopy video according to their location on the colon. Therefore, we presented the spatial distribution of bleeding and ulcers using a colon scheme (cf Figure 7.2).

We consider the colon by its curvilinear abscissa (cf subsection 7.2.1), the frames were supposed linearly distributed in the video (cf subsection 7.2.2). We link

the video to its position on the colon curvilinear abscissa using the linear interpolation.

We generate the map corresponding to the annotations provided by the doctors (cf subsection 7.3.1) and for the detections of the linear models (cf subsection 7.3.2) as well.

The obtained results demonstrate a new and more complete representation of the disease state that cannot be provided by the current diagnosis method i.e. endoscopic scores. It offers to doctors a clear representation of the spatial extent of the lesions and provides a fine distinction between patients with the same endoscopic severity score.

The information about the global lesions spatial extension, degree of inflammation and global mucosa state between patients have not been reported in previous works. Despite the fact that our work is just the beginning and needs some improvement due to some supposed hypothesis (cf subsection 7.2.2), it brings complementary information to the endoscopic scores about the disease severity. It offers to doctors the possibility of differentiating between patients of the same class of disease depending on the nature of the mucosa, and maybe leads to proposing new treatment procedures.

7.6 Perspective

The generated visualization is a new way to study the activity of the disease, and can complement the endoscopic scores used actually in medical practice. The visualization presented in this chapter, can be applied to the following up of the state of the patient and the efficiency of a given treatment. It can show if the disease continues propagating in the colon or it is receding meaning the adopted medical treatment procedure is effective. Otherwise, it should be changed.

Consequently, the work of this chapter can be extent to supervise UC patients over time (comparing maps of different colonoscopy) and controlling the adopted medications.

Another possible extension, is to use the lesion detection results of the intelligence artificial algorithms suggested in [Gho+18; GFW18; SBW14] for bleeding detection and in [Sum+17b; KG15] for ulcer detection in spite of the linear models detection here presented in subsection 7.3.2.

Chapter 8

Modelling the severity of the Ulcerative Colitis disease

Abstract

The severity scores such as UCEIS and MAYO subscore take into consideration only the most severe lesions in a colonoscopy video regardless of their spatial extension in the colon (chapter 4). The representation of the lesions generated in the chapter 7 suggests the necessity of fine grading for the severity assessment to improve the disease severity estimation and the long-time following-up. In this chapter, we propose to model the severity of UC considering the spatial distribution of the lesions.

For a first attempt, we investigated a linear model of the severity scores based on the spatial distribution of the lesions regarding their positions in the colon. To do that, we decide to investigate the set {Functions: patient $\implies$ $\mathbb{R}$}. The patient is represented by the count (or area) of annotations found in the colonoscopy video. The colon is parametrized similarly to chapter 7. For each segment of the colon, we inspected the lesion's count (and area) locally. The final score counts for all the segments.

On a set of 13 patients, we obtain good agreement between the proposed model and the scores provided by the gastroenterologists using the L_2 distance. In addition, we tested different colon configurations i.e. colon segments decomposition. For UCEIS, considering the colon by 6 segments is optimal, whereas the MAYO subscore is better estimated considering the colon into only 4 segments.

Although the modelling procedure was helpful to estimate the UCEIS and MAYO (sub) scores using the localisation of the lesions in the colon, it depends on the gastroenterologist scores. More, the proposed scores model does not involve objective biomarkers of the severity such as albumin and C-reactive protein.

8.1 Introduction

In chapter 4, we presented diverse endoscopic scores used to numerically assess the UC severity after undergoing a colonoscopy examination. The representation of the disease severity state of the patient is limited to a "small" set of parameters of type clinical, endoscopic and biological and sometimes the personal interpretation of the doctor as well.

The scores currently used in clinical practice, UCEIS and MAYO subscore are based on the evaluation of six (cf Table 4.6) and three (see Table 4.10) parameters respectively depending on the most relevant lesions of the UC disease. They took into count the bleeding, erosions/ulcerations, vascular pattern and mucosal erythema and friability. Likely to almost all the endoscopic indices, they don't consider all relevant information about the disease such as the spatial extension and the distribution of the lesions along the colon wall and their abundance as well.

In fact, in chapter 7, we have generated the geometrical distribution of the bleeding and ulcers lesion along the colon. We have also brought details about the inflammation's grade/ degree on each part of the colon. It was important to notice that the lesions are not uniformly distributed for all the colon segments. For instance, in Figure 7.5, in the representation at the right, bleeding lesions are spreading in the whole colon, unlike ulcers that are confined to the rectum and the descending colon. This fact is occurring in almost all the lesions profile for UC patients (cf Figure 7.6 and Figure 7.7). This phenomenon isn't considered during the computation of the severity indices. On the other hand, the severity score UCEIS was equal to patients having completely distinct lesions profile/spatial extension. Same for the Mayo subscore (see for instance Figure 7.8, Figure 7.9, Figure 7.10).

As a consequence, there is an imprecise evaluation of the disease severity without taking into account information about the lesions'spatial extent and their abundance. Later, it can affect the following-up of the disease course and the adaptation of proposed therapy strategies. Further, the risk of complications such as irreversible colon damage or colorectal cancer will increase.

Therefore, the endoscopic scores are almost incomplete to measure the state of the intestinal mucosa, and they need to encounter the spatial localisation of the lesions in the colon, in addition to the local state of their amount.

For all these reasons, we propose to model the severity of the UC disease by taking into consideration the local severity of the lesions (count or surface) regarding their position on the colon based on the parametrization expanded in chapter 7.

We propose to explore the set {Functions: patient $\Longrightarrow$ Severity $\in \mathbb{R}$} of type linear. We remind that bleeding and ulcer lesions are only considered in all the works of the thesis. The patient is thus represented by the list of bleeding and ulcers seen in the colonoscopy video. The amount of frames in a colonoscopy video is very big (the average video timing ranges from 6 to 12 minutes, with a rate of 30 images per second), therefore to have robust results and avoid over-learning, we must reduce the parameters of the model. To do this, we decide to consider some colon decompositions and consider locally the concentration of lesions represented by their count or area. The optimal colon decomposition is then computed through an optimisation problem using the L_2 distance. This work will be performed to the UCEIS and MAYO-sub-score.

The modelling technique will be helpful for the next chapter to study later the evolution of the disease severity with the amount the bleeding and ulcer lesions in the colonoscopy video. On the other hand, the gastroenterologist personal evaluation is a key parameter for the computation of the Sutherland [Sut+87] and MAYO [Lew+08] scores as one can see in Table 4.4 and Table 4.5 respectively. In addition, for all the endoscopic scores the evaluation of the lesions state based on colonoscopy video indirectly implies the personal point of view of the gastroenterologist. These elements make the scores not reproducible for the same gastroenterologist and between gastroenterologists evaluating the state of the same patient. Hereafter, we propose to study/measure the variability of severity scoring between the gastroenterologists.

The rest of the chapter is organized as follow. The section 8.2 encounters for the database of endoscopic scores, the tested colon decompositions to compute the severity scores and the mathematical model as well. In section 8.3, we present the obtained results by considering the count or the surface of the lesions, and the computation of the best colon decomposition as well.

8.2 Modelling setups

We will explore the linear models of the set: {Functions: patient $\Longrightarrow$ Severity}. For this process, we will use a set of 13 patients of Vatic database.

8.2.1 Representation of the patient

Due to the consideration of the spatial extent of lesions along the colon, the patient will not be represented by the most severe disease descriptors found in colonoscopy videos. Rather, we will consider locally the amount (or area) of the lesions in the colon. Mathematically, the patient is an element of $\mathbb{R}^2$, the first component designs the count of bleeding, the second component corresponds to the total count of ulcers. We obtain the following values for the considered 13 patients:

```
Bleeding annotations total count for 13 patients is:
[1946, 1101, 67, 2728, 1026, 674, 157, 673, 1081, 202, 8294, 766, 1574]
Ulcers annotations total count for 13 patients is:
[2332, 3888, 20, 0, 0, 92, 956, 905, 904, 0, 50, 1292, 0]

Bleeding annotations area for 13 patients is:
[0.43, 0.05, 0.02, 0.7, 0.08, 0.24, 0.01, 0.29, 0.08, 0.08, 0.92, 0.11, 0.99]
Ulcers annotations area for 13 patients is:
[0.37, 0.81, 0.02, 0.0, 0.0, 0.07, 0.2, 0.06, 0.09, 0.0, 0.02, 0.11, 0.0]
```

The main objective of this work is to count for the spatial distribution of UC lesions such as bleeding and ulcers to model the disease severity. The segments of the colon are therefore essential to localise the distribution of the lesions.

In this direction, we continue the use of the colon parametrization based on the curvilinear abscissa given in Figure 7.2. The spatial information about the lesions can be extracted according to their position (point-to-point) in the colon. Nevertheless, the number of the model parameters will be very huge and the model will be more complicated to be analyzed. Therefore, we decide to reduce the number of the parameters to the number of the colon segments. Afterwards, we compute the count (or surface) of the lesions on each segment. We decide to test some possible colon decompositions to find the best one.

Consequently, the patient will be considered in $\mathbb{R}^{2k}$, where k value is chosen to cover the principal segments of the colon of the Figure 7.2. Herein, we consider the maximum number of colon segments is $k = 6$ as one can see in Figure 8.1.

In Figure 8.1, we assign a specific colour for every segment to facilitate the differentiation between the considered decompositions. For instance, the colon is considered as two segments, we have considered the ascendant and transverse colon as one segment (denoted by a light blue colour), and the other colon parts as one segment (designed by a medium blue colour). For the case of the decomposition into $k = 3$, we considered the ascendant colon, also called the right colon, the transverse colon as two segments, and the colon parts as the third segment. For $k = 4$, we considered the principal colon segments associated with the anatomy of the colon. For the upper values of colon segments, namely for the cases of $k = 5$ and $k = 6$, we consider sub-parts of the sigmoid colon limited/confined by points D to G.

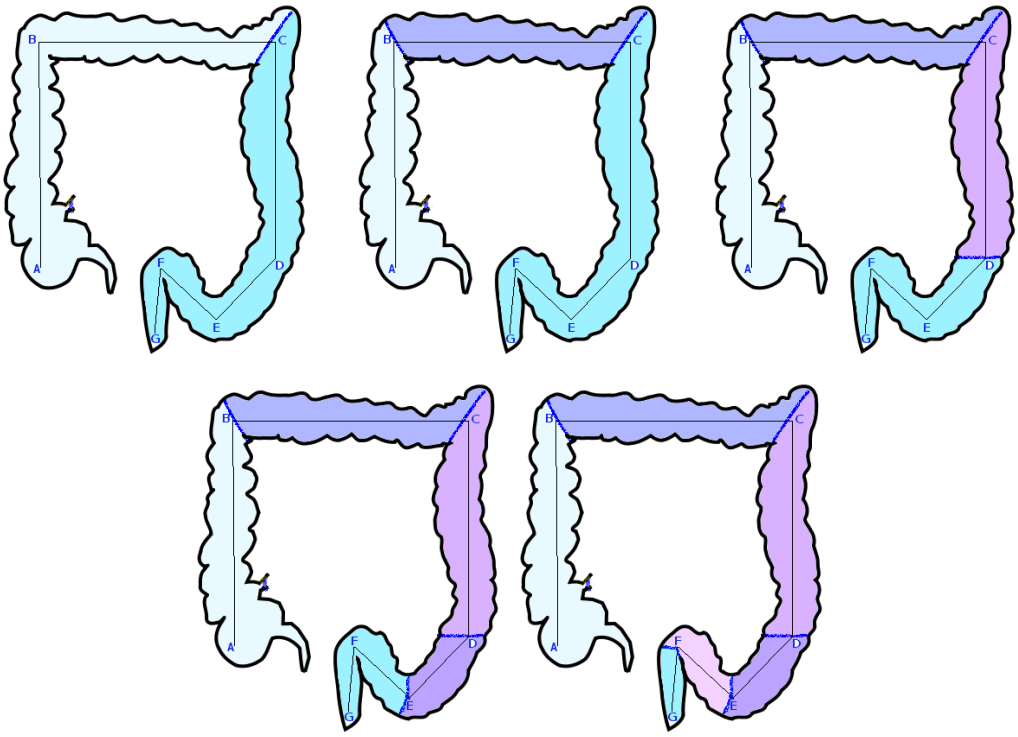


Figure 8.1: Decomposition of the colon into k segments. First line, the colon is decomposed into $k = 2$, $k = 3$, $k = 4$ segments respectively. Second line, the colon is considered by $k = 5$ and $k = 6$ segments

We will explicit the proposed representation for patient of Vatic database. He has all the colon affected by bleeding and ulcers with the spatial distribution given by Figure 8.2.

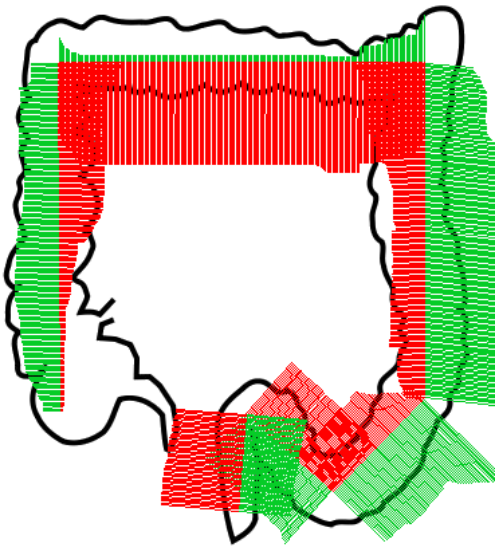


Figure 8.2: Distribution of bleeding (in red) and ulcers (in green) according to annotations provided by a gastroenterologist.

To compute the number of lesions on a segment "MN" for example, we will consider the number of annotations for all frames having the curvilinear abscissa included into the abscissa M and the abscissa N.

For the patient of Figure 8.2, the count of the lesions for the different colon decompositions are, by taking V_b the bleeding vector and V_u the ulcers vector:

```
Patient is represented by the vector V=[Vb,Vu] where for:
k=1,Vb=[1946.0] & Vu=[2332.0]
k=2,Vb=[ 555. 1391.] & Vu=[ 275. 2057.]
k=3,Vb=[ 385.  180. 1381.] & Vu=[ 219.   79. 2034.]
k=4,Vb=[385. 180. 527. 854.] & Vu=[ 219.   79. 1159.  875.]
k=5,Vb=[385. 180. 527. 294. 560.] & Vu=[ 219.   79. 1159.  424.  451.]
k=6,Vb=[385. 180. 527. 294. 328. 232.]
    & Vu=[ 219.   79. 1159.  424.  272.  179.]
```

When considering the area of lesions for each colon decomposition, we will compute the amount of the annotations surface for all the frames included in this colon segment weighted by the length of the considered segment. For the patient that we took as an example, we get the following values for the lesion surface for each considered segment, for all the proposed decompositions:

```
Patient is represented by the vector V=[Vb,Vu] where for:
k=1,Vb=0.427132227741 & Vu=0.369273124895
k=2,Vb=['0.476', '0.378'] & Vu=['0.171', '0.567']
k=3,Vb=['0.195', '0.750', '0.372'] & Vu=['0.274', '0.076', '0.575']
k=4,Vb=['0.195', '0.750', '0.290', '0.451']
    & Vu=['0.274', '0.076', '0.641', '0.511']
k=5,Vb=['0.195', '0.750', '0.290', '0.333', '0.522']
    & Vu=['0.274', '0.076', '0.641', '0.621', '0.445']
k=6,Vb=['0.195', '0.750', '0.290', '0.333', '0.493', '0.556']
    & Vu=['0.274', '0.076', '0.641', '0.621', '0.471', '0.414']
```

8.2.2 Representation of disease severity

The severity of the disease is evaluated using the endoscopic scores i.e. it is considered as an element $\in \mathbb{R}$. During our collaboration with gastroenterologists (XT, CS and YB), we obtained the scores UCEIS and MAYO sub-score for 13 UC patients from the Vatic database that we give in the code chunk:

Code chunk 21: Severity-modelling.py (part 21)

```
# UCEIS score
UCEIS_XT=[8,7,2,1,3,1,5,5,5,5,7,5,7]
UCEIS_CS=[8,6,2,1,2,4,5,5,6,3,6,4,7]
UCEIS_YB=[7,8,3,1,1,0,4,5,3,2,3,3,4]

# MAYO sub-score
MAYO_XT=[3,3,1,1,1,1,3,3,3,3,3,3,3]
MAYO_CS=[3,3,1,1,1,2,3,3,3,1,3,3,3]
MAYO_YB=[3,3,2,1,1,0,3,2,2,1,2,2,2]
```

Next, we will state the linear model adopted to the considered colon decompositions.

8.2.3 The model

We decide to model the severity by the linear functions defined as follow:

$$\begin{aligned} \mathbb{R}^{2k} &\rightarrow \mathbb{R} \\ (D^b, D^u) &\rightarrow \text{Severity} \end{aligned}$$

where $k \in \mathbb{N}^*$ used as the counter of the colon decomposition as mentioned previously, D^l is the data information about the lesions (count or surface l) used for bleeding ($l = b$) and ulcers ($l = u$).

In this work, we have only the severity of the UC in terms of the endoscopic scores UCEIS and MAYO sub-score. For this reason, we decide to compute the model's parameters based on the scores available for patients in the database and the mean square error minimization (MSE):

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^{2k+1}}{\operatorname{argmin}} \sum_{i=1}^{i=N_p} (Y_i - X.\beta)^2 \underset{MSE}{\Longrightarrow} \boxed{\hat{\beta} = (X^t.X)^{-1}.X^t.Y}$$

where $X \in \mathcal{M}(N_p, (2k+1)N_p)$ is the data matrix, N_p is the number f patients (herein 13), and Y is the UCEIS score or the MAYO subscore.

8.3 Results: application to patients

In this section, we will compute the severity scores using the colon decompositions and the linear model discussed of section 8.2. We have to be precise that the annotations are anonymous, namely, we do have not the annotations of each gastroenterologist. Thereby, the data extracted from the spatial distribution of lesions used to compute the model of a score is the same for all the experts. The difference is in the value of the scores that every expert associates with a patient. As a consequence, the parameters' values for each model are distinct.

The scores of the gastroenterologists will be drawn by a continuous point marker line while the model results will be designed by a dashed point marker line.

In Figure 8.3, we present the obtained score for all the gastroenterologists when considering the lesions data in the whole colon, mainly we don't consider "special" segments for this case, (i.e. $k = 1$). The count and the area of lesions bring almost similar results for the estimation of UCEIS and Mayo subscore. For UCEIS given by Dr.XT, there is a difference of 2 points with the real score for four patients. The Mayo subscore is likely to be better estimated using the count of lesions in the colon than their surface. In contradiction, for Dr.CS, the area of lesions seems to be more suitable to estimate the Mayo subscore. For Dr. YB, the scores are better estimated with the surface of lesions. There is a difference of a maximum of one point between the given score UCEIS and the estimated score for almost all the patients (9 over 13 patients / 69%). Using the lesions count, the results are less better with coincidence for only 54% of the patients. For the Mayo subscore, the results are less good.

In the case of $k = 2$ given in Figure 8.4, the curves of estimated UCEIS are near the real one for approximately all the patients. Same remark to Mayo subscore and all the doctors. There is no relevant differnce with results obatined in the case $k = 3$, see Figure 8.5.

From Figure 8.6 to Figure 8.8, we remark that the highest the number of considered colon segments, the more we approximate the severity scores, especially in the case of the UCEIS score.

In comparison of the results of $k = 6$ in Figure 8.8 to results of $k = 1$ in Figure 8.3, one can see the estimated score UCEIS coincides/overlaps with the real one for most of the patients. This phenomenon is less visible to the MAYO sub-score, maybe to the limited values of the scores (0 to 3) compared to the UCEIS, which make the difference more striking/remarkable.

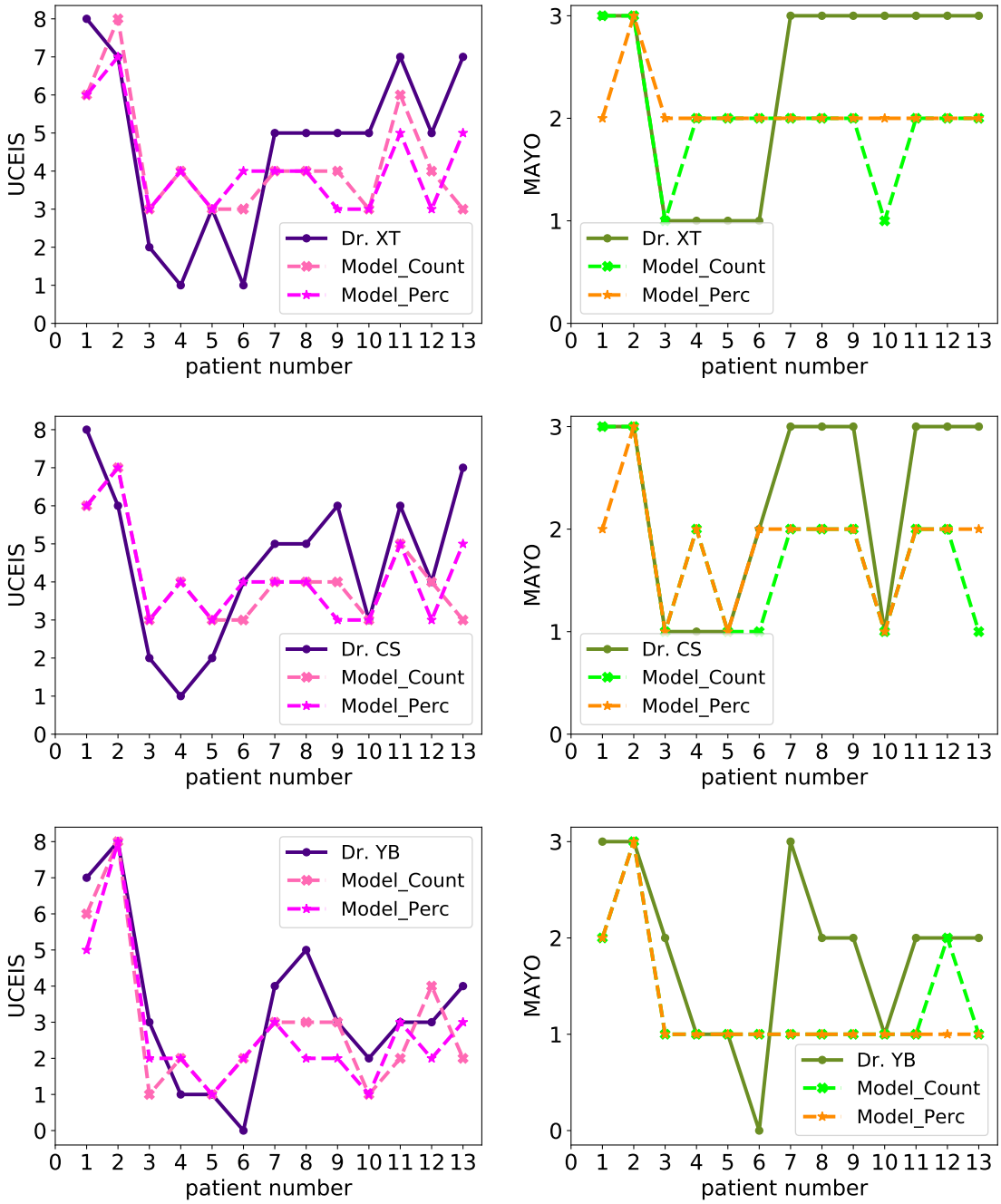


Figure 8.3: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 1$

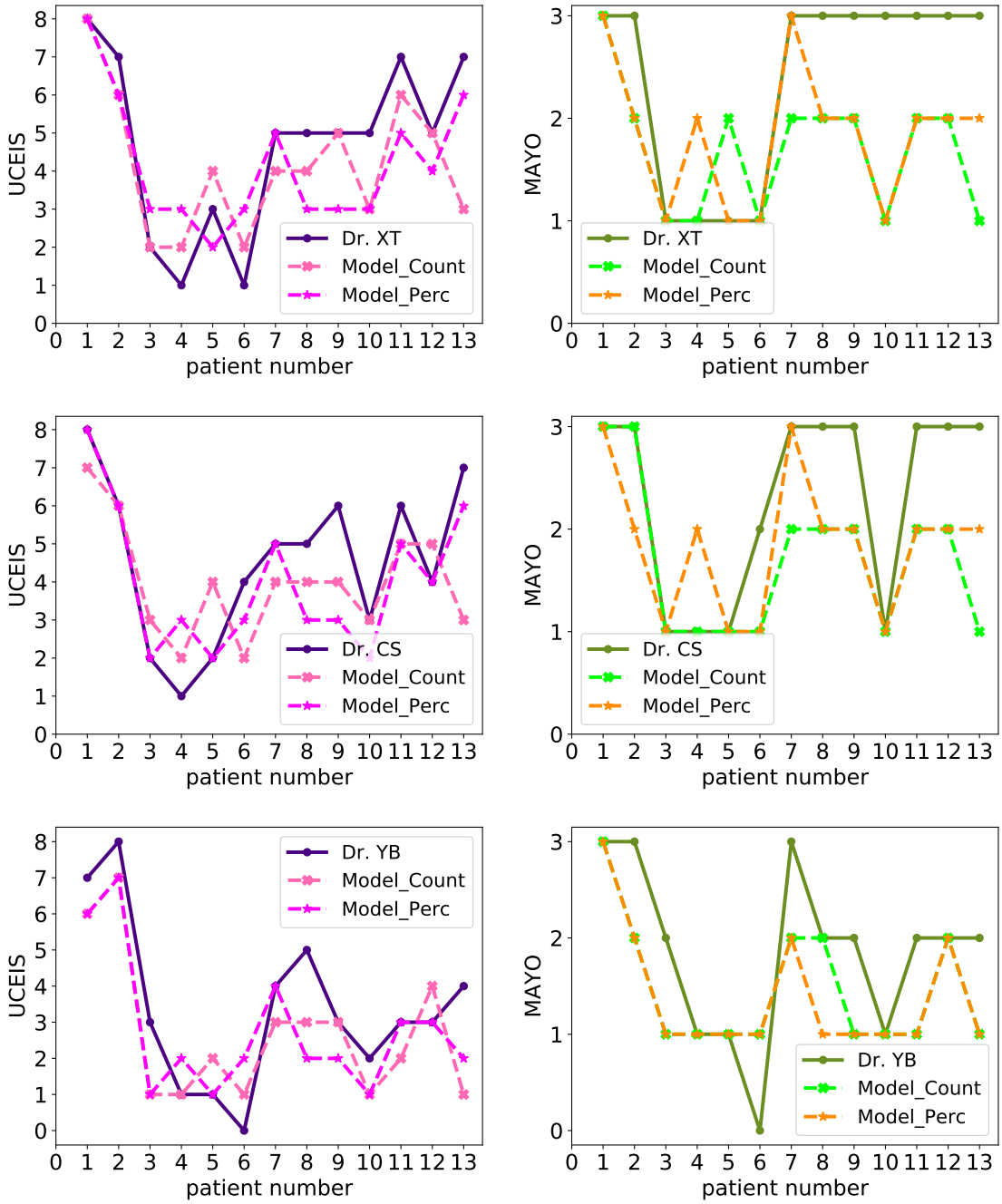


Figure 8.4: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 2$

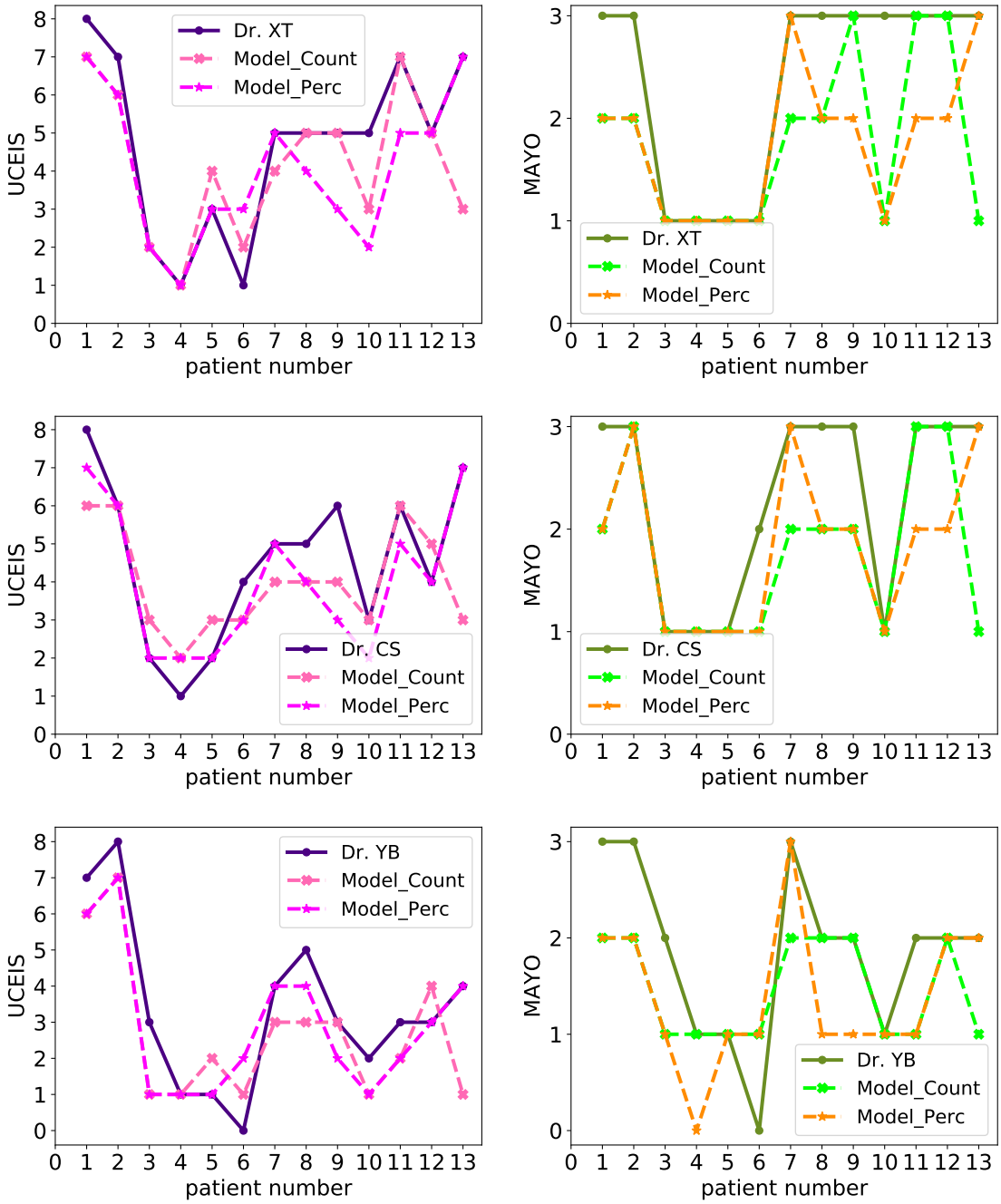


Figure 8.5: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 3$

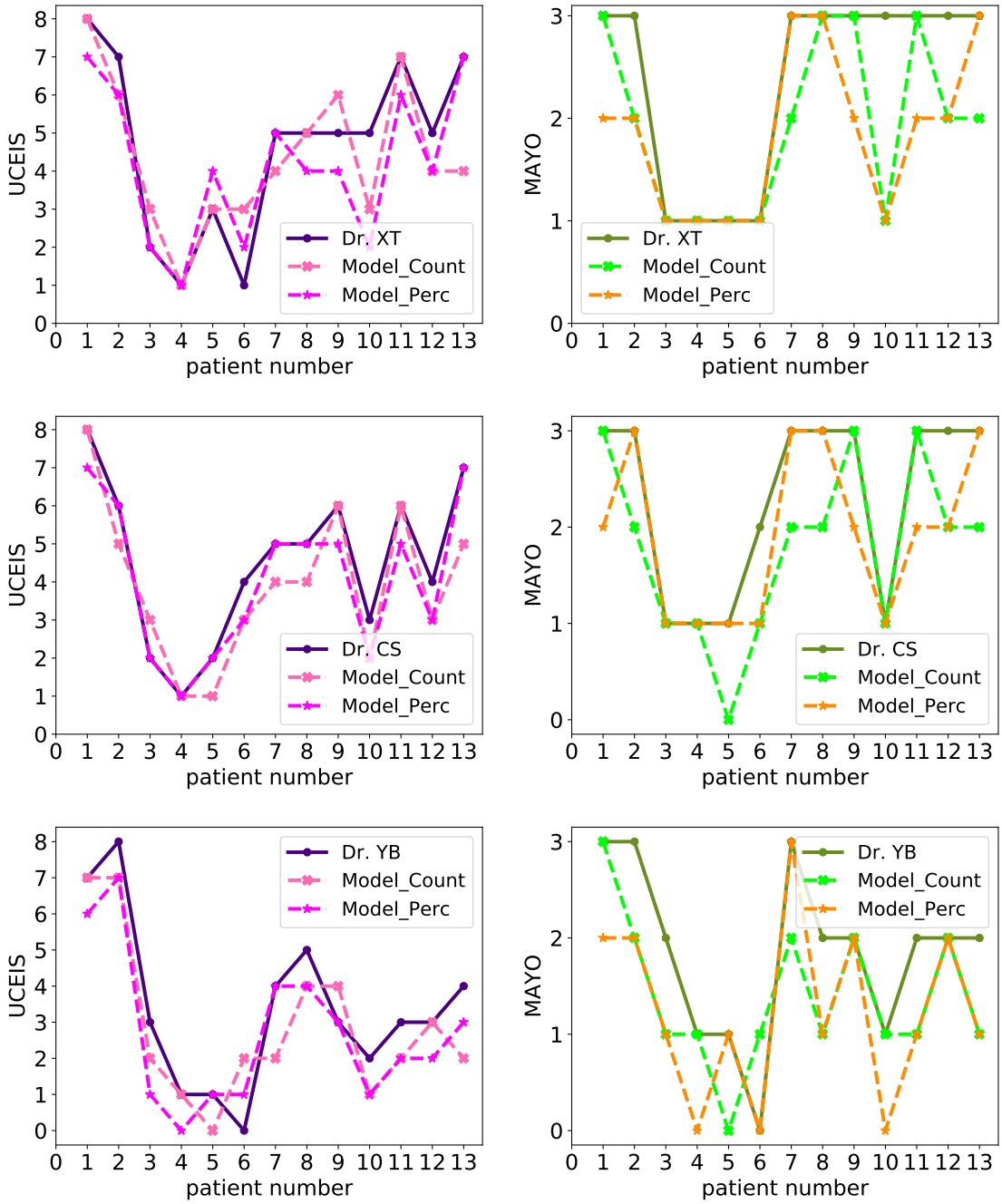


Figure 8.6: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 4$

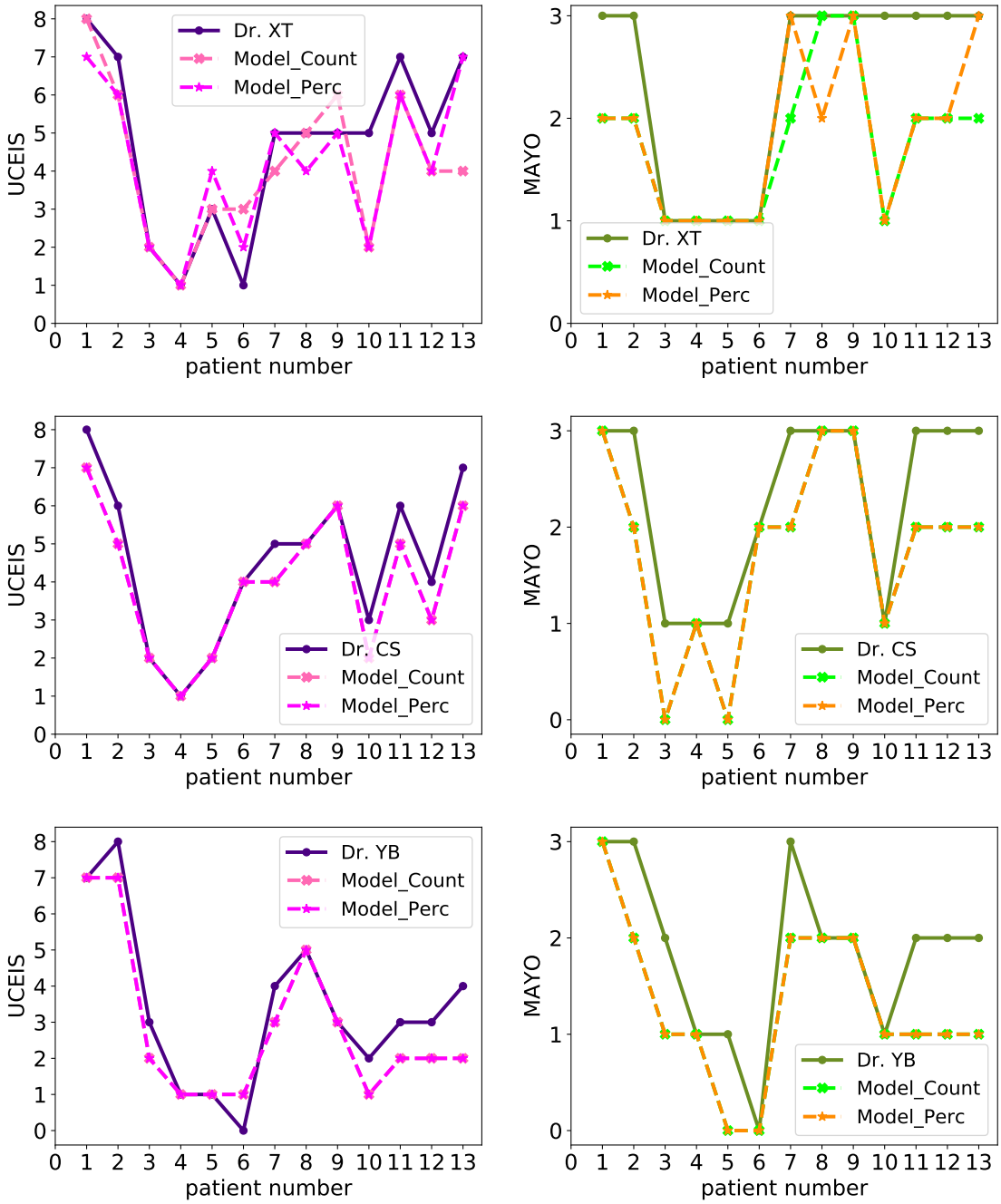


Figure 8.7: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 5$

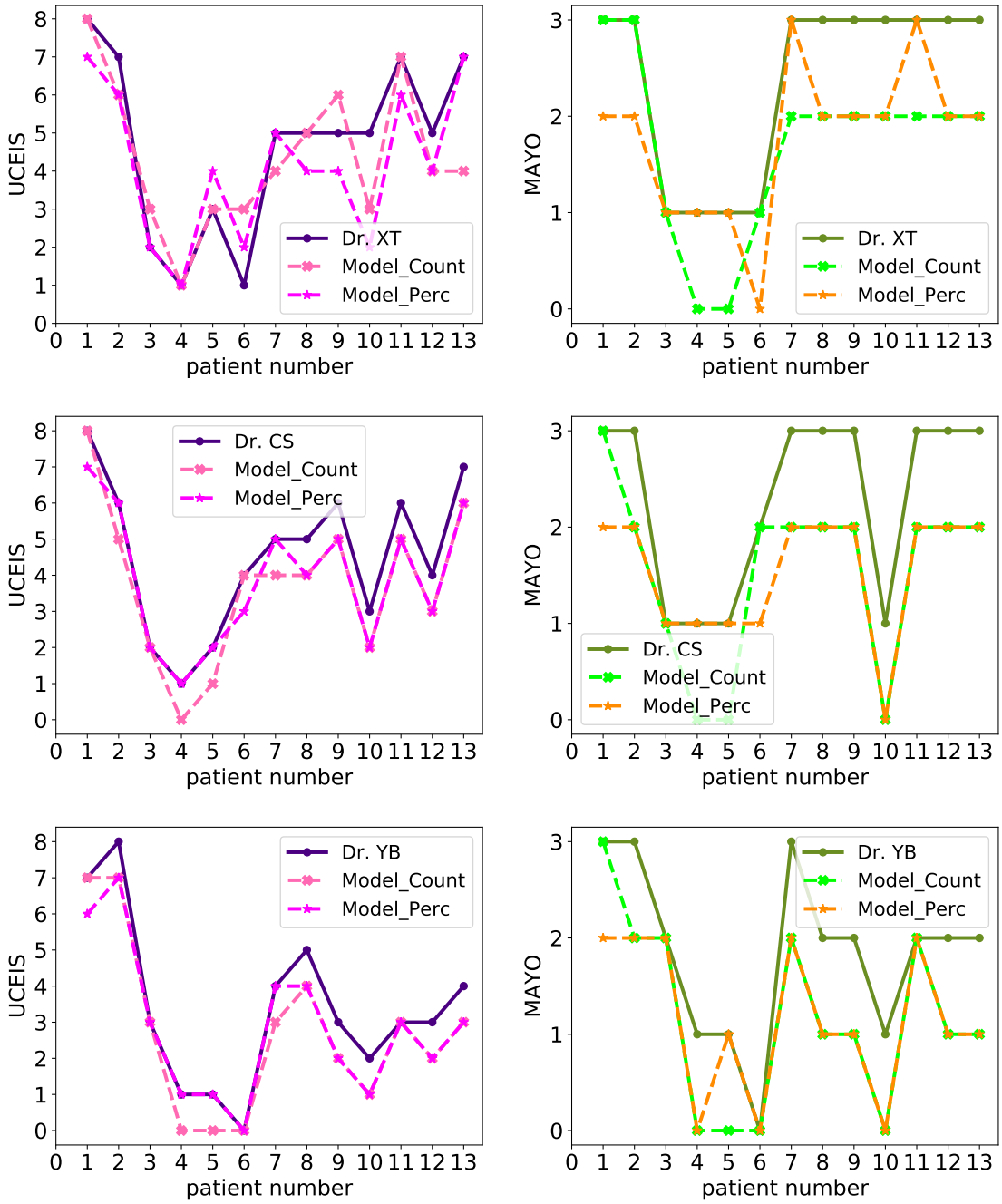


Figure 8.8: Estimated UCEIS at the left and estimated MAYO at the right for the gastroenterologist XT (first row), CS (second row) and YB (third row) in the case of $k = 6$

In the next section, we will search for the optimal colon decomposition for each gastroenterologist and then for all of the gastroenterologists.

8.3.1 Best colon decomposition

From the previous section, we have noticed that when the number of considered colon segments increases, the error between the estimated score and the standard score is decreasing. To visualise the errors of all the decomposition, we used the bar representation for the two types of data (count and surface).

8.3.1.1 For one gastroenterologist

We must specify that the annotations are anonymous, which means namely we do not have the annotations of each gastroenterologist. Thereby, the data information used to compute the model of a gastroenterologist is the same for all the experts. The difference is in the value of the scores that every expert associate with a patient. In this section, we compute the L_2 errors between the severity scores provided by the gastroenterologists and the estimated scores based on the count (or area) of the annotations.

From Figure 8.9, one can conclude that the best colon decomposition corresponds for $k = 6$ for Drs. XT and YB and $k = 5$ for Dr. CS. We can see a sign of variability in the UCEIS scoring between experts.

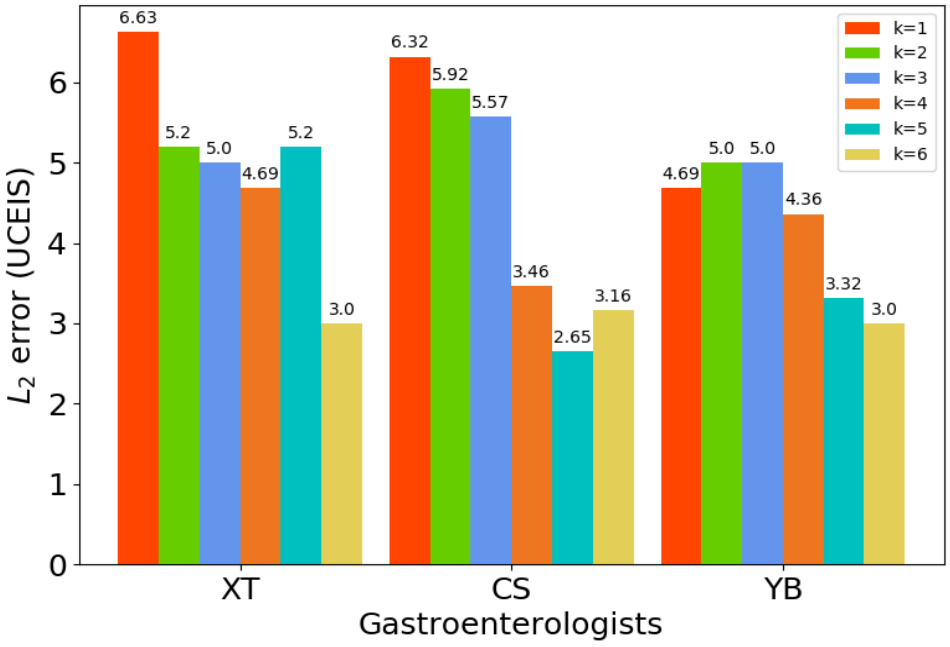


Figure 8.9: L_2 errors between the standard UCEIS score and the estimated score using the count of lesions

By denoting the model $S(B, U) = a_i C_i^b + b_i C_i^u$, for $1 \leq i \leq k$ the index of the considered colon segment, we can extract the best models of the UC severity

for the three gastroenterologists based on UCEIS score:

```
For Dr. XT, the best model with the amount of lesions is for k=6:
a_i= [-0.016, 0.0048, 0.0244, -0.0437, 0.0499, -0.0012]
b_i= [0.0878, -0.0345, -0.0073, -0.0132, 0.0444, -0.1103]
intercept= 2.3411
```

```
For Dr. CS, the best model with the amount of lesions is for k=5:
a_i= [-0.0036, -0.0001, 0.0114, -0.0188, 0.0016]
b_i= [0.0331, -0.0196, 0.0007, 0.0007, -0.003]
intercept= 2.6791
```

```
For Dr. YB, the best model with the amount of lesions is for k=6:
a_i= [-0.0116, 0.0011, 0.0156, 0.0196, 0.0165, -0.0016]
b_i= [0.0496, -0.0066, -0.0073, -0.0172, 0.0609, -0.1146]
intercept= 1.629
```

The best estimation of the MAYO score using the spatial distribution of the count of the lesions corresponds to the decomposition parameter $k = 4$ for Dr. XT, for $k = 4$ and $k = 5$ for Dr. CS, whereas the values $k = 2, 3, 5$ segments lead to the optimal model for Dr. YB (see Figure 8.10).

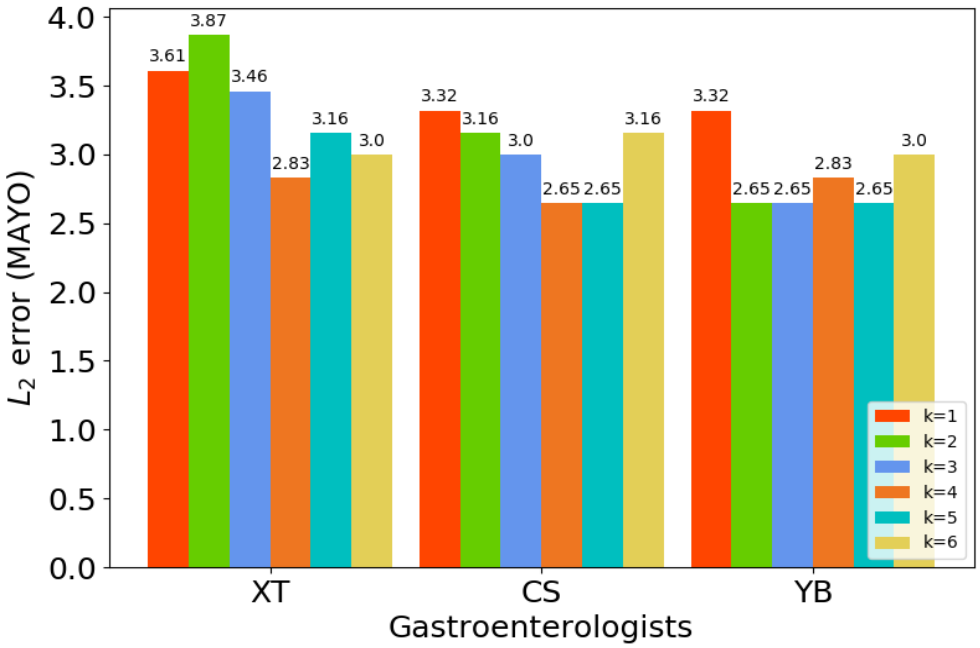


Figure 8.10: L_2 errors between the standard MAYO subscore and the estimated score using the count of lesions

The parameters of the best models by calibration with the MAYO sub-score are:

```
For Dr. XT, the best model with the amount of lesions is for k=4:
a_i= [-0.0002, -0.0006, 0.0015, -0.0008]
```

```
b_i= [-0.0009, 0.0007, 0.0007, 0.0005]
intercept= 1.9309
```

```
For Dr. CS, the best model with the amount of lesions is for k=4:
a_i= [-0.0003, -0.0005, 0.002, -0.0012]
b_i= [0.0021, -0.0014, 0.0008, 0.0004]
intercept= 1.7635
```

```
For Dr. YB, the best model with the amount of lesions is for k=2:
a_i= [1e-06, 0.000201]
b_i= [1.6e-05, 0.000865]
intercept= 1.189278
```

The same work is done for the estimated scores based on the surface of lesions on each segment of the colon. From Figure 8.11, we can notice that the estimation of the UCEIS taking into consideration the surface of the lesions is better in the case of $k = 6$ for the Drs. XT and YB. The best parameter is $k = 5$ for Dr. CS. In parallel, the error of the estimation of the MAYO score is minimal for $k = 6$ in the case of Dr XT and $k = 4$ for the CS scoring, while the minimum error is obtained for $k = 5$ for the Dr. YB (cf Figure 8.12).

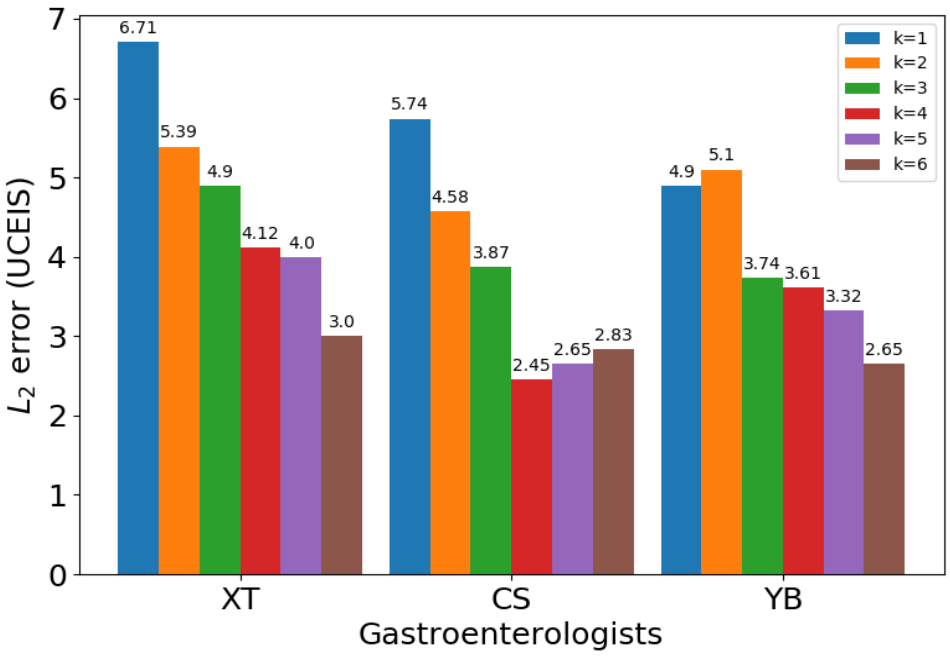


Figure 8.11: L_2 errors between the standard UCEIS score and the estimated score using the surface of lesions

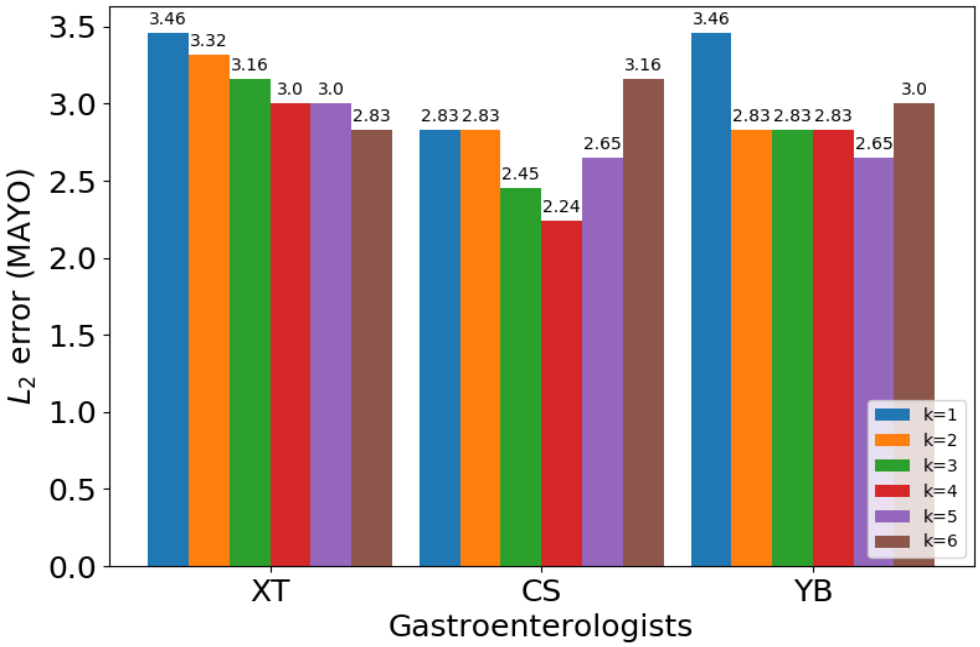


Figure 8.12: L_2 errors between the standard MAYO sub score and the estimated score using the surface of lesions

We designed the score model by $S(B,U) = c_i P_i^b + d_i P_i^u$, and we find the best models for UCEIS are:

```
For Dr. XT, the best model with the surface of lesions is for k=6:
c_i= [12.8202, -23.0941, -22.3747, -110.295, 151.1595, -0.2719]
d_i= [-77.7351, 34.4317, 73.6002, -127.6214, 100.8813, -10.5435]
intercept= -0.6865
```

```
For Dr. CS, the best model with the surface of lesions is for k=4:
c_i= [11.4487, -16.396, 13.3557, -3.8324]
d_i= [0.2204, -17.1345, 30.0981, -10.4851]
intercept= 3.0995
```

```
For Dr. YB, the best model with the surface of lesions is for k=6:
c_i= [11.6301, -25.8716, -153.3957, 70.3061, 103.667, -0.3672]
d_i= [3.1796, 11.5913, -6.7382, 67.5971, -336.2431, 278.0887]
intercept= -1.737
```

The optimal models for the calibration with the MAYO sub-score are:

```
For Dr. CS, the best model with the surface of lesions is for k=6:
c_i= [4.4749, -12.465, 14.3398, -78.957, 80.1198, -4.8484]
d_i= [-26.6621, 8.062, 31.9028, -67.4165, 58.9363, -7.8604]
intercept= 0.4239
```

```
For Dr. CS, the best model with the surface of lesions is for k=4:
c_i= [3.8453, -6.6425, 3.6344, 1.328]
```

```
d_i= [-4.1214, -3.9306, 9.4453, -0.586]
intercept= 1.1703
```

```
For Dr. YB, the best model with the surface of lesions is for k=5:
c_i= [2.422, 0.5116, -46.7343, 38.9877, 6.4196]
d_i= [-37.4402, 22.0243, 14.6967, -0.895, -2.4345]
intercept= 0.1942
```

To summarise this part, we conclude that for Drs. XT and YB, the UCEIS computed based on the colon decomposition $k = 6$ is the best, independently of the data type (count/surface). Dr. CS shows minimum errors when considering the decomposition of the colon into $k = 5$ and $k = 4$ segments to compute the estimated score based on the count and surface of lesions respectively. For the MAYO sub-score, the optimal values are for $k \geq 4$, except for Dr. YB who shows the same behaviour for $k = 2, 3, 5$ in the case of the count of lesions (cf Figure 8.10).

8.3.1.2 For all the gastroenterologists

In this section, we proceed to the research of the best colon segments number k with respect to the minimum error to the UCEIS and the MAYO (sub) scores, and this for all the gastroenterologists. We decide to compute the L_2 error using the optimal estimated scores $\bar{S}$ and the data scores {UCEIS,MAYO} designed by S, and from gastroenterologist G in {XT,CS,YB} and patients number N_p:

$$\hat{k} = \underset{k\in\{1,\dots,6\}}{\operatorname{argmin}} \sqrt{\sum_{k=1}^{k=6} \sum_{G\in\{XT,CS,YB\}} \sum_{p=1}^{p=N_p} \left(\bar{S}(G,p) - S(G,p)\right)^2} \tag{8.1}$$

The best models are:

```
UCEIS-Count: the min error 5.292 is obtained for k= 6
UCEIS-Perc: the min error 4.899 is obtained for k= 6
MAYO-Count: the min error 4.796 is obtained for k= 4
MAYO-Perc: the min error 4.69 is obtained for k= 4
```

We conclude that based on the disease scoring for the three gastroenterologists, the best colon decomposition is $k = 6$ segments in the case of UCEIS and $k = 4$ segments in the case of the MAYO sub-score regardless of the type of the data (amount or surface).

8.4 Conclusion

The visualisation of the spatial distribution of the bleeding and ulcer lesions exhibited in the previous chapter emphasizes the variability of disease activity between UC patients. The colon segments are not usually similarly affected by the lesions. Bleeding may be absent or more abundant than ulcers and vice

versa. Therefore, the endoscopic numerical assessment of disease severity (refer to chapter 4) that take into account only the most severe lesions and discard their extent along the colon cannot reveal the precise state of the disease. As a consequence, we propose a UC severity modelling process that takes into account the spatial extent of the lesions in the colon. The severity of the disease is represented by the endoscopic scores applied for clinical practice, mainly the Mayo subscore (cf section 4.6) and UCEIS score (cf section 4.9).

We decide to use the linear functions $\in$ {Functions: patient $\Longrightarrow$ $\mathbb{R}$} and the colon parametrization of chapter 7. Instead of considering the patient with the most severe lesions like Mayo and UCEIS, we rather consider the count (or surface) of the lesions on each colon segment. We have proposed different colon decompositions to extract the local number of lesions. We do the same work by using the area of the lesions on each segment. The final score counts for all the colon segments through the linear model as discussed in section 8.2.

We tested the method on 13 videos from the vatic database for which three gastroenterologists have associated the severity score UCEIS as defined in Table 4.10 and the MAYO sub-score given in Table 4.6. According to L_2 distance (cf section 8.3), as per reported results, we find that considering the colon into its four principal/main segments is optimal to estimate the MAYO sub-score. Whereas considering the colon into six segments is optimal to compute the UCEIS score.

In the future, we want to go further in the exploration of all the decompositions of the colon and in parallel the exploration of other types of mathematical models. A larger database, perhaps including new objective markers of UC, such as calprotectin could enhance this part of the work.

In the following chapter, we will go further into the application of the lesions spatial extent to bring responses to a set of medical statements like the evolution of the severity with the count of the lesions, numerical evaluation of the scoring variability between gastroenterologists and study of the abundance of the lesions regarding the colon segments.

Chapter 9

Mathematical proofs for medical hypotheses

Abstract

This chapter was motivated by discussions with our collaborators on the results presented in the previous chapters, in particular the lesions profiles explored in chapter 7. We try to bring some proof to deal with medical practice issues and/or hypotheses evolving the UC severity and the distribution of the lesions along the colon. Due to the limited number of patients considered for this study (i.e. only 13 patients), we introduced the compatible models that have performance near the best model performance with some defined/accepted error threshold.

From a medical point of view, the inflammation of ulcerative colitis is recognized to start with the rectum. Mathematically, we seek for analysis of bleeding and ulcers lesions distribution bias through linear and ordinary differential equation solutions. On the other hand, we study the severity growth according to the total annotations count found in colonoscopy videos. Finally, we look into the lesions assessment variability between gastroenterologists. The correlation between bleeding and ulcers was also studied to examine if the disease begins with a special type of lesion.

The obtained results suggest that the UC lesions are almost concentrated near the rectum, however, in comparison to ulcers, the bleeding is spreading preferably from the descent segment to the end of the colon. Despite the severity increase with the lesions count for all the doctors, they seem to not evaluate similarly the bleeding during the severity assessment. The lesions showed uncorrelated, there was no response to the type of disease beginning lesions.

The works of this chapter elucidate some medical hypotheses not yet proved and show the necessity of the information extracted from the distribution of the lesions.

In general, the formulation of a biological problems by the mathematicians is very complicated.

9.1 Lesions repartition bias

We mentioned in the introduction of this manuscript that "*ulcerative colitis is generally characterized by the onset of inflammation starting in the rectum*" and spreading continuously upstream to the colon. This medical hypothesis has not yet been validated either by a medical study or by a mathematical approach or by a computer algorithm. According to the strategy of data considered along with the thesis works, the statement can be considered in terms of lesions as follows:

- **Statement 1:** the lesions are hosted advantageously between the descent colon and the rectum
- **Statement 2:** the lesions are more abundant near the rectum

We continue using the colon parametrization of chapter 7, where the colon is represented by the segment $[0,1]$. Along with this work, we will consider the count of lesions i.e. bleeding and ulcers separately. This assumption means that the patient is represented by a real number corresponding to the count of bleeding or ulcers. More, the case of different colon decompositions will not be tested as in the previous chapter.

As a consequence, we need to explore the set of functions $\{[0,1] \to \mathbb{R}\}$ of infinite dimension. We therefore restrict the search for two sub-class of models:

- Set of affine functions i.e. $\mathcal{M}_a = \{y = a\,s + b \text{ for } a \text{ and } b \in \mathbb{R}\}$
- Set of solutions of the first order ordinary differential equation i.e. $\mathcal{M}_o = \{y' = ay, a \in \mathbb{R}\}$

The affine functions space $\mathcal{M}_a$ can be represented by the couple $\{a,b\}$ of $\mathbb{R}^2$. In particular, one can represent the affine functions compatible (that we will define later in this chapter) with observations as subsets of $\mathbb{R}^2$, and scientific statements or mathematical properties correspond to regions of the plane. Therefore, proving a scientific statement corresponds to showing that all repetitions of the data belong to the same region of the plane, or that all compatible models belong there.

On the other hand, the space $\mathcal{M}_o$ consists of the exponential functions $\{y = Ce^{as}\}$. It is also a two-parameter space that can be represented in the plane $\mathbb{R}^2$. In this models' space, the parameter C can be interpreted as a total number of lesions, and the parameter a as a distribution' indicator. If $a \geq 0$ then the function is increasing, and the lesions are more abundant close to the rectum than in the other colon segments. A value of a close to 0 indicates that the lesions are uniformly distributed through the colon.

An appropriate definition for **Statement 1** can be given by:

Definition 1. *Let p be a function defining the state of patient i.e. its colon state according to the chosen lesion. If p satisfies:*

$$\int_0^{0.5} p(s)ds < \int_{0.5}^{1} p(s)ds \tag{9.1}$$

where s is the curvilinear asbcissa $\in [0,1]$, then the lesions of patient with profile p are spreading advantageously from the descent colon segment to the rectum.

The patient will be then represented by zero or one. In the following code chunk, we perform the computation for 37 patients from Vatic database:

Code chunk 22: bias_repartition

```
bias_b, bias_u = np.full(37,True), np.full(37,True)
for i in range(0,36):
    p = Data_perc_b[i]
    L,L2 = len(p), (int) (len(p)/2)
    bias_b[i] = sum(p[0:L2])/L < sum(p[(L2+1):L])/L
    p = Data_perc_u[i]
    bias_u[i] = sum(p[0:L2])/L < sum(p[(L2+1):L])/L

print sum(bias_b), "patients over ", len(bias_b),"check Statement 1 " \
"(left < right) in the case of bleeding."

print sum(bias_u), "patients over ", len(bias_b),"check Statement 1 " \
"(left < right) in the case of ulcers."
```

Interpret with `python2`

```
28 patients over  37 check Statement 1 (left < right) in the case of bleeding.
21 patients over  37 check Statement 1 (left < right) in the case of ulcers.
```

We conclude that according to our database, 76% of patients satisfy **Statement 1** for bleeding profile while only 57% of patients have ulcers preferably distributed between the descent colon and the rectum. Consequently, based on the 37 patients of our database, we can say that:

"*The ulcerative colitis disease generates lesions that are preferably hosted from the descent colon to the rectum.*"

In practice, all UC patients will not necessarily verify statement 1. We are therefore led to assess the level of proof provided by the number of patients above. For this, we seek to construct a statistical test belonging to the **Statement 1** region against the absence of bias. Under the assumption of no bias and since we have considered that all the colon positions are equally probable, the lesions are therefore evenly distributed in the colon, between 0 and 1. As a consequence, the fact that a patient verifies statement 1 is an event of 0.5

probability (by symmetry), and the number of patients verifying statement 1 follows a binomial distribution. The statistical test presented below leads us to reject the hypothesis $\lambda = 0.5$ for bleeding and ulcer:

Code chunk 23: bias_repartition (part 2)

```
from scipy.stats import binom
# Bleeding
print "Proba[B >= ",sum(bias_b), "sur", len(bias_b),"] =", \
1-binom.cdf(sum(bias_b),len(bias_b),0.5)
# Ulcer
print "Proba[U >= ",sum(bias_u), "sur", len(bias_u),"] =", \
1-binom.cdf(sum(bias_u),len(bias_u),0.5)
```

Interpret with `python2`

```
Proba[B >=  28 sur 37 ] = 0.0003764485300052911
Proba[U >=  21 sur 37 ] = 0.1620043000439182
```

A more complete analysis can be done in this case. Indeed, UC refers to a set of patients with similar typical symptoms of the disease. Through statement 1, we project this set of patients onto a set of Boolean numbers as discussed before. The space of observations for a statement on the UC is, therefore, $\{0,1\}^{37}$. The following assumptions can be made by "construction":

- independence: UC is not a transmissible disease, until proven otherwise the lesions are specific for each patient,
- same distribution: since we assumed that patients are identical/indistinguishable.

Each patient can therefore be considered as the realization of an independent, identically distributed Bernoulli random variable. The set of HCR models is the set of Bernoulli variables with parameter $\lambda \in [0,1]$. The likelihood $\mathbb{P}(\text{data}|\lambda)$ indicates whether the observed values are plausible for a model, or vice versa whether a model is compatible with the data. The **Statement 1** is corresponding to Bernoulli variables with $\lambda \in]0.5,1]$.

In Figure 9.1, we arbitrarily define the notion of compatible model by $\ln(\mathbb{P}(\text{data}|\lambda)) > -5)$, and we plot the set of models parameters λ compatible with the 37 patients in our database. The Figure 9.1 indicates that models with $\lambda \leq 0.55$ aren't compatible with bleeding distribution bias. As the set of compatible models for bleeding is included in $]0.5,1]$, we conclude that statement 1 is well demonstrated by accumulation of models. On the other hand, the models $\lambda > 0.75$ are incompatible with the ulcer distribution bias. As the set of models compatible with the ulcer bias is neither included in E1 nor in its complement, it is not possible to conclude for ulcers.

Code chunk 24: bias_repartition (part 3)

```
import matplotlib.pyplot as plt
def plot_compat(figname,n,k,p):
    logp = binom.logpmf(k,n,p)
    p_compat = p[logp>-5]
    fig,ax = plt.subplots(figsize=(6,3.5))
    _ = ax.margins(y=0.12)
    ax.plot(p,logp)
    ax.plot(p_compat,p_compat*0-5)
    ax.plot([0.5, 0.5],[-15, 0])
    ax.text(float(k)/n,np.max(logp),'Max Likelihood',\
            va='bottom',ha='center')
    ax.text(np.mean(p_compat),-5,'Compatible models',\
            va='top',ha='center')
    ax.set_xlabel('lambda parameter')
    ax.set_ylabel('Likelihood (log)')
    fig.savefig(figname,bbox_inches = 'tight')
    plt.close()

plot_compat('bias_compat_b.pdf',\
    len(bias_b),sum(bias_b),np.arange(0.4,0.98,0.01))

plot_compat('bias_compat_u.pdf',\
    len(bias_u),sum(bias_u),np.arange(0.2,0.9,0.01))
```

9.2 Evolution of the severity with the lesions

In this section, we study the following medical question:

Does the UC severity increasing with the number of lesions in the colonoscopy video?

This statement corresponds to considering the severity in the subset of increasing models among the set of functions of the type "$\texttt{N\_bleed} \to \mathbb{R}$" in the case of bleeding and "$\texttt{N\_ulcer} \to \mathbb{R}$" in the case of ulcers. Here, we denote by `N_bleed` the number of bleeding annotations in the colonoscopy video whereas the `N_ulcer` represents the amount of the ulcer annotations.

We decide to use the total number of lesions extracted from that colonoscopy video considering the colon as one segment ($k = 1$). The set of models is restricted to linear functions $\{f(l) = al + b\}$, ($l = B$ for bleeding and $l = U$ for ulcers). Consequently, the increasing functions are the functions with a positive slope, otherwise, the severity will be decreasing with the number of lesions.

In our database, we are limited to the scores of 13 patients which limited the accuracy of the conclusion/response (see subsection 8.2.2). In the absence of repetition, we propose to verify the obtained response by the accumulation of compatible models, that is to say, we want to know if all the compatible models

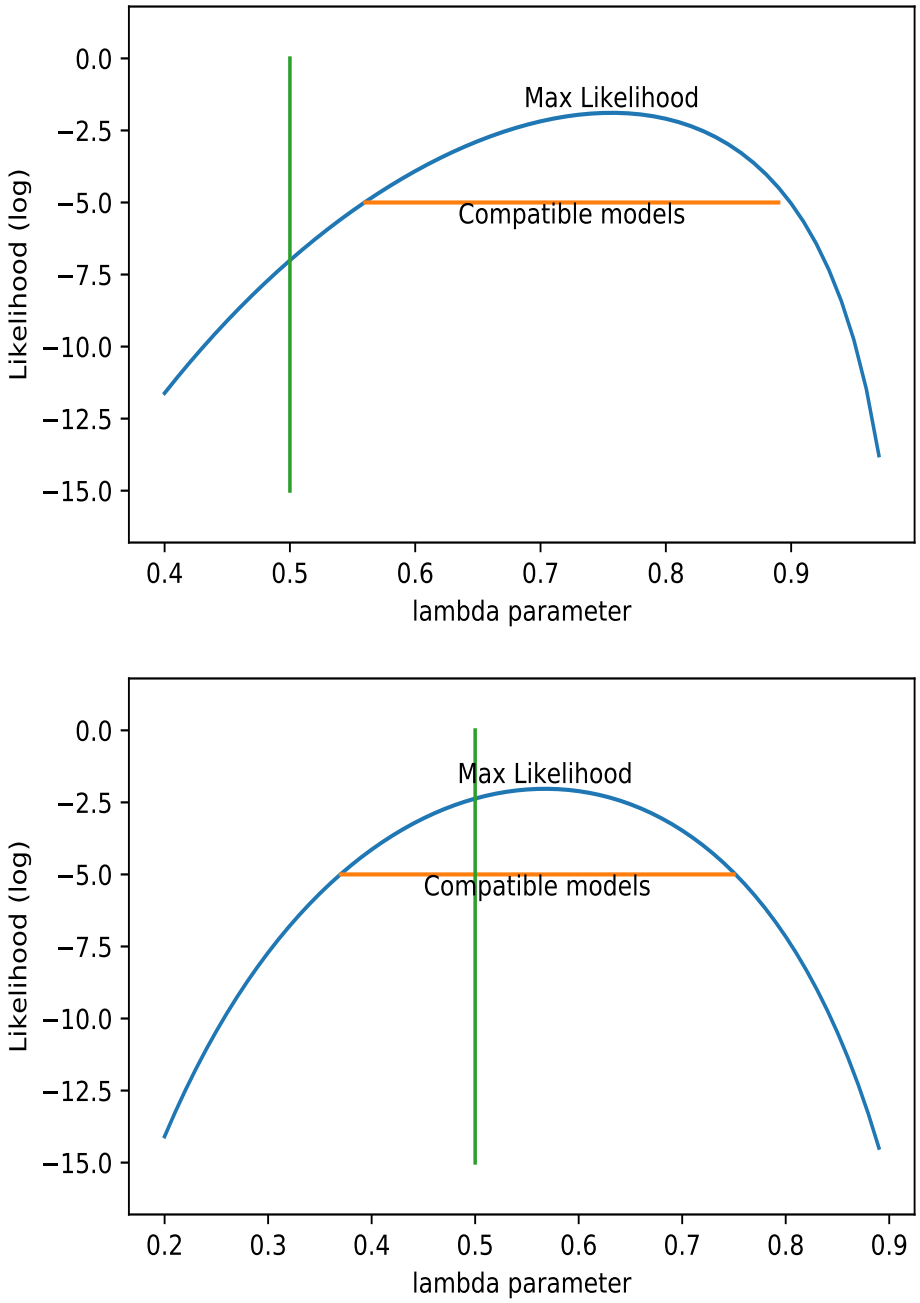


Figure 9.1: Compatible Bernoulli models according to the observed distribution bias. All compatible models belong to $]0.5, 1]$ in the case of bleeding (top), but the situation is indeterminate for ulcers (bottom).

of the severity are increasing functions.

We define the $\alpha-$compatible models, the models that are α distance far from the optimal models:

Definition 2. *The $\alpha-$compatible models $\bar{m}$ according to the optimal model $\hat{m}$ are defined as: $d(D,\bar{m}) \leq (1+\alpha)d(D,\hat{m})$, where d can be the distance L_1, L_2, standard deviation or any other type of discrete distance type, D denotes the data and α the error threshold.*

The objective is then to search the slopes of the 0.05-compatible linear models defined as $\{d(\bar{a}B+\bar{b},Y) < 1.05d(\hat{a}B+\hat{b},Y)\}$, where $\hat{a}$ and $\hat{b}$ are the linear regression parameters of the best model, Y can be the UCEIS score or the MAYO sub-score. Same work will be done in the case of ulcers (the choice of error's threshold is not restricted to 5%). The standard deviation, denoted by `std` in `python` library, was used as distance to search for the compatible models in the following chunks:

Code chunk 25: Severity-modelling.py

```
def compatible_linmodels(x,y,alpha,candidates):
    # Return array containing alpha-compatible slopes among candidate slopes
    best_slope, best_intercept, _,_,_ = stats.linregress(x,y)
    best_dist = np.std(y - best_slope * x - best_intercept)
    return np.sort(np.append([ s for s in candidates
        if np.std(y - s*x) < (1+alpha)*best_dist], best_slope))
```

We study the 0.05-compatible models for bleeding:

Code chunk 26: Severity-modelling.py (part 2)

```
x = np.array(N_bleed).transpose()
fig,ax = plt.subplots(figsize=(6,4))
for (i,uc,l) in zip(range(3),[UCEIS_XT,UCEIS_CS,UCEIS_YB],['XT','CS','YB']):
    s = compatible_linmodels(x,uc,0.05,np.arange(-0.02,0.02,3e-5))
    _ = ax.plot(s,s*0+5-i,'.-',label='UCEIS_'+l)

for (i,mayo,l) in zip(range(3),[MAYO_XT,MAYO_CS,MAYO_YB],['XT','CS','YB']):
    s = compatible_linmodels(x,mayo,0.05,np.arange(-0.02,0.02,3e-5))
    _ = ax.plot(s,s*0+2-i,'.-',label='MAYO_'+l)

_ = ax.plot([0,0],[0,5])
_ = ax.legend(loc=4); _ = ax.set_xlabel('Slope')
fig.savefig('severite-bleed-compatible.pdf')
plt.close()
```

and for ulcers as well:

Code chunk 27: Severity-modelling.py (part 3)

```
x = np.array(N_ulcer).transpose()
fig,ax = plt.subplots(figsize=(6,4))
for (i,uc,l) in zip(range(3),[UCEIS_XT,UCEIS_CS,UCEIS_YB],['XT','CS','YB']):
    s = compatible_linmodels(x,uc,0.05,np.arange(-0.025,0.025,3e-5))
    _ = ax.plot(s,s*0+5-i,'.-',label='UCEIS_'+l)

for (i,mayo,l) in zip(range(3),[MAYO_XT,MAYO_CS,MAYO_YB],['XT','CS','YB']):
    s = compatible_linmodels(x,mayo,0.05,np.arange(-0.025,0.025,3e-5))
    _ = ax.plot(s,s*0+2-i,'.-',label='MAYO_'+l)

_ = ax.plot([0,0],[0,5])
_ = ax.legend(loc=4); _ = ax.set_xlabel('Slope')
fig.savefig('severite-ulcer-compatible.pdf')
```

From Figure 9.2, we conclude that the compatible models for bleeding are increasing, namely the slopes of the compatible models are almost positive, except of the case of the severity grading for the gastroenterologist YB. All the compatible models based on ulcers have positive slopes (cf Figure 9.3). This means that the severity of the UC disease is increasing with the number of lesions.

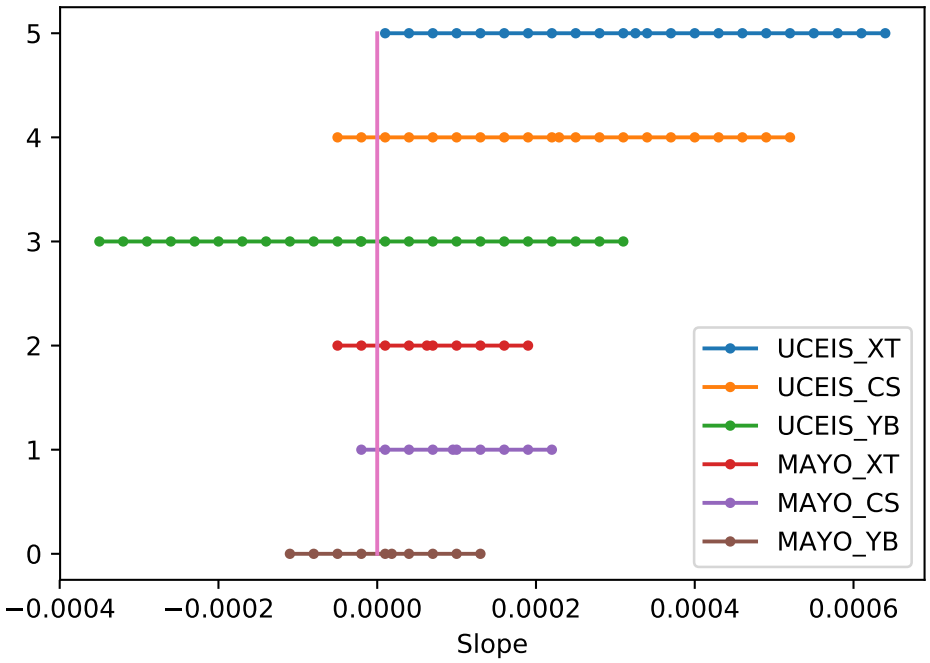


Figure 9.2: Slopes 5% compatible for the UCEIS and MAYO scores assigned by the three gastroenterologists for bleeding. Compatible models almost tend to have a positive slope, except in the case of the doctor YB

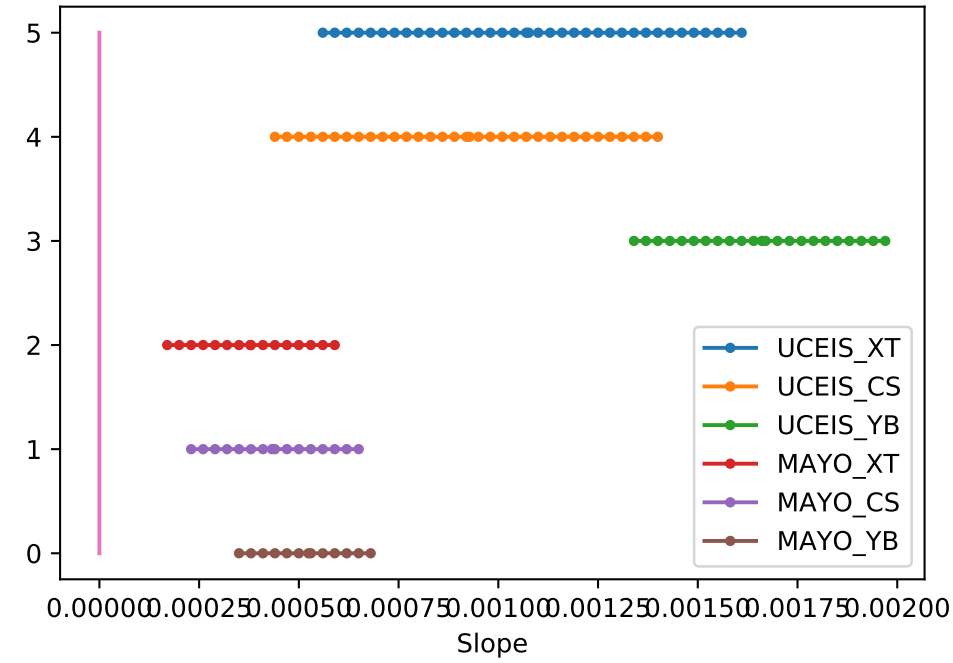


Figure 9.3: Slopes 5% compatible for the UCEIS and MAYO scores assigned by the three gastroenterologists for ulcers. All the compatible models have a positive slope

9.3 Study of the inter-gastroenterologists variability

In this section, we aim to measure the variability of the UC severity scores (UCEIS/MAYO sub-score) between the gastroenterologists based on the obtained results in the section 8.3 for the case $k = 1$, i.e. when the colon is considered as one segment.

The severity can be written as function f such as $f(B,U) = aB + bU + c$, $(a,b,c) \in \mathbb{R}^3$. We thus obtain three linear models for each score corresponding to the three gastroenterologists. In the following code chunk we apply the function `LinearRegression` from the python library to compute the models parameters:

Code chunk 28: Severity-modelling.py (part 4)

```
from sklearn.linear_model import LinearRegression
# UCEIS
Nb_annot = np.array([N_bleed, N_ulcer]).transpose()
models_uceis = [LinearRegression().fit(Nb_annot,g)
                for g in [UCEIS_XT, UCEIS_CS, UCEIS_YB]]
for (m,g) in zip(models_uceis,['XT','CS','YB']):
    print 'f(B,U)_',g, '=', round(m.coef_[0],6), 'B + ', \
                            round(m.coef_[1],6), 'U + ', round(m.intercept_,3)
```

Interpret with python2

```
f(B,U)_ XT = 0.000416 B +  0.001185 U +  3.092
f(B,U)_ CS = 0.000306 B +  0.001007 U +  3.252
f(B,U)_ YB = 0.000108 B +  0.001689 U +  1.859
```

Code chunk 29: Severity-modelling.py (part 5)

```
# MAYO sub-score
models_mayo = [LinearRegression().fit(Nb_annot,g)
                for g in [MAYO_XT, MAYO_CS, MAYO_YB]]
for (m,g) in zip(models_mayo,['XT','CS','YB']):
    print 'f(B,U)_',g, '=', round(m.coef_[0],6), 'B + ', \
                            round(m.coef_[1],6), 'U + ', round(m.intercept_,3)
```

Interpret with python2

```
f(B,U)_ XT = 9.3e-05 B +  0.000402 U +  1.916
f(B,U)_ CS = 0.000131 B +  0.000467 U +  1.728
f(B,U)_ YB = 6e-05 B +  0.00054 U +  1.32
```

Even though that the coefficients for the number of ulcers are roughly the same for all the models and gastroenterologists, the obtained models present significant differences between the coefficients for bleeding for all the experts. This illustrates the inter-expert variability of the scoring process.

We also study the evolution of the severity scores with the surface of the lesions:

Code chunk 30: Severity-modelling.py (part 6)

```
Pr_annot = np.array([P_bleed, P_ulcer]).transpose()
models_uceis_p = [LinearRegression().fit(Pr_annot,g)
                  for g in [UCEIS_XT, UCEIS_CS, UCEIS_YB]]
for (m,g) in zip(models_uceis_p,['XT','CS','YB']):
    print 'f(B,U)_',g, '=', round(m.coef_[0],6), 'B + ', \
                             round(m.coef_[1],6), 'U + ', round(m.intercept_,3)
```

Interpret with python2

```
f(B,U)_ XT = 2.781648 B +  5.908825 U +  3.041
f(B,U)_ CS = 2.735894 B +  5.298106 U +  2.983
f(B,U)_ YB = 1.309063 B +  8.475124 U +  1.841
```

Code chunk 31: Severity-modelling.py (part 7)

```
models_mayo_p = [LinearRegression().fit(Pr_annot,g)
                  for g in [MAYO_XT, MAYO_CS, MAYO_YB]]
for (m,g) in zip(models_mayo_p,['XT','CS','YB']):
    print 'f(B,U)_',g, '=', round(m.coef_[0],6), 'B + ', \
                             round(m.coef_[1],6), 'U + ', round(m.intercept_,3)
```

Interpret with python2

```
f(B,U)_ XT = 0.569324 B +  1.67093 U +  1.985
f(B,U)_ CS = 0.924414 B +  2.147158 U +  1.734
f(B,U)_ YB = 0.301337 B +  2.534526 U +  1.412
```

We find that taking into consideration the total surface of annotated lesions in a given video, the variability of the severity measurement on the bleeding and ulcer for all the scores.

Despite the simplicity of the procedure, we obtained a quantitative measurement of teh variability of the grading systems between doctors.

9.4 Correlation between bleeding and ulcers

We tested the Pearson [Sti89], Spearman [MWLJ13] and Kendall [Ken38] correlation indexes to evaluate the relation between bleeding and ulcers due to the lack of information about the possible type of relation (linear, monotone ..). They were first apply on the count of lesions:

Code chunk 32: Severity-modelling.py (part 8)

```
import scipy.stats
# Pearson
Corre_Pearson, pvalue_Pearson = scipy.stats.pearsonr(N_bleed,N_ulcer)
print "Pearson_R= "+str(round(Corre_Pearson,3))+\
        " and the Pearson_Pval= "+str(round(pvalue_Pearson,3))

# Spearman
rho, pvalue_Spearman = scipy.stats.spearmanr(N_bleed, N_ulcer)
print "Spearman_R= "+str(round(rho,3))+\
        " and the Spearman_Pval= "+str(round(pvalue_Spearman,3))

# Kendall
tau, p = scipy.stats.kendalltau(N_bleed, N_ulcer)
print "Kendall_R= "+str(round(tau,3))+" and the Kendall_Pval= "+str(round(p,3))
```

Interpret with python2

```
Pearson_R= -0.141 and the Pearson_Pval= 0.646
Spearman_R= -0.022 and the Spearman_Pval= 0.942
Kendall_R= 0.0 and the Kendall_Pval= 1.0
```

The Kendall tau for the couple (N_bleed,N_ulcer) is zero, which leads to affirm that the count of bleeding and the count of ulcers are independent. In Figure 9.4, one can easily see that there is no relation between the count of lesions.

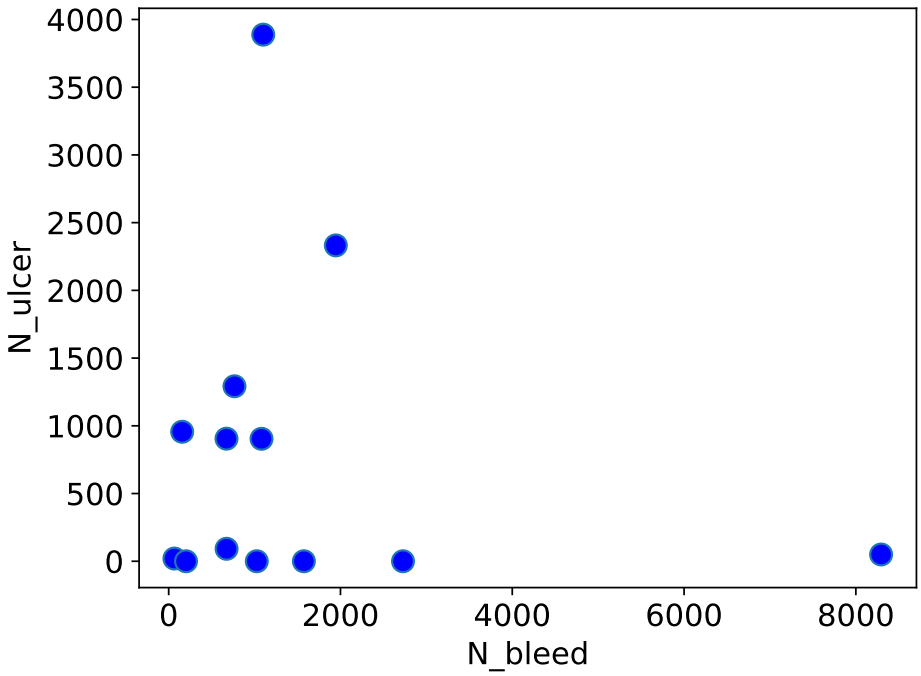


Figure 9.4: Representation of the `N_bleed` and `N_ulcer` for the 13 patients

The total surface of annotated bleeding and ulcers lesions in the colonoscopy video are also not related by studying the correlation indexes and visualising their plot (Figure 9.5).

Code chunk 33: Severity-modelling.py (part 9)

```
Corre_Pearson, pvalue_Pearson = scipy.stats.pearsonr(P_bleed,P_ulcer)
print "Pearson_R= "+str(round(Corre_Pearson,3))+\
    " and the Pearson_Pval= "+str(round(pvalue_Pearson,3))

rho, pvalue_Spearman = scipy.stats.spearmanr(P_bleed, P_ulcer)
print "Spearman_R= "+str(round(rho,3))+\
    " and the Spearman_Pval= "+str(round(pvalue_Spearman,3))

tau, p = scipy.stats.kendalltau(P_bleed, P_ulcer)
print "Kendall_R= "+str(round(tau,3))+\
    " and the Kendall_Pval= "+str(round(p,3))
```

Interpret with `python2`

```
Pearson_R= -0.278 and the Pearson_Pval= 0.358
Spearman_R= -0.353 and the Spearman_Pval= 0.236
Kendall_R= -0.288 and the Kendall_Pval= 0.189
```

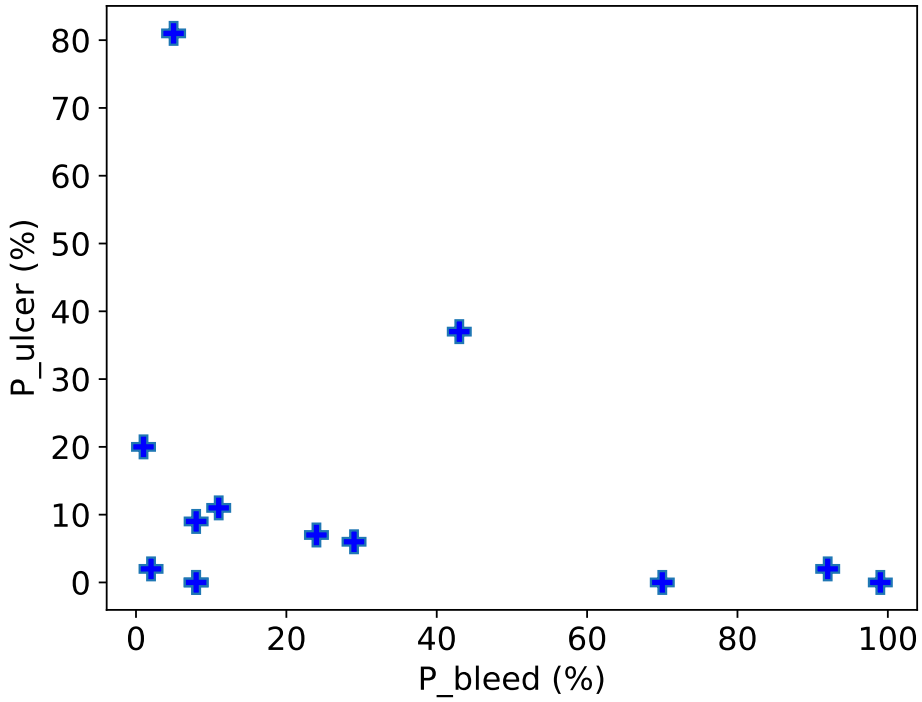


Figure 9.5: Representation of the `P_bleed` and `P_ulcer` for the 13 patients

We conclude that using the classical correlation indexes like the Pearson, Spearman and Kendall, the bleeding and ulcers are not correlated. Consequently, there is no rule to decide the type of model to define the severity of the UC based on the lesions of the colonoscopy video, and therefore the linear model type can be treated/applied.

9.5 Conclusion

This chapter covers trials to verify/prove a set of medical hypotheses or statements based on lesions found in the colonoscopy video. The main parameters were the lesions distribution and the colon parametrization previously introduced. The major difficulty was the formulation of the medical statements due to a lack of a priori medical knowledge. As mathematical techniques, for ease of interpretation, we focus on models of type linear and ordinary differential equation solutions. Due to their simplicity and the limited number of patients considered for his study (i.e. only 13 patients), the obtained results will be unprecise. Consequently, we introduce the notion of compatible models that have performed less than/near 5% in comparison to the best model performance.

We were interested in responding to four statements that were inspected due to the results of previous chapters. Firstly, we considered the bias of lesions distribution in the colon. Indeed, the UC is characterized by lesions that propagate continuously along with the colon beginning with the rectum, but it was not yet proved by a medical study. Secondly, we study the evolution of the severity with the count of lesions using a linear model found in chapter 8. Thirdly, we studied the variability of endoscopic scores between gastroenterologists at evaluating the bleeding and ulcers. Finally, motivated by the results of the lesions maps in chapter 7, we investigated the relationship between the bleeding and ulcer lesions to know if one generates the other.

Through compatible models, we show that the lesions are concentrated at the rectum. However, the bleeding lesions appear to be more abundant between the descent colon and the rectum in comparison to ulcers. The endoscopic scores are likely to increase with the number of lesions found in colonoscopy videos. On the other hand, there is remarkable variation between gastroenterologists while assessing the bleeding severity. Finally, we find that the bleeding and ulcer are not correlated based on statistical criteria Pearson, Spearman and Kendall.

The introduction of the compatible models' concept was helpful to confirm some medical conclusions. Although our works are not yet completely validated, the distribution of lesions appears to be informative for the severity evaluation. A database containing much more patients with variable lesions profiles may lead to more appropriate conclusions about the studied statements.

Chapter 10

Reaction diffusion equations and Ulcerative Colitis disease

Abstract

In this chapter, we are interested in modelling the spatial distribution of ulcerative colitis lesions in the colon.

We introduce a reaction-diffusion equation of type Fisher involving surface of bleeding (or ulcers) as a function of the colon curvilinear abscissa/space. The patient considered in an asymptotic regime (colonoscopy taken at large time t after the beginning of the disease) is considered as a travelling wave solution for Fisher's equation.

Under certain conditions, the asymptotic travelling wave solutions for Fisher's equation are unique up to a translation and dilation depending on the diffusion parameter. We proceed by an inverse problem with respect to the L_2 distance to estimate the compatible dilation with the patient's inflammation. A set of compatible diffusion parameters is extracted. We also compute the set of compatible velocities and the lesion's invasion time is provided by numerical simulations.

Despite the limitations of the proposed model to be realistic, the reported results present the first attempt to model the spatial distribution of UC lesions leading to obtaining a possible prognosis of the disease. On the other hand, these results require medical validation to be used in the medical practice of the UC.

10.1 Introduction

Inflammatory bowel diseases are known to be autoimmune disorders of the digestive tract affecting the mucosa of the gut by inflammation. Their development involves a genetic susceptibility, deregulated immune system's response and an environmental trigger. In general, bowel mucosal homeostasis includes interactions among the microbiota, the intestinal epithelium, and the gut-associated immune system. When a failure in these interactions occurs, an inflammation may precipitate.

Mathematics may help in exploring the consequence of manipulating many parameters involved in the biological phenomenon associated with particular desired scenarios. Indeed, many studies are exploring mathematical models to understand the immunopathogenesis, and diseases arising from an inadequate immune system response as well. In [RBC10], the authors describe a skin disease using mathematical modelling relating the densities of some immune cells such as dendritic and T helper cells to Keratinocytes. In [ML04], the authors analyzed a mathematical model of three non-linear ordinary differential equations for chronic myelogenous leukaemia (CML), cancer affecting the blood cells. In [HJ18], the authors investigate the interaction between tumour cells, $CD4^+$ T cells, cytokines, and normal tissue cells for tumour regression.

In the case of IBDs, multiple genes play a crucial role in the immune system response such as the ones that regulate innate immunity, adaptive immunity, and epithelial barrier function. The Crohn's and ulcerative colitis diseases were identified as Th1 and Th2 mediated processes respectively [BKL11; Mat+04]. The primary function of Th1 cells is to protect the body from intracellular pathogens while the Th2 cells confer protection against extracellular pathogens.

For IBDs, we can cite the work of Yates et al. [YCS04], where a mathematical model of ordinary differential equations is presented to understand the differentiation of naïve T cells following the antigen stimulation into Th1 and Th2 cells.

In 2010 [Wen+10], Wendelsdorf et al. proposed a mathematical model based on a system of 29 ordinary differential equations to understand the complex interactions between cells and genes with bacteria in the colonic lumen involved in the IBDs's physiopathology. They found a positive inflammatory feedback loop by inflammatory M1 macrophage activation of T-cells under the immunopathology of IBD. However, due to the complexity of the proposed system, qualitative properties of cell's interactions are not easy to be extracted or understood.

Later in 2013, Lo and Friedman [LAF13b] extend the work of Yates et al., to a differential equations system including regulatory T cells (Treg cells), and

additional cytokines that are involved in IBDs. Their objective was to study the relationships between Th1, Th2, and Treg cells in the initial phase of the inflammatory process. They show that if the inhibitions of the Treg cells are downregulated, then the Th1 or Th2 response to cytokines induced by macrophages and dendritic cells in response to commensal flora is abnormally high leading to inflammation. Otherwise, when the inhibitions rate of the Treg cells is upregulated, then the cell responses Th1 and Th2 type are abnormally low. The latter case is similar to a bacterial infection. In [Lo+16], the authors expand their previous model to include the Th17 cells pathway which plays an important role in inhibiting Treg cell differentiation that is associated with autoimmune disorders and inflammation.

In the work of Maiti et al. [Mai+15], a system of ordinary differential equations was employed to model the interaction between pro- and anti-inflammatory signaling. The model predictions present a good agreement with experimental data.

In the presented works, a diversity of parameters was considered depending on the desired biological phenomena and the addressed issues such as pro-inflammatory cytokines, anti-inflammatory mediators or density of targeted cells. Most of the obtained results in the referenced papers and the references therein are based on stability analyses of the studied system to find possible equilibrium steady states and supported by numerical simulations for some estimated parameters. In addition, they encounter for sensitivity analysis of some parameters supposed to influence in the biological scenarios like in the work of..

The shared property between mathematical models to IBDs is to study the microscopic level, i.e. cell levels interactions to understand the physiopathology of these diseases. In this chapter, we present a different approach to deal with the spatial distribution of endoscopic lesions observed in colonoscopy videos. We aim to provide a simplified mathematical model framework for the macroscopic level, to understand the long time behavior of the UC lesions in the colon and then be able to estimate the speed of propagation. Our mathematical model is based on the reaction-diffusion equation with a monostable reaction term.

In the previous chapters, we studied the general properties of the UC disease like the map of the spatial distribution of the lesions in the colon (cf chapter 7), the modelling of the endoscopic severity scores in chapter 8 and bias of the distribution of the bleeding and lesions by the notion of compatible models (cf chapter 9). In this chapter, we studied the spatial models of the lesions's distribution. We want to estimate the prognosis of the disease with respect to the present lesions in the colon and their spatial distribution.

Ulcerative colitis is generally characterized by continuous inflammation affecting the innermost lining of the colon. It begins at the rectum and propagates proximally through the colon to stop at a given point of the colon. In Figure 10.1, we represent the surface of annotated bleeding as a function of the colon space represented by its curvilinear abscissa. One can see the mucosa near the cecum (annotated by a green rectangle) is healthy, with no inflammation present ($u = 0$), while the mucosa near the rectum (annotated by a red rectangle) is bleeding and represent the highest level of inflammation ($u = 1$). We will explain later the value u according to the data used in this chapter.

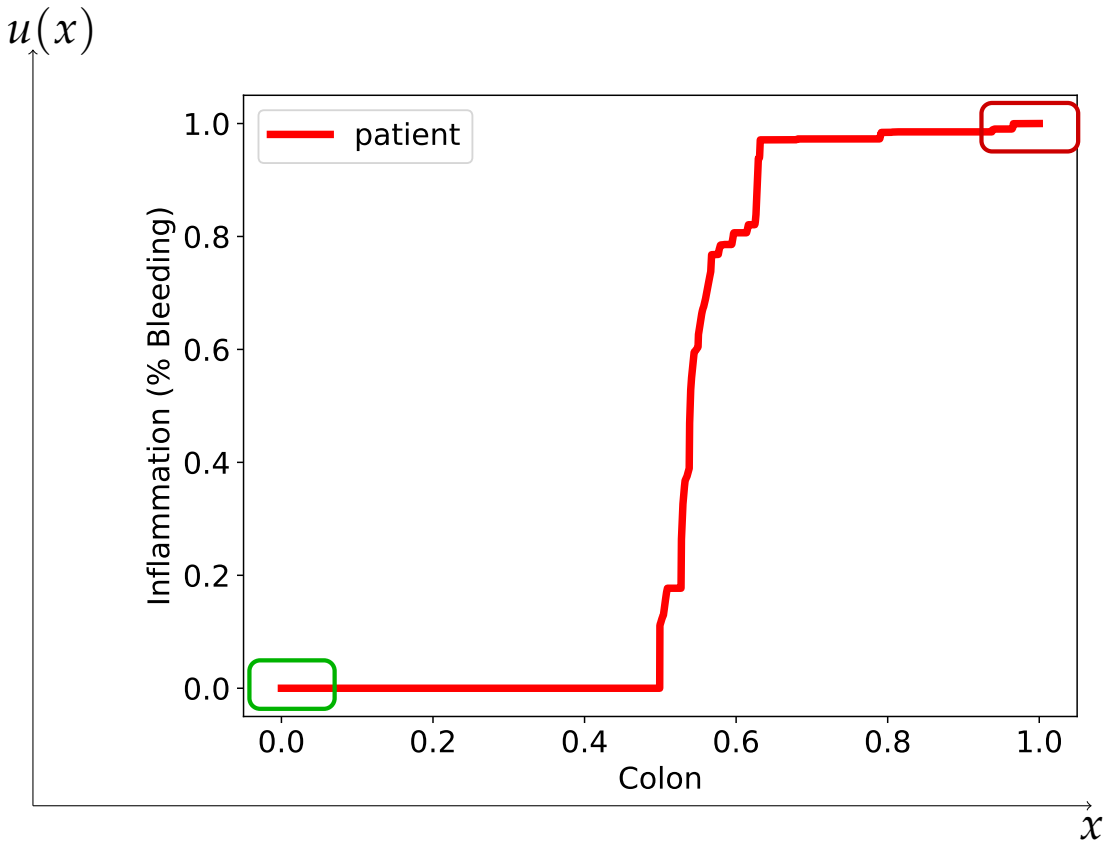


Figure 10.1: The surface of annotated bleeding denoted by u as function of the colon curvilinear abscissa denoted by x.

The barrier point between an inflamed and healthy mucosa is not yet understood and no medical studies have explained it. Therefore, in this study, it will not be taken into account. In addition, we will consider that the colonoscopy video is taken during a period of activated symptoms i.e. the disease is out of control and will continue to spread in the colon. This hypothesis suggests that the inflammation will affect healthy tissue and spread throughout the colon. Such a phenomenon can be seen as a wave front moving with a certain speed. The reaction-diffusion equations are the famous mathematical partial differential equations able to model a spatial motion dependent on the density

and to admit solutions of traveling waves. This could help infer the disease prognosis by estimating how quickly lesions spread to the colon.

The rest of the chapter is organised as follow: section 10.2 is dedicated to summarizing some theoretical properties of the reaction-diffusion equations and possible applications for the dynamic populations. In subsection 10.2.3, we describe the travelling wave solutions for reaction-diffusion equations, and in particular for the Fisher equation. We present the mathematical model in section 10.3 and the inverse problem to estimate the parameters in section 10.4. The computation of the speed of propagation of the inflammation is provided as well. The numerical results for patients of Vatic database are given in section 10.5. We conclude the chapter with some possible extensions.

10.2 Reaction-diffusion equations

10.2.1 Some known properties

The mathematical models of reaction-diffusion type are widely used to describe the Spatio-temporal spreading phenomenon generated by a substance or population which can be a virus, a special type of cells, inflammation, dominant genotype that spatially diffuses by random walks, Brownian motion, hydrodynamic turbulence, or other mechanisms, reacting to each other and their surroundings and then affecting their local densities and dispersion [Kri08]. They are utilized in many areas of real-world application like dynamical populations, chemistry [Tur90], genetic [BT11] and combustion theory [Bar+85].

The reaction-diffusion equations are adapted to ecological phenomenons such as the formation of a spatial pattern. In this sense, we recall the paper of Alain Turing [Tur90] who show that any repeating phenomena/pattern in nature could be a result of the interaction between two species such as cells or molecules through specific conditions. He proposed the reaction-diffusion equations and demonstrate that the two species spread at different velocities to produce patterns. The faster one is called the inhibitor while the slower one is called activator. Such kind of nature patterns can be seen in zebra's stripes, leaopard's spots and wood spirals (cf Figure 10.2).

Whereas a special class of solution for this type of equation, namely the travelling wavefront solutions are useful to model biological invasions [AM17] or cardio-vascular diseases such as atherosclerosis [VP09]. The class of travelling wave solutions for the reaction-diffusion equations are of great interest in this work, we will discuss it in subsection 10.2.3.

Figure 10.2: Examples of nature patterns that can be modeled by the reaction-diffusion equations. First row: zebra patterns and woods. Second row: propagation of the Bubonic plague in europe during the 14th century.

The general form of a reaction–diffusion equation can be written as follow:

$$\frac{\partial}{\partial t}u - D\frac{\partial^2}{\partial x^2}u = f(u,t,x), \quad t \geq 0,\ x \in \mathbb{R}^n \tag{10.1}$$

with some initial $u_0 = u(0,x)$, and boundary conditions $u(t,-\infty), u(t,+\infty)$.

The system parameter $u(t,x)$, a function of two real variables, the time $t \geq 0$ and the spatial coordinate $x \in \mathbb{R}^n$, describes the density of a substance, population, viruses or another type of species.

The term $\frac{\partial}{\partial t}$, also denoted by ∂_t represents the partial derivative with respect to time t, and $\frac{\partial^2}{\partial x^2}$, also denoted by Δx is the spatial Laplacian, is the sum of the second order partial derivatives with respect to space $\frac{\partial^2}{\partial x_i^2}$ with $x = (x_i), i \in \{1, ..., n\}$. D and f are the diffusion matrix and the reaction terms respectively.

The term of diffusion D, of units of length squared over time, is crucial for describing the diffusion mobility of the population or the substance u at a position x and time t. Diffusion is used to represent the resulting movement from an object or body making many short movements in random directions or the same direction of movement [AM17]. The higher the value of this term, the faster the diffusion is.

When u is a scalar quantity, the Equation 10.1 can be represented by a single diffusion-reaction equation of the form:

$$\frac{\partial}{\partial t}u - D\frac{\partial^2}{\partial x^2}u = f(u, t, x), \quad t \geq 0, \ x \in \mathbb{R} \tag{10.2}$$

Many cases can arise depending on the nature of the reaction term f and the model parameter u. In the case of u and f do not depend on x, we get a simple system of ordinary differential equation that it is an asset in a temporally heterogeneous media:

$$\partial_t u = f(u, t)$$

In the case of u and f are time-independent, we get a so-called system of weakly coupled elliptic partial differential equations (PDEs) that is active in a spatially heterogeneous media:

$$-D\Delta_x u = f(u, x)$$

When f only depends on u, the system is referred to as an asset in a homogeneous media.

Several types of reaction term were employed in literature for reaction-diffusion systems and which can be classified into (cf Figure 10.3):

- Monostable if $f(u) > 0$ for $0 < u < 1$
- KPP type if f is monostable and $f(u) \leqslant f'(0)u$ for $0 < u < 1$
- Ignition type if for some $0 < \theta < 1$, the term f satisfies

$$\begin{cases} f(u) = 0, & for \quad 0 \leqslant u \leqslant \theta \\ f(u) > 0, & for \quad u > \theta \end{cases}$$

- Bistable, if sor some $0 < \theta < 1$

$$\begin{cases} f(u) < 0, & for \quad 0 \leqslant u \leqslant \theta \\ f(u) > 0, & for \quad \theta < u < 1 \end{cases}$$

In population dynamics, common choices for non linearity are $f(u) = ru$(linear growth), $f(u) = (u)ru(1 - uK)$ (logistic growth), or $f(u) = ru(u - a)(1 - uK)$ with $a \in (0, K)$ (growth with Allee effect).

For systems, typical reaction terms are those that occur in non-spatial models for competition, mutualism, or predator-prey interactions. Those include Lotka-Volterra models, but also more general models such as predator-prey models with functional response. In the case of systems, the stability analysis often involves the eigenvalues of matrices obtained by linearising the equilibria. Equilibria and eigenvalues play a similar role in the analysis of reaction-diffusion models, but the eigenvalues generally are associated with differential operators rather than matrices.

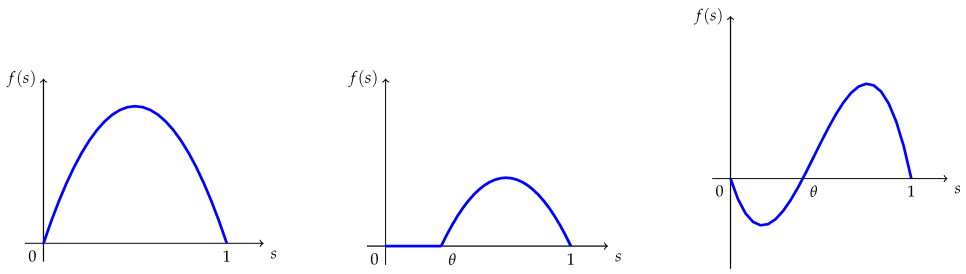


Figure 10.3: Monostable, bistable and ignition type non linearities

In general, stationary states are of interest, corresponding to the absence of modification/change in population size over time [Kut11]. If the diffusion terms (i.e. spatial effects) are ignored, then the PDEs of Equation 10.1 are ODEs,

$$\partial_t u = f(u, t) \tag{10.3}$$

where f can take various forms to describe the time evolution of the studied population. Typical choices are $f(u) = au$ where a is the growth rate leading to an exponential growth, $f(u) = au(1 - \frac{u}{K})$, where K is the carrying capacity to model a logistic growth of the population. In addition, $f(u) = au(\frac{n}{u_0} - 1)(1 - \frac{n}{K})$, which is based on a logistic growth but if the population is too small, it will die out. Such a phenomenon may appear due to the necessity to find a mate for reproduction or to defend the group against predators. This leads to an additional term $(\frac{n}{u_0} - 1)$.

In the next section, we wil present a classical non-linear reaction-diffusion equation known as Fisher Kolmogorov-Petrovski-Puskinov equation that will be used to model the spatial distribution of bleeding and ulcers in this chapter.

10.2.2 Fisher Kolmogorov-Petrovski-Puskinov equation

In 1930, Fisher proposed to model the spatial diffusion of an advantage gene such as a wave in one-dimensional habitat [Kol37]. Kolmogorov, Petrovski, Puskinov was the first to study it mathematically. The Equation 10.2 is called Fisher Kolmogorov-Petrovski-Puskinov equation, abbreviated Fisher-KPP, if the non linearity term f is monostable i.e. $f(u) = ru(1-u)$ and only depends on the population density u. The Fisher-KPP equation is given by the following system:

$$\frac{\partial}{\partial t}u - D\frac{\partial^2}{\partial x^2}u = ru(1-u), \quad t \geq 0,\ x \in \mathbb{R} \tag{10.4}$$

where r represents the growth rate of the population under study (like cell growth rate or birth and die rate..) and $f(u)$ statifies the following conditions:

$$f(0) = f(k) = 0 \quad \text{for some} \quad k > 0; \quad \text{and} \quad f(u) > 0 \quad \text{for} \quad u \in (0,k).$$

In this work, we consider the parameter r is one all the time and the Equation 10.4 remains as follow:

$$\frac{\partial}{\partial t}u - D\frac{\partial^2}{\partial x^2}u = u(1-u), \quad t \geq 0,\ x \in \mathbb{R} \tag{10.5}$$

Remark: A solution u for the Fisher-KPP equation given by the Equation 10.5 can be represented by the couple (D, u_0) where D and $u_0 = u(x, t = 0)$ denote the diffusion parameter and the initial spatial condition for time $t = 0$ respectively.

In this work, we decide to use the solutions of the Fisher-KPP equation to model the spatial distribution of the UC lesions in the colon and then predict the disease evolution for the patient using the compatible models with the observations, which are the colonoscopy video. Therefore, the lesions spatial model will based on two parameters: D and u_0.

In the next section, we will explore an interesting class of solutions for the Fisher-KPP equation, kwown as the travelling waves.

10.2.3 Travelling Wave solutions (TW)

Travelling waves (TW) are PDEs solutions moving in a particular direction with constant speed c while maintaining their shape over their propagation axis (see Figure 10.4). They play an important role in describing the long time behaviour

of solutions to initial value problems modelled by reaction-diffusion equations. They connect two steady states of the PDEs system.

Travelling waves are observed in many areas such as a result of some chemical reactions like combustion [VVV94], propagation of flames, migration of biological species, tumour growth, Schock profiles in fluid dynamics[Smo12] and many other applications.

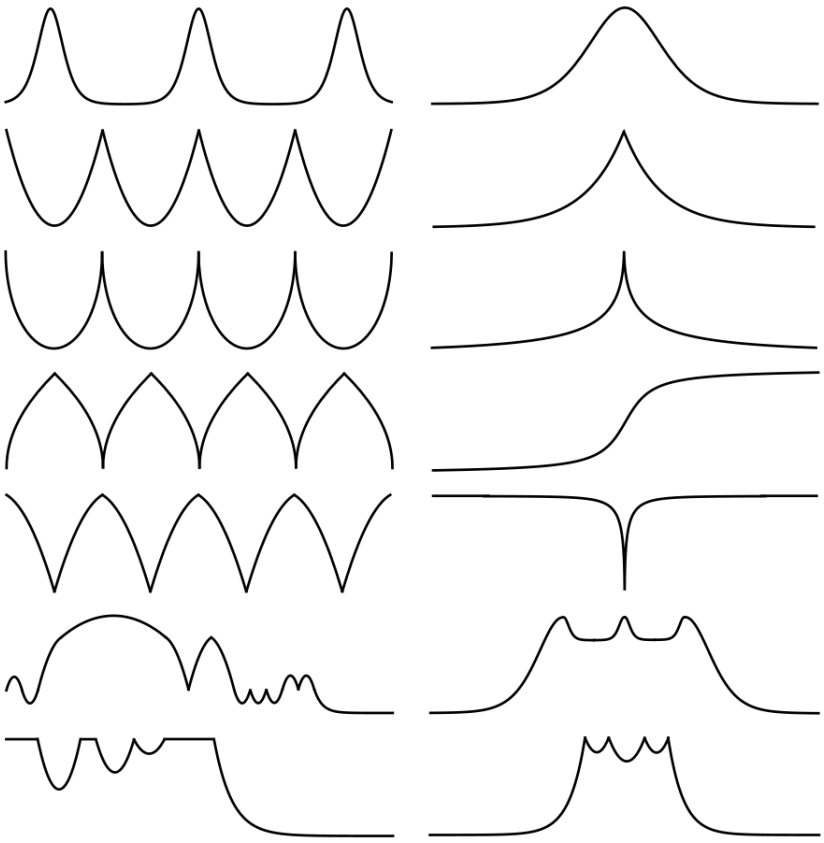


Figure 10.4: A selection of some travelling waves. The waves on the left side from top to bottom are of the following types: smooth periodic, peaked periodic, cusped periodic, periodic with peaked crests and cusped troughs, periodic with peaked crests and troughs, composite, composite with plateaus. Right side top to bottom: smooth solitary, peaked solitary, cusped solitary, wavefront, compactly supported anticusp, multi-crest with decay, multi-peak with decay [GQ18].

From mathematical point of view, we can define the travelling wave solutions for a partial differential system like the Equation 10.1:

Definition 3. *[Per15] A travelling wave solution is a solution of the form* $u(t,x) = v(x-ct)$ *with* $c \in \mathbb{R}$ *a constant called the wave speed or the traveling speed. We say it connects the states 1 and 0 if* $v(-\infty) = 1, v(\infty) = 0$. *The function* v *is the wave profile will describe the evolution of the system between two*

steady states and c the wave velocity.

In particular, when the velocity c is equal to zero then we get standing waves that do not move at all [San02].

There are many types of travelling waves with some of them can be given by the following definition:

Definition 4.

- *A travelling wave solution v is called* ***Wave front*** *if $v(-\infty) \neq v(+\infty)$*
- *A travelling wave solution v is called* ***Pulse wave*** *if $v(-\infty) = v(+\infty)$ and $v(\xi)$ is not a constant function*
- *A travelling wave solution v is called* ***Wave trains*** *if the wave profile is periodic*

In this part of the manuscript, we seek wavefront solutions because the inflammation value at the beginning of the colon (at the cecum) is in general different from its value at the rectum.

In the case of reaction diffusion equation with monostable reaction term, and more precisely for the Fisher-KPP equation, there esxits a travelling wave solutions for which we can compute the movement's velocity. In the first work of Kolmogorov et al. [Kol37], it has been demonstrated that all initial conditions with compact support for Equation 10.5 in the case of $D = 1$ evolve to a travelling wave solution with the minimum speed $c = 2$ [SM96].

For a general case of the diffusion parameter, we refer to [MZ05] for the existence of a travelling wave. Let denote by $c_{min} = 2\sqrt{D}$ the minimal speed satifiying [MZ05]:

- For every $c > c_{min}$ there exists a travelling wavefront of the form $u(x,t) = v(x+ct)$ with $v(s)$ increasing and $v(-\infty) = 0, u(+\infty) = k$
- For $c \leq c_{min}$, there is no such monotone wavefront with speed c
- The wavefront is unique up to translation
- The wavefront is stable according to some class of initial functions such as compactly support functions.

FKPP equation exhibits only two non negative equilibrium [Per15]: $u = 0$, unconditionnally unstable (i.e. for any small initial perturbation close to zero, the solution will grow exponentially) and $u = 1$ (i.e for any small perturbation near 1, the solution will relax exponentially to 1).

The Fisher-KPP equation is unvariant under change of sign of the spatial coordinate x, therefore one can say:

Proposition 1. *Fisher-KPP equation admits a travelling wave front solution of the form $u(t,x) = v(x+ct)$ with $c \in \mathbb{R}$, with respect to the boundary conditions $v(-\infty) =$*

$0, v(+\infty) = 1$

Let take the Fisher-KPP equation in the case of $D = 1$, we introduce the co-moving frame $\xi = x + ct$ (travelling wave coordinate transformation), and we denote by c the dimensionless wave speed and we substitute u by v in the Fisher KPP equation, we obtain the following ordinary differential equation on v:

$$\begin{cases} -v''(\xi) + cv'(\xi) = f(v(\xi)); \quad x \in \mathbb{R} \\ v(-\infty) = 0, v(+\infty) = 1, \quad \text{boundary conditions} \end{cases} \tag{10.6}$$

The explicit expression of the travelling wave solutions of the Equation 10.6 is not kown except for particular velocity values [MF96; BD15]. In case of Equation 10.5 with $D = 1$, Ablowitz and Zeppetella [AZ79] computed the explicit form of the solution, given by $u = [1 + c \exp(\frac{(x-ct)}{\sqrt{6}})]^{-2}$ for the velocity value $c = \frac{5}{\sqrt{6}}$. Many published works studied the existence and stability of such particular solutions. For theorectical review, see for example references [Per15; TV20; VP09].

Let u be a Fisher-KPP solution with $D = 1$ and initial condition $u_0 = u(x, 0) \ll x^\alpha e^{-x\sqrt{f'(0)}}$, for $\alpha \leq -2$, and u_0 is zero starting from a certain value. Hence, the asymptotic behavior of u is a wave front u^A independent of the initial condition u_0, moving with velocity $c = 2\sqrt{f'(0)} = 2$ (see [BD15]).

Therefore, for general value of D, we can say:

Proposition 2 (Asymptotic behavior). *The wave front profile v is independent of u_0, and it only depends on the value of the diffusion parameter D. The asymptotic behavior of u is a front wave moving with the velocity $2\sqrt{D}$*

Example: We simulate the Fisher-KPP equation for $x \in X = [-20, 20]$ with space step $\Delta_x = 0.1$ and time $t \in [0, 6]$ with step $\Delta_t = 0.002$ and diffusion parameter $D = 2$ with different initial conditions (represented in black in the Figure 10.5):

- Case 1: $u_0(x) =$
- Case 2: $u_0(x) = 1_{[7:20]}$

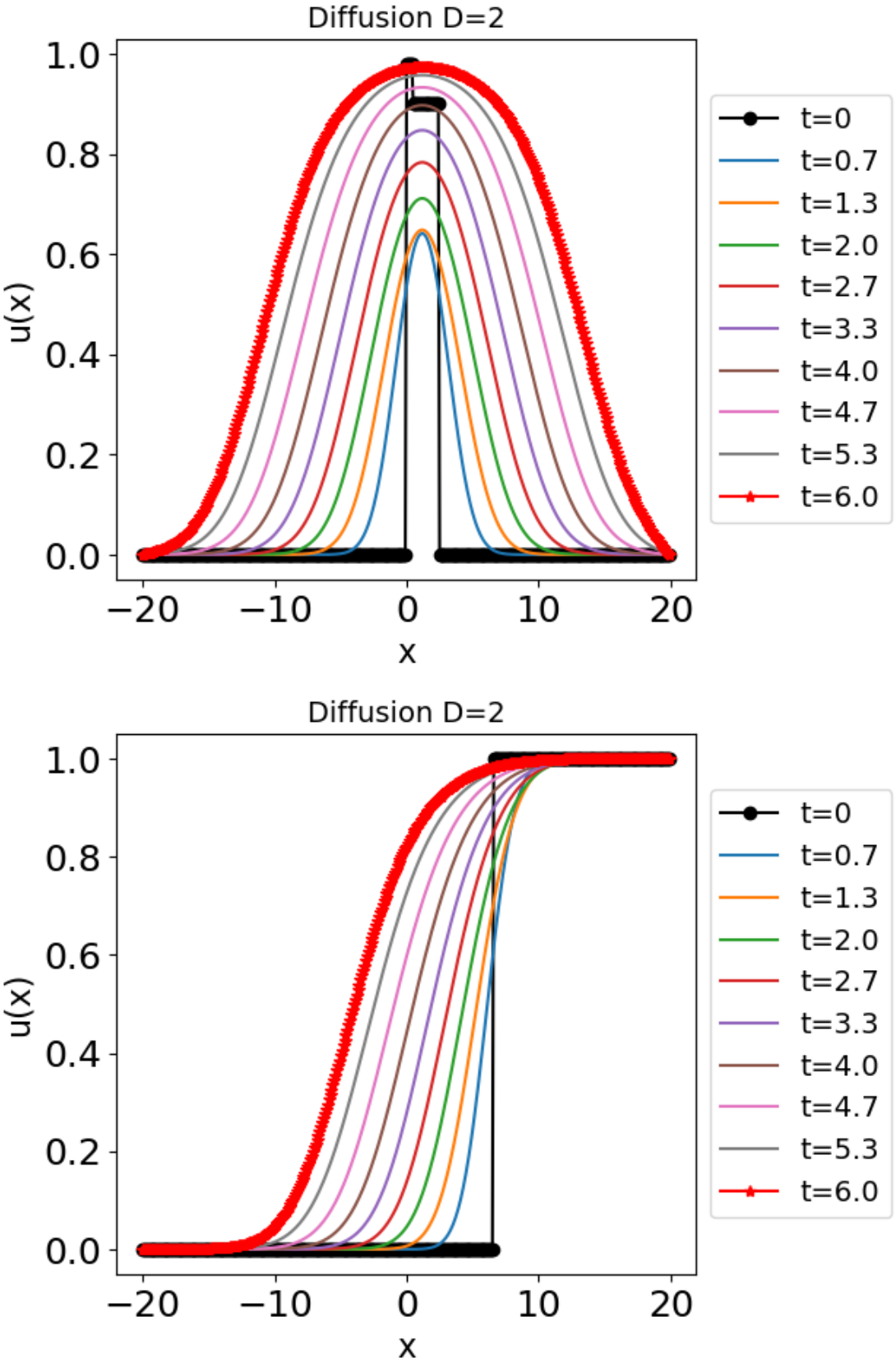


Figure 10.5: Simulation of Fisher-KPP solution with two different initial conditions u_0 compactly supported for the same value of the diffusion parameter $D = 2$.

From Figure 10.5, we can conlude that for all initial conditions u_0 decreasing fast enough, there exists travelling wave fronts. If u_0 is with compact support, there exit two wave fronts: the first from right and the second from left. If $u_0(0,t) = 0$ and $u_0(L = 1, t) = 1$, only one left moving wave front appears [BD15].

Example: Another example, we simulate the Fisher-KPP equation for $x \in X = [-20, 20]$ with space step $\Delta_x = 0.1$ and time $t \in [0, 6]$ with step $\Delta_t = 0.002$ and different diffusion parameters $D_1 = 2$ and $D_2 = 4$ with different initial conditions (represented in black in the Figure 10.6):

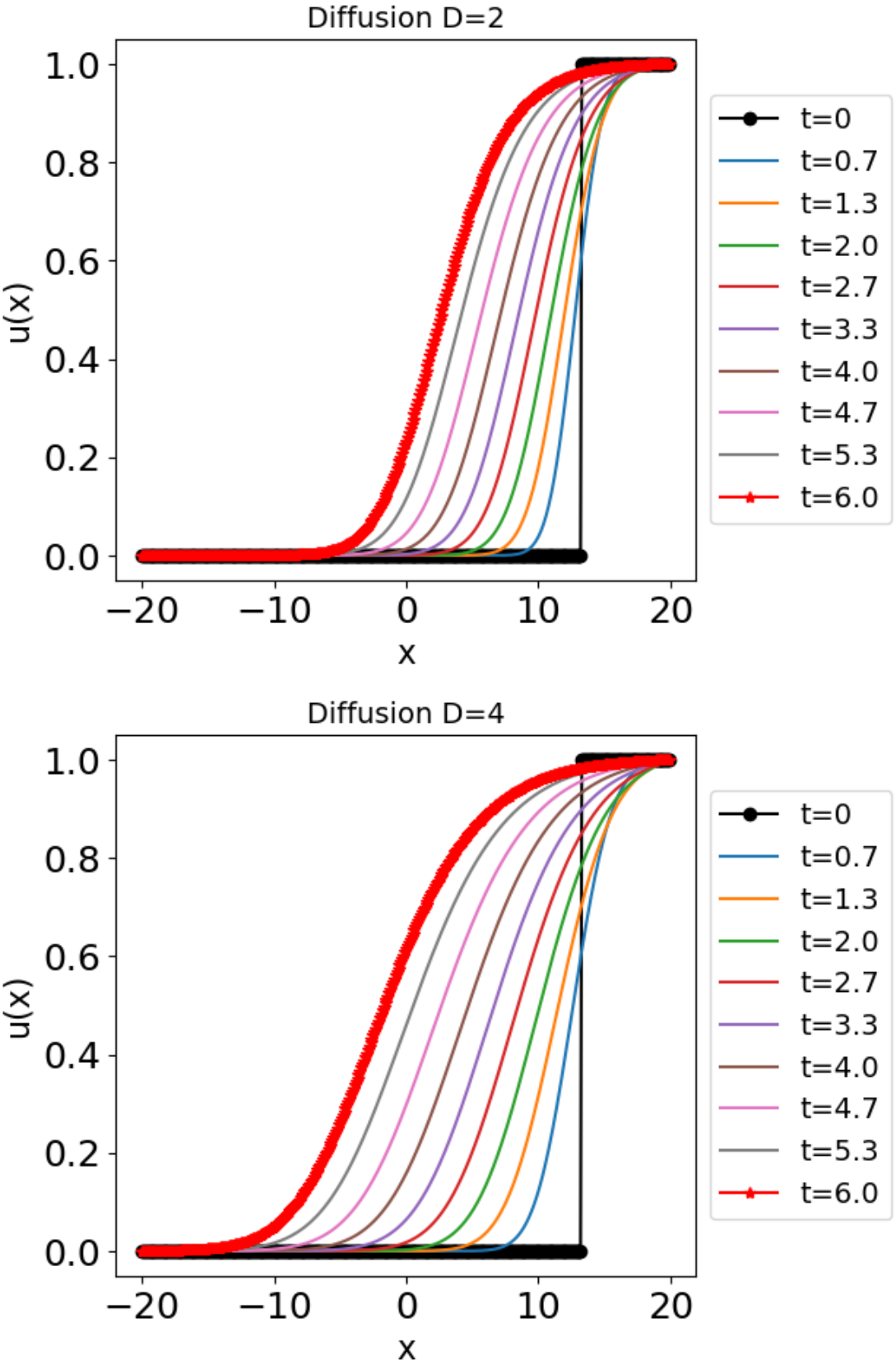


Figure 10.6: Simulation of Fisher-KPP solution with initial condition u_0 compactly supported (black) with two different diffusion parameters.

Proposition 3. *The FKPP solution (D, u_0) for compact support initial condition u_0 can be obtained by spatial scale changing for FKPP solution with diffusion parameter $D = 1$ i.e. $(1, u_0(\frac{1}{\alpha}x))$ where $\alpha = \sqrt{D}$.*

Theorem 10.2.1. *Let denote by S, the space of TW solutions for the FKPP equation with initial condition u_0 with compact support. S is a one dimensional space*
In this manuscript, we are rather interested in the application of FKPP equation to model the long time behavior of UC disease, or explicitly we need to get a numerical estimation about the inflammation in the colon: if it will continue progressing in the colon (new parts of the colon will be affected) or regressing.

10.3 The mathematical model

In this section, we will exhibit the modelling of the spatial distribution of UC lesions as solution of the Fisher-KPP equation in one dimensional space for time and space. We decide to model the inflammation of the UC disease according to the bleeding and ulcer lesions for which the information is available from our data-base Vatic. The model reads:

$$\frac{\partial}{\partial t}u(x,t) - D\frac{\partial^2}{\partial x^2}u(x,t) = u(1-u), \quad t \geq 0,\ x \in \Omega = [0,1] \in \mathbb{R} \tag{10.7}$$

We will consider the Dirichlet boundary conditions $u(0,t)$ and $u(1,t)$, and initial condition-data $u(0,x) = u_0(x)$ for all $x \in \Omega$.
The density parameter u can be the surface of annotated lesions (bleeding or ulcers) or the surface of the detected pixels by the work of the chapter 6. The spatial coordinate x denotes the colon's curvilinear abscissa denoted by Ω like the parametrization used in chapter 7. The coefficient $D > 0$ is the diffusion rate of the lesions in the colon.

In the previous section, we mentionned that the Fisher-KPP solution is parametrized by the couple (D, u_0). To simulate numerically the Fisher-KPP solution, we decide to use the finite difference method, and more precisely the Forward in time and centered in space numerical scheme (FTCS).

In order to distinguish between space and time coordinates, superscript index n is used for time coordinate where as a subscript i is used to represent the space position i.e. the value of variable u at time $t = n$ and position $x = i$ will be written as $u_i^n = u(x = x_i, t = n)$, $x_j = j \times \Delta_x$, for $i \in 1, .., N$, supposing that we discretise the space domain on N equal points.

The explicit time derivative formula:

$$u_t = \frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta t} + O(\Delta t) \tag{10.8}$$

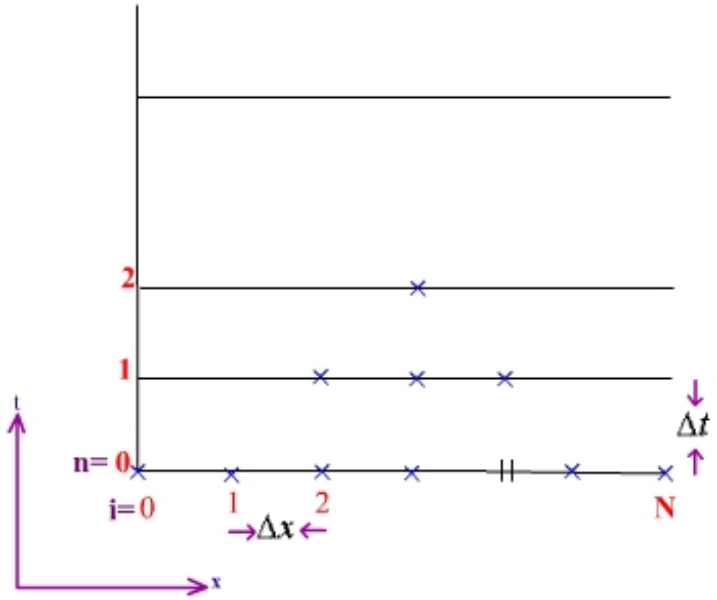


Figure 10.7: Space and time discretization for FTCS numerical scheme

And the centered space derivative (Taylor):

$$u_{xx} = \frac{u(x_{i-1}, t_j) - 2u(x_i, t_j) + u(x_{i+1}, t_j)}{\Delta x^2} + O(\Delta x^2) \tag{10.9}$$

We suppose that we want to solve the reaction diffusion equation in a one dimensional domain i.e. $\Omega = \{x \in \mathbb{R}, 0 < x < L = 1\}$ and time interval $\{0 < t < T, T \in \mathbb{R}^*\}$. And, the initial and boundary conditions are written as follow:

$$\begin{aligned} u(x,0) &= g(x), \text{ initial space distribution} \\ u(0,t) &= g_1(t) \\ u(L,t) &= g_2(t) \end{aligned}$$

So the new discretise form of the reaction diffusion in Equation 10.2 will be:

$$\begin{aligned} \frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta t} = D \times \frac{u(x_{i-1}, t_j) - 2u(x_i, t_j) + u(x_{i+1}, t_j)}{\Delta x^2} \\ + f(u(x_i, t_j)) + O(\Delta t + \Delta x^2) \end{aligned} \tag{10.10}$$

And, the estimated value of u at any node/point x_i and time step $t+1$ will be formulated by:

$$\begin{aligned} u(x_i, t_{j+1}) = u(x_i, t_j) + \frac{D\Delta t}{\Delta x^2}(u(x_{i-1}, t_j) - 2u(x_i, t_j) + u(x_{i+1}, t_j)) \\ + \Delta t \left(u(x_i, t_j) - u(x_i, t_j)^2\right) \end{aligned} \tag{10.11}$$

or also

$$u_i^{j+1} = u_i^j + \frac{D\Delta t}{\Delta x^2}(u_{i-1}^j - 2u_i^j + u_{i+1}^j) + \Delta t\left(u_i^j - (u_i^j)^2\right) \tag{10.12}$$

From equation (10.12), the FTCS scheme is classified as explicit because the value of u_i^{n+1} at the $(n+1)$th time level may be calculated directly from known value of u_i^n at previous time levels.

It is well-known that propose scheme approximates the continuous operator to order $O(\Delta_t + \Delta_x^2)$. At the space position x_n when $n = 1$, the unknown variable u is first calculated using the initial conditions at time $t = 0$ and boundary conditions applied to values $x = 0$ and $x = L$. Once the solution at time level t_1 $(= t_0 + \Delta_t)$ is obtained, the solution at level $n = 2$ is calculated in the same manner by making use of the solution at $n = 1$ and the boundary conditions at $x = 0$ and $x = L$. The same procedure is repeated until the solution reaches a steady state or until the required time maximum level.

Let denote $CFL = D\frac{\Delta_t}{\Delta_x^2}$, the Courant-Friedrichs-Levy also called CFL condition, named according to researchers who first described this requirement. CFL required to be bounded by a constant depending upon the particular numerical scheme. Due to absence of better method to get an approximate difference formula, the only requirement is that the obrained differentiate formula, must pass/justify some tests for accuracy, consistency, stability and convergence [AM17]. By Von Neumann stability analysis, the FTCS scheme is always conditionally stable, for $0 < CFL \leq \frac{1}{2}$.

Therefore, if we consider the colonoscopy video as the initial distribution of the lesions i.e. the colonoscopy taken at time $t = 0$ of the disease, we can simulate the Fisher-KPP with a test value of the diffusion parameter $D = 100$ by the following code chunk **??**:

Code chunk 34: fisherkpp

```
L=len(Data_perc_b[23])
x0, xL, delta_x = 0, L-1, 1
t0, tF, delta_t = 0, 50, 0.003
Nx, Nt = int((xL-x0)/delta_x+1), int((tF-t0)/delta_t+1)
X, T = np.linspace(x0,xL,Nx), np.linspace(t0,tF,Nt)

# initial condition
U1=np.zeros((Nx,Nt))
U1[0,:] = 0.0
# at position x=0 for all times
U1[L-1,:] = 1.0
# at position x=L-1 for all times
U1[:,0]= Data_perc_b[23] / max(Data_perc_b[23]) # at t=0

# simulation code
D=100 # Diffusion parameter
CFL=D*delta_t/delta_x**2
for t in range(Nt-1):
    U1[1:-1,t+1] = CFL*U1[2:,t] + CFL*U1[:-2,t] + \
                (-2*CFL + 1 + delta_t)* U1[1:-1,t] - delta_t*U1[1:-1,t]**2

# figure code
_ = plt.figure()
_ = plt.xlabel('Frame number',fontsize=14)
_ = plt.ylabel('Surface of Bleeding',fontsize=14)
ts = np.floor(np.linspace(0,Nt-1,10))
for t in ts.astype(int):
    roundedt = round(T[t],1)
    _ = plt.plot(X,U1[:,t], label='t=%s'%roundedt)

_ = plt.legend()
_ = mpl.rc('xtick',labelsize=16)
_ = mpl.rc('ytick',labelsize=16)
_ = plt.rc('legend',fontsize=14)
_ = plt.tick_params(axis='both', which='minor', labelsize=14)
plt.savefig('test-patient-23.pdf',bbox_inches = 'tight')
plt.close()
```

Interpret with `python2`

```
```

We then compute the velocity at which the inflammation is moving using the code chunk:

Code chunk 35: fisherkpp (part 2)

```
U1_trapz = np.trapz(U1,dx=delta_x,axis=0)
len(U1_trapz) # should be equal to nb of time points, here 16667

import scipy
from scipy import stats
lm = stats.linregress(T[16000:],U1_trapz[16000:])
print(lm.slope,lm.rvalue)
print 'Velocity of the propagation = '+str(lm.slope)# should be 2.sqrt D = 20
(U1_trapz[-1] - U1_trapz[-2]) / (20*delta_t) # should be 1
```

Interpret with python2

```
16667
(19.65732103691607, 0.9999999967132336)
Velocity of the propagation = 19.65732103691607
0.9832079143128188
```

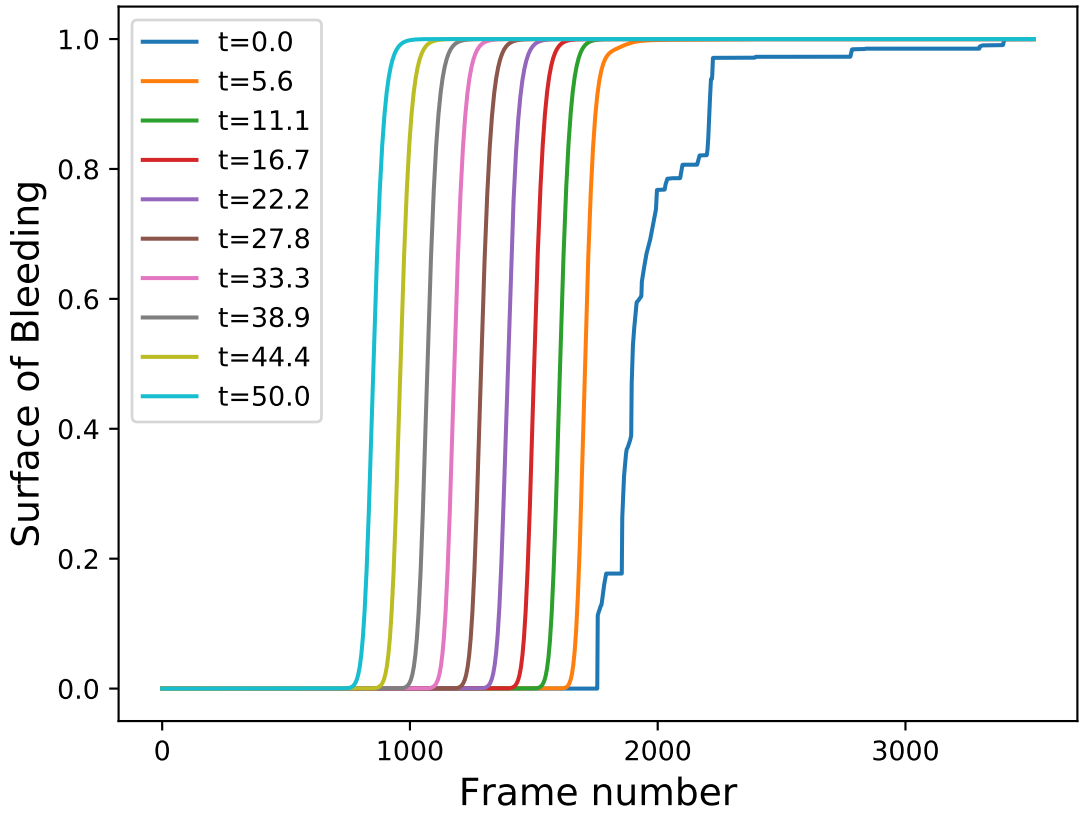


Figure 10.8: Simulation of the bleeeding lesions's evolution as solution Fisher-KPP model with a diffusion parameter $D = 100$. A travelling wave front appears with movinf velocity $c = 2\sqrt{100} = 20$.

The symptoms of the UC disease appear after some time $t > 0$ from the beginning of the disease. Thus a patient $p(x)$ is in general observed at time t which is not the initial time of the disease. In other words, the colonoscopy

video is in general taken after a time $t > 0$ that is not the begining of the disease denoted by T_{initial}. On the other hand, the biological's laws modeled by the Fisher-KPP equation can be applied at any time of the patient's life. The inflammation due to the disease heve evolved during an unknown time t before the patient have made the clinical examinations.

??? Therefore, if we consider that the patient is observed at time $t = 0$ of the disease, the problem will be ill-posed. Due to the fact, that only one colonoscopy video is available per each patient, we are conducted to consider the initial spatial distribution $u_0(x) = p(x)$ and D unkown parameter.

The aim of the modelling process is thus to find a model (D, u_0) such that there is a time t for which the Fisher-KPP solution fits with the inflammation $p(s)$ observed during the colonoscopy. If we consider that tis time is large enough, namely that the patient is observed in the asymptotic regime.

In Proposition **??**, we mentionned that the asymptotic behavior of a Fisher-KPP solution with initial compactly supported condition u_0 ($u_0(x) = 0$ when $x \to -\infty$ and $u_0(x) = 1$ when $x \to +\infty$) is a travelling wave moving with velocity depending on the parameter D. As the lesions begin from the rectum before propagating throught the colon, and the colon's parametrization considered in this thesis, we can take $u(0,t) = 0$ and $u(1,t) = 1$ for any time t. In this case, we showed that the asymptotic behavior of the model (D, u_0) can be obtained by a dilation of the asymptotic solution u^A obtained in the case of $D = 1$.

To sum up, the space of models is reduced to the set of travelling wave fronts of the form $\sqrt{D}.u^A$. If we test a set of random values of $\sqrt{D}$, namely dilations, we are ..

10.4 Inverse problem

We can estimate the set of compatible dilations that are compatible with the observations according to the notion introduced in chapter 9:
Proposition 4.

$$\begin{aligned} D^* &= \underset{D \in S_D}{\operatorname{argmin}} J(D) \\ &= \underset{D \in S_D}{\operatorname{argmin}} \|u_p - u^A_{\sqrt{D},T}\|^2, \end{aligned} \tag{10.13}$$

where S_D is the set of tested diffusion parameters D and u_p is the inflammation data for a patient p.

Using the compatible notion of the chapter 9, we can compute a set of compatible dilation parameters with α error tolerance as follow:

Proposition 5. $D_{comp} = \{D \in S_D, \|u_p - u^A_{\sqrt{D},T}\|^2 \leq 1.1 \times \|u_p - u^A_{\sqrt{D^*},T}\|^2\}$, *where S_D is the set of tested diffusion parameters D and u_p is the inflammation data vector for a patient p.*

Final, we can compute the set of the compatible velocities:

Proposition 6.

$$v_{comp} = \{\mathbf{2}\sqrt{\mathbf{D}} \textit{ for } D \in D_{comp}\}$$

10.5 Numerical results

We applied the proposed Inverse modelling process to some patients of the data base Vatic, and we used the

10.5.1 Asymptotic travelling wave

We estimate the asymptotic travelling wave for Fisher-KPP equation for the space interval $[-10 : 60]$ with spatial step $\Delta_x = 0.1$, and for time discretization we consider the interval $[0 : 150]$ with time discretization step $\Delta_t = 0.001$.

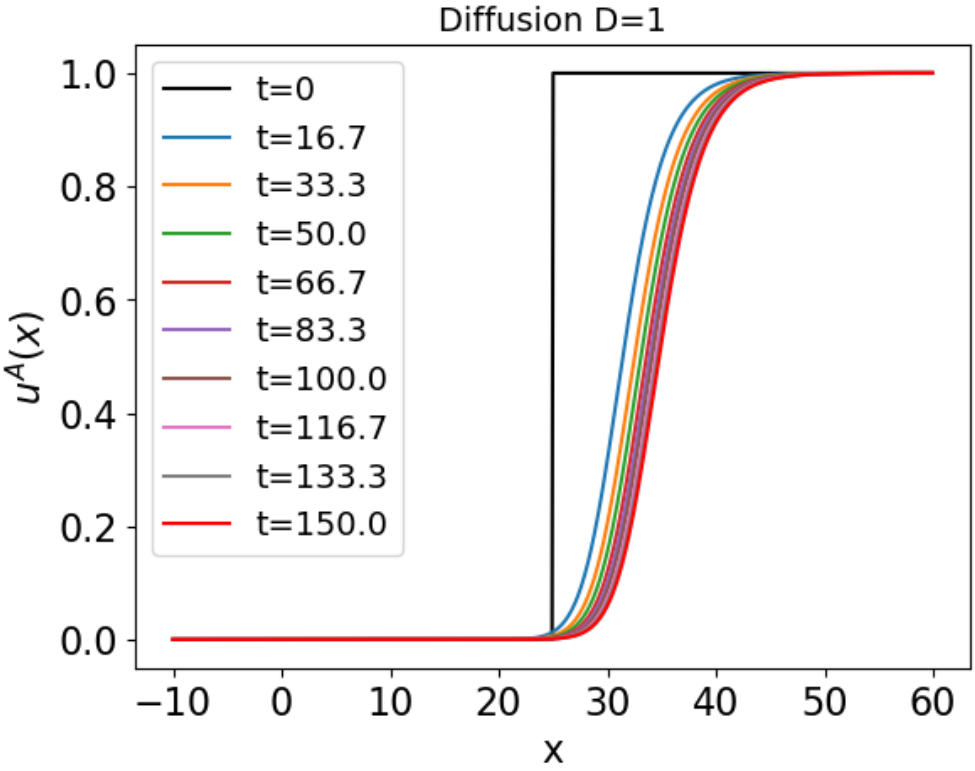


Figure 10.9: Simulation of the asymptotic travelling wave solution for Fisher-KPP equation in the case of $D = 1$

10.5.2 Application to Vatic data base

Wa can estimate the translation value by supposing that the integrals of the patient function (red plot) and asymptotic TW FKPP solution for $D = 1$ (blue dashed plot) are equal. Many numerical analysis methods as trapezoidal, rectangle , simpson rules can be applied to approximate integrals of functions. In our simulation, we used the trapezoidal rule [SM03] that already exists in python library.

Code chunk 36: fkpp_recalage (part 2)

```
def fkpp_interp(p,D,figname=None):
    Xp = np.linspace(0,1,num=len(p))
    sqrtD = np.sqrt(D)
    Xi = sqrtD * X \
        + (np.trapz(U1,dx=delta_x*sqrtD) \
        - np.trapz(p,dx=Xp[1]-Xp[0]) - (xL*sqrtD-1))
    Ui = np.interp(Xp,Xi,U1)
    distD = np.trapz( (p-Ui)**2,dx=Xp[1]-Xp[0] )
    if figname != None:
       fig,ax = plt.subplots()
       ax.plot(Xp,p,'r',linewidth = 4, label ='patient')
       ax.plot(Xi,U1,"b", linewidth = 3, label='Asymptotic TW')
       plt.xticks(fontsize = 16)
       plt.yticks(fontsize = 16)
       plt.xlabel('x', fontsize=16)
       plt.ylabel('Inflammation (Bleeding)', fontsize=16)
       leg = ax.legend(prop={"size":18})
       plt.legend()
       ax.set_title("sqrtD = " + str(round(sqrtD,3)) + \
                    "  distance = " + str(round(distD,3)), fontsize=20)
       fig.savefig(figname)
       plt.close()
    return distD
```

We test a set of diffusion parameters $S_D = [0.01, 0.02]$ with step of 0.0005 as follow:

Code chunk 37: fkpp_speed

```
D = np.arange(0.01,0.02,0.0005) ** 2
dists = [fkpp_interp(p,d) for d in D]
alpha, min_dist = 0.05, np.amin(dists)
D_comp = np.extract(dists<(1+alpha)*min_dist,D) #D compatibles
scompat = 2 * np.sqrt(np.extract(dists<(1+alpha)*min_dist,D))
print 'Vitesses compatibles', scompat
Xp = np.linspace(0,1,num=len(p))
print 'Temps d\'invasion',np.round((1-np.trapz(p,dx=Xp[1]-Xp[0]))/scompat,2)
```

Interpret with python2

```
Vitesses compatibles [0.026 0.027 0.028 0.029 0.03 ]
Temps d'invasion [21.53 20.73 19.99 19.3  18.66]
```

In addition, we plot the functional J to minimise and given by Equation 10.13 by the following code chunk:

Code chunk 38: fkpp_speed (part 2)

```
plt.figure()
fig,ax=plt.subplots()
plt.plot(D,dists,'o-b',linewidth = 2)
plt.plot(D[np.argmin(dists)],np.amin(dists),'Pr')
ax.annotate(r'$D^*$', xy=(D[np.argmin(dists)],np.amin(dists)),  xycoords='data',
            xytext=(0.5, 0.4), textcoords='axes fraction',
            arrowprops=dict(facecolor='black', shrink=0.0),
            horizontalalignment='right', verticalalignment='top',
            )

#plt.plot(sD_comp,dists_comp,'*r')
plt.xlabel(r'$D$',fontsize=16)
plt.ylabel(r'$J(D)$',fontsize=16)
plt.xticks(fontsize = 13)
plt.yticks(fontsize = 13)
plt.savefig('Test-Plot-sDcomp.png',bbox_inches = 'tight')
plt.close()
```

Interpret with python2

```
<Figure size 640x480 with 0 Axes>
[<matplotlib.lines.Line2D object at 0x7fa09e6b4d90>]
[<matplotlib.lines.Line2D object at 0x7fa09e682750>]
Text(0.5,0.4,'$D^*$')
Text(0.5,0,'$D$')
Text(0,0.5,'$J(D)$')
(array([5.0e-05, 1.0e-04, 1.5e-04, 2.0e-04, 2.5e-04, 3.0e-04, 3.5e-04,
       4.0e-04]), <a list of 8 Text xticklabel objects>)
(array([0.0012, 0.0014, 0.0016, 0.0018, 0.002 , 0.0022, 0.0024]), <a list of 7 Text yticklabel obje
```

We plot the optimal dilatation of u^A comparing to the patient's inflammation:

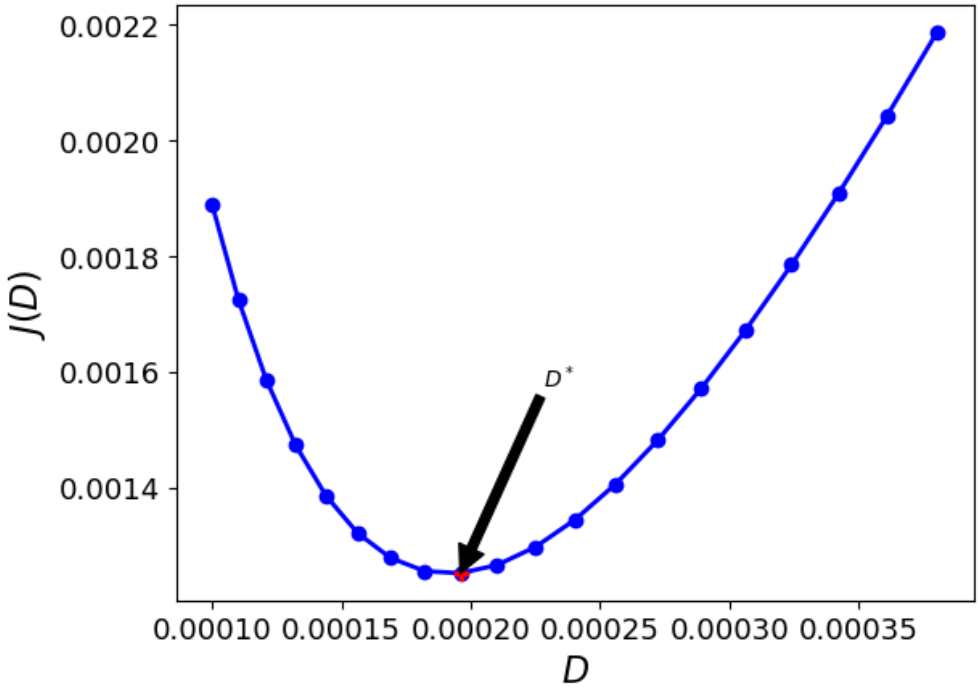


Figure 10.10: Representation of the function $J(D)$ given in Equation 10.13 for the diffusion parameters $D \in S_D$

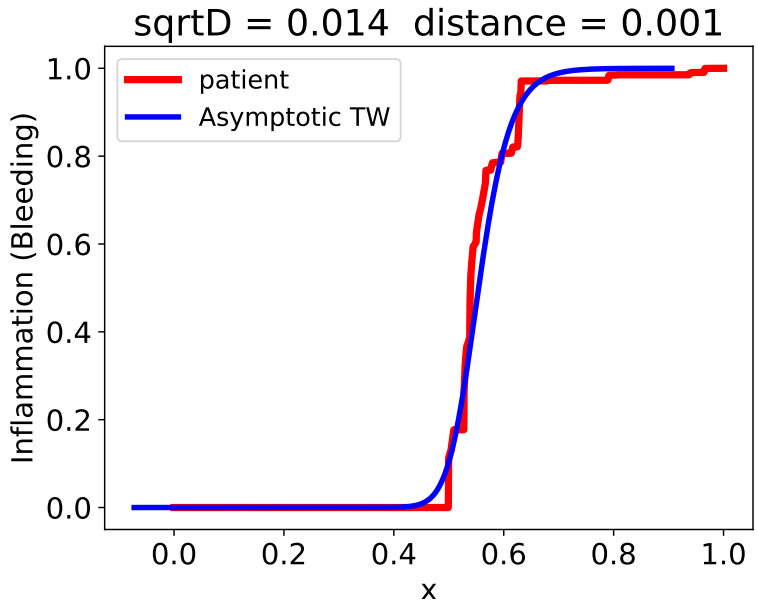


Figure 10.11: Representation of the inflammation of the patient 23 (in red) and u^A (blue) with the optimal dilation parameter $D^* = 0.014^2$

From Figure 10.11, we conclude that the optimal dilation $D^* = 0.014^2$ fits with the spatial distribution of the bleeding lesions found in colonoscopy video.

We study the spatial distribution of bleeding and lesions for patient of video 3, for which the colon is affected by bleeding and ulcers as shown in Figure 10.12:

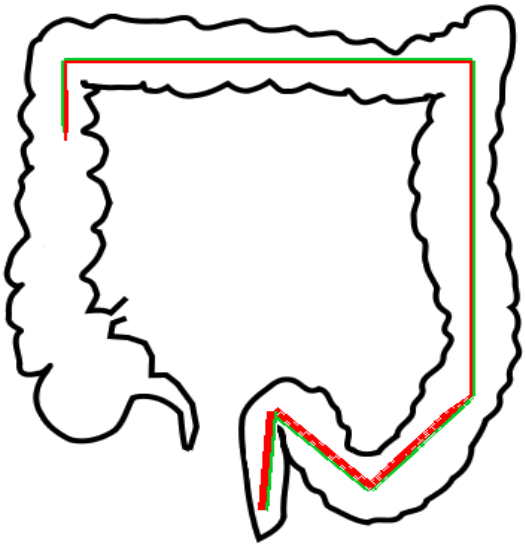


Figure 10.12: Representation of the doctor's annotations for patient of video 3 using the technique presented in chapter 7

We compute the speed of propagation for the two lesion's type. We denote by *fkpp_interp_ulcer* the function to compute the rescaling with modification of the color of ulcer's plot.
For bleeding:

Code chunk 39: fkpp_speed (part 4)

```
D = np.arange(0.01,0.02,0.0005) ** 2
dists_b = [fkpp_interp(pb_3,d) for d in D]
alpha, min_dist_b = 0.05, np.amin(dists_b)
D_comp_b = np.extract(dists<(1+alpha)*min_dist_b,D) #D compatibles
scompat_b = 2 * np.sqrt(np.extract(dists_b<(1+alpha)*min_dist_b,D))
print 'Compatibles velocities for bleeding', scompat_b
Xp = np.linspace(0,1,num=len(pb_3))
print 'Temps invasion',np.round((1-np.trapz(pb_3,dx=Xp[1]-Xp[0]))/scompat_b,2)
```

Interpret with python2

```
Compatibles velocities for bleeding [0.035 0.036 0.037 0.038 0.039]
Temps invasion [16.47 16.01 15.58 15.17 14.78]
```

Code chunk 40: fkpp_speed (part 5)

```
D = np.arange(0.01,0.02,0.0005) ** 2
dists_u = [fkpp_interp_ulcer(pu_3,d) for d in D]
alpha, min_dist_u = 0.05, np.amin(dists_u)
D_comp_u = np.extract(dists<(1+alpha)*min_dist_u,D) #D compatibles
scompat_u = 2 * np.sqrt(np.extract(dists_u<(1+alpha)*min_dist_u,D))
print 'Compatibles velocities for ulcer', scompat_u
Xp = np.linspace(0,1,num=len(pu_3))
print 'Temps invasion',np.round((1-np.trapz(pu_3,dx=Xp[1]-Xp[0]))/scompat_u,2)
```

Interpret with python2

```
Compatibles velocities for ulcer [0.02]
Temps invasion [9.96]
```

We plot the optimal scalings for bleeding and ulcers:

Code chunk 41: fkpp_speed (part 6)

```
ind_b = np.argmin(dists_b)
fkpp_interp(pb_3,D[ind_b],'recalageOptimal_Bleeding_Patientvideo_3.pdf')

ind_u = np.argmin(dists_u)
fkpp_interp_ulcer(pu_3,D[ind_u],'recalageOptimal_Ulcer_Patientvideo_3.pdf')
```

Interpret with python2

```
0.08681243557484017
0.006370388191950768
```

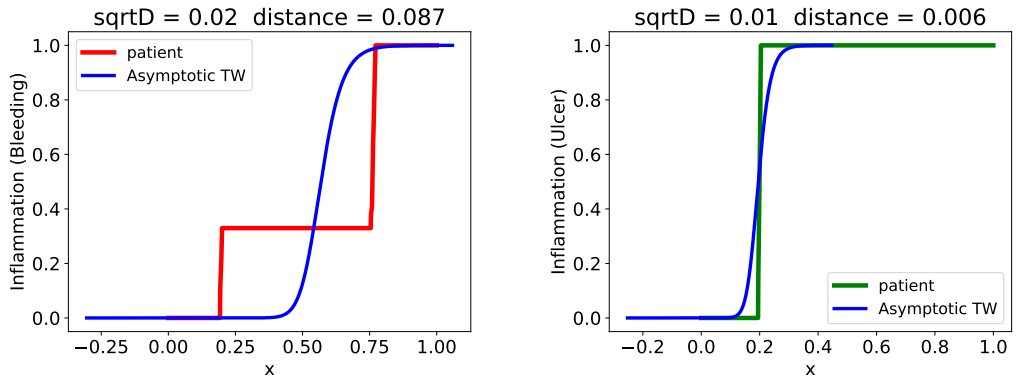


Figure 10.13: Rescaling of the travelling wavefront u^A for patient of video 3 with optimal diffusion for bleeding lesions $D_b^* = 0.02^2$ and ulcers $D_u^* = 0.01^2$

According to our results, we find that the bleeding lesions will diffuse in the colon ($v(bleeding) = 0.037$) more faster than the ulcers lesions ($v(ulcers) = 0.02$).
We compile the velocities for patient having more affected parts of the colon as represented in Figure 10.14:

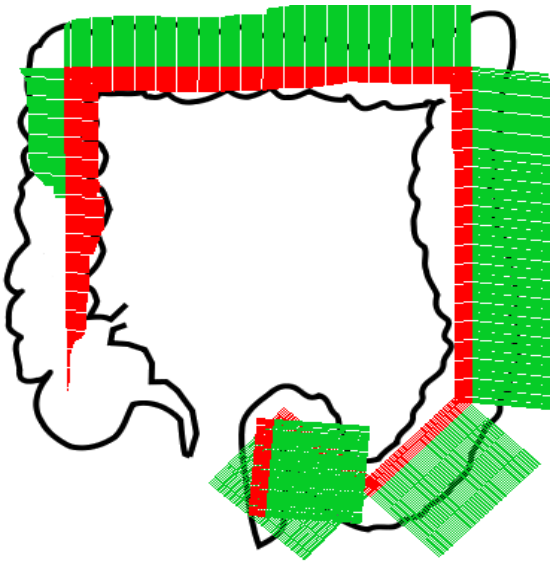


Figure 10.14: Representation of the doctor's annotations for patient of video 42 using the technique presented in chapter 7

And we find the set of compatibles 0.05-velocities:

– – – – – – – – – – – – – *todo* – – – – – – – – – – – – –

We compile the velocities for patient having more affected parts of the colon as represented in Figure 10.15:

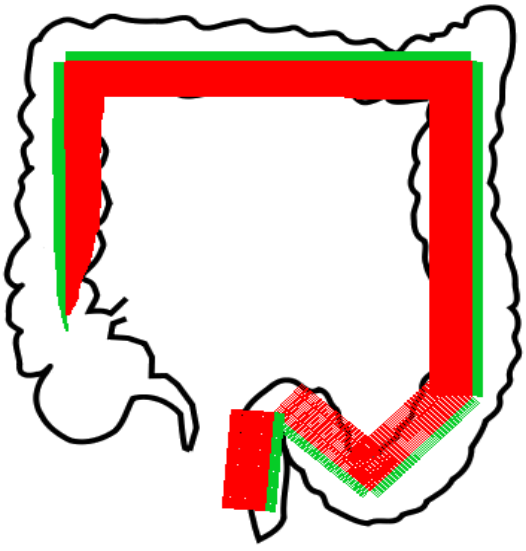


Figure 10.15: Representation of the doctor's annotations for patient of video 7 using the technique presented in chapter 7

Code chunk 42: fkpp_speed (part 7)

```
# Velocities of bleeding's movement
D = np.arange(0.01,0.02,0.0005) ** 2
dists_b = [fkpp_interp(pb,d) for d in D]
alpha, min_dist_b = 0.05, np.amin(dists_b)
D_comp_b = np.extract(dists<(1+alpha)*min_dist_b,D) #D compatibles
scompat_b = 2 * np.sqrt(np.extract(dists_b<(1+alpha)*min_dist_b,D))
print 'Compatibles velocities for bleeding', scompat_b
Xp = np.linspace(0,1,num=len(pb))
print 'Temps invasion',np.round((1-np.trapz(pb,dx=Xp[1]-Xp[0]))/scompat_b,2)
```

Interpret with python2

```
Compatibles velocities for bleeding [0.036 0.037 0.038 0.039]
Temps invasion [19.93 19.39 18.88 18.4 ]
```

Code chunk 43: fkpp_speed (part 8)

```
# Velocities of ulcer's movement
D = np.arange(0.008,0.018,0.0005) ** 2
dists_u = [fkpp_interp_ulcer(pu,d) for d in D]
alpha, min_dist_u = 0.05, np.amin(dists_u)
D_comp_u = np.extract(dists<(1+alpha)*min_dist_u,D) #D compatibles
scompat_u = 2 * np.sqrt(np.extract(dists_u<(1+alpha)*min_dist_u,D))
print 'Compatibles velocities for ulcer', scompat_u
Xp = np.linspace(0,1,num=len(pu))
print 'Temps invasion',np.round((1-np.trapz(pu,dx=Xp[1]-Xp[0]))/scompat_u,2)
```

Interpret with python2

```
Compatibles velocities for ulcer [0.031 0.032 0.033 0.034 0.035]
Temps invasion [20.84 20.19 19.58 19.01 18.46]
```

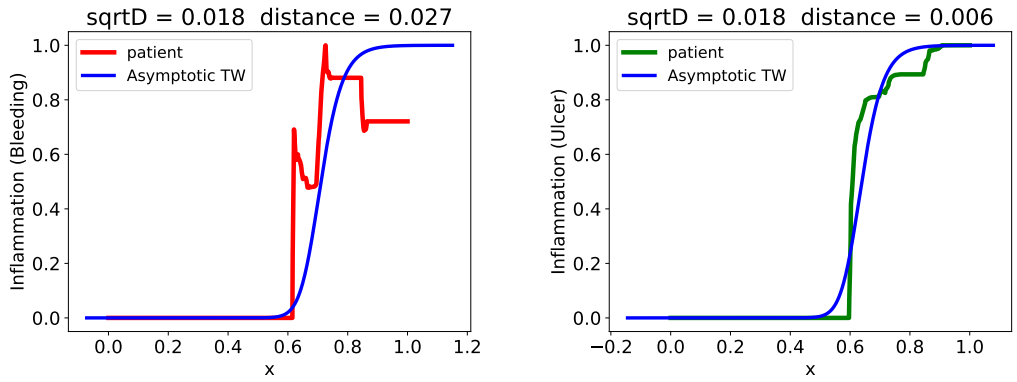


Figure 10.16: Rescaling of the travelling wavefront u^A for patient of video 7 with optimal diffusion for bleeding lesions $D_b^* = 0.02^2$ and ulcers $D_u^* = 0.01^2$

10.6 Conclusion

This chapter is devoted to the spatial modelling of the spatial distribution of UC lesions found in a colonoscopy video. Although, the works presented to study the cell interactions to understand the physiopathology of the disease, any of them have presented a model to study the spatial distribution of the (surface) of the lesions according to their position in the colon.

We have investigated the class of the one-dimensional reaction-diffusion equations, and more precisely the classical Fisher-KPP with mono-stable reaction term. The variable of the model was either the surface of bleeding or the surface of ulcers as a function of the colon curvilinear abscissa. The colonoscopy video is considered taken after a long time of the beginning of the disease, the patient is thus considered in the asymptotic regime. We then considered the patient as a travelling wave solution of the Fisher-KPP equation.

The travelling wave solutions for Fisher-KPP with the compactly supported initial condition are one-dimensional space indexed by the diffusion scalar D. In addition, they are unique up to translation and dilation of order $\sqrt{D}$ compared to the case $D = 1$. We proceed with an inverse problem to estimate the compatible dilation of the asymptotic travelling wave that fits with the inflammation of the patient. Moreover, we computed the velocities at which the lesions will continue propagating in the colon and the invasion time as well.
The numerical results obtained using the FCTS numerical scheme for a set of patients from the database Vatic bring useful information about the prognosis of the disease, namely the velocity of displacement in the colon.

10.7 Extension

The assumptions adopted in this work

Chapter 11

Conclusion and future directions

The main goal of this thesis is to provide an objective evaluation about the severity state of the ulcerative colitis disease. The UC severity assessment is generally made by the gastroenterologist who rates, via/through the endoscopic scores, the state of the principle lesions of the disease such as the bleeding, the ulcers and the vascular pattern which are found in the endoscopic videos produced by the colonoscopy examination. This evaluation is very subjective and changing/variable between doctors. We have used a set of image analysis techniques and mathematical modelling approaches to handle the issue.

Firstly, we proposed to detect all the bleeding and the ulcer lesions in a colonoscopy video through an automatic computer aided algorithm (cf chapter 6). From the state of the art of works, the learning data base.

We used a data base of 37 colonoscopy videos for which the gastroenterologists have delineated the lesions by rectangles. However, the complex geometry of the lesions invoked many errors of correspondence between the exact lesion and its delineation, and the detector detection as well. To remedy this problem, we have modified the detector's sensitivity criteria which demonstrated a good adaptation to encounter the labeling errors that may occur during the detector's training phase. In addition, we speeded up the computation of the detector parameters by only focusing on the non-trivial models/detectors.

Next in chapter 7, we generated the mapping of the spatial distribution of the detected lesions all along the colon. By this way, we succeeded in visualizing/displaying the disease severity state and hence differentiating between the patients. The obtained results have shown high variability of lesions distribution and has lead to study further the evolution of the severity and the inter-experts variability (cf chapter 8).

We then proposed in chapter 8 to integrate the spatial information about lesions on the different parts of the colon to the computation of the severity scores and

we showed good agreement with the evaluations made in medical practice.

Since the UC disease is almost incurable, the doctors need to control its progressing speed. Therefore, we proposed in chapter 10 to compute the speed of the disease propagation with the help/by applying mathematical model the spatial distribution of the lesions along the colon using partial differential equations of type reaction diffusion. The obtained results can be used as diagnosis assistance to the gastroenterologist's diagnosis and treatment decisions to the long time disease behavior/evolution.

Appendix A

Codes

A.1 General settings

We import all necessary python librairies to generate the codes:

Code chunk 44: python

```
import numpy as np
import cv2, io, string
from math import sqrt,exp
```

Code chunk 45: Comparaison_RL_EDO1st.py

```
import matplotlib as mpl
mpl.use('PDF')
import matplotlib.pyplot as plt
import numpy as np
import cv2, io, string
from math import sqrt,exp
import pickle
```

A.2 Vatic database image storage

Code chunk 46: Comparaison_RL_EDO1st.py (part 2)

```
# Data base informations
num_video=[1,2,3,5,6,7,9,10,11,12,13,14,15,16,17,19,21,23,24,27,29,31,32,33,34,37,38,
39,40,41,42,43,44,46,47,48,49]
fichier_image_video=['MVideo_p1_0',
'MVideo_p1_1','MVideo_p1_2','MVideo_p1_3','MVideo_p1_4','MVideo_p1_5','MVideo_p1_6',
'MVideo_p2_0','MVideo_p2_1','MVideo_p2_2','MVideo_p2_3','MVideo_p2_4','MVideo_p2_5',
'MVideo_p2_6','MVideo_p2_7','MVideo_p2_8','MVideo_p3_0','MVideo_p3_1','MVideo_p3_2',
'MVideo_p3_3','MVideo_p3_4','MVideo_p4_0','MVideo_p4_1','MVideo_p4_2','MVideo_p4_3',
'MVideo_p4_4','MVideo_p4_5','MVideo_p4_6','MVideo_p5_0','MVideo_p5_1','MVideo_p6_v42',
'MVideo_p5_2','MVideo_p5_3','MVideo_p5_4','MVideo_p5_5','MVideo_p5_6','MVideo_p5_7']
```

A.3 Data computation: Count of lesions

Code chunk 47: Comparaison_RL_EDO1st.py (part 3)

```
def abscvideo(nframe):
    abscvid = np.linspace(0, 1, num=nframe )
    return abscvid
```

Code chunk 48: Comparaison_RL_EDO1st.py (part 4)

```
def Count_Lesions(num_video):
    fichier_txt="/users/alali/ownCloud/PremierProjet/VideoTXT/video%d.txt"%(num_video)
    fichier = io.open(fichier_txt, 'r')
    Lignes = fichier.readlines()
    Maxannot = Lignes[len(Lignes) - 1]
    lMaxannot = Maxannot.split()
    Nfmax = int(lMaxannot[5])
    NannotS_frame,NannotU_frame=np.zeros(Nfmax), np.zeros(Nfmax)
    for frame in range(0, Nfmax):
        for ligne in Lignes:
            l = ligne.split()
            if (int(l[7]) + int(l[6])) == 0 and int(l[5]) == frame:
                if '"Saignement"' in l:
                    NannotS_frame[frame]+=1
                elif '"Ulceration"' in l:
                    NannotU_frame[frame]+=1

    return NannotS_frame,NannotU_frame
```

For time saving, we make a copy of data information to computation complexity reduction. The dataset is divided into 2 subsets, the first one for treating videos number 1 to 39, the second for videos number 40 to 49 as in the code chunk:

Code chunk 49: python (part 2)

```
#num_video=[1,2,3,5,6,7,9,10,11,12,13,14,15,16,17,19,21,23,24,27,29,31,32,33,34,37,38,39]
num_video=[40,41,42,43,44,46,47,48,49]

y_b_count, y_u_count=[[]]*len(num_video),[[]]*len(num_video)
for index in range(len(num_video)):
    y_b,y_u= Count_Lesions(num_video[index])
    y_b_count[index].append(y_b)
    y_u_count[index].append(y_u)

#with open('/users/alali/Cartographie-Lesions-DrAnnotation-and-Auto/Estim-EDO1st-Count-\
# Lesions-AxeColon-DrAnnotation/'
#          'New-Data-BleedUlcer-Count-DrAnnotation-vid1a39.pickle', 'wb') as f:
#           pickle.dump([y_b_count,y_u_count],f)
with open('/users/alali/Cartographie-Lesions-DrAnnotation-and-Auto/Estim-EDO1st-Count-\
#Lesions-AxeColon-DrAnnotation/'
'New-Data-BleedUlcer-Count-DrAnnotation-vid40a49.pickle', 'wb') as f:
          pickle.dump([y_b_count,y_u_count],f)
```

Code chunk 50: python (part 3)

```
## DATA USED Example
with open('/users/alali/Cartographie-Lesions-DrAnnotation-and-Auto/Estim-EDO1st-Count-\
Lesions-AxeColon-DrAnnotation/New-Data-BleedUlcer-Count-DrAnnotation-vid1a39.pickle','rb') as
          y_b_count,y_u_count=pickle.load(f)

# bleeding
Data_Count_b = [ [individualArray] for individualArray in y_b_count ]
Data_Count_b=Data_Count_b[0][0]
print "Array Count vid 1= "+str(len(Data_Count_b[0]))
print "Array Count vid 2= "+str(len(Data_Count_b[1]))
print "Array Count vid 3= "+str(len(Data_Count_b[2]))
# ulcer
Data_Count_u = [ [individualArray] for individualArray in y_u_count]
Data_Count_u=Data_Count_u[0][0]
print Data_Count_b[0][100],Data_Count_b[0][200]
```

Interpret with python2

```
Array Count vid 1= 812
Array Count vid 2= 378
Array Count vid 3= 2058
2.0 2.0
```

A.4 Data computation: Percentage of lesions

For a given frames, after an application of extraction of unuseful information (pixels in black), we extract the amount of abnormal labeled pixels according to the total masque pixels using the function *Data_perc_annotations*.

Code chunk 51: Comparaison_RL_EDO1st.py (part 5)

```
# FIV: fichier_image_video
def Data_perc_annotations(num_video,FIV):
    fichier_video="/users/alali/Rapports-Stage-Lepton/Modelisation_Video_Frame/%s"%(FIV)
    fichier_txt="/users/alali/ownCloud/PremierProjet/VideoTXT/video%d.txt"%(num_video)
    fichier = io.open(fichier_txt, 'r')
    Lignes = fichier.readlines()
    Maxannot = Lignes[len(Lignes) - 1]
    lMaxannot = Maxannot.split()
    Nfmax = int(lMaxannot[5])
    Masque = "%s/Test-Fig-Masque.png" % (fichier_video)
    M = cv2.imread(Masque)
    M = cv2.erode(M[:, :, 1], 255 * np.ones((5, 5), np.uint8))
    M1 = cv2.imread(Masque)
    M1 = cv2.erode(M1[:, :, 1], 255 * np.ones((5, 5), np.uint8))
    n_masque = np.sum(M == 255)
    per_bleed, per_ulcer=np.zeros(Nfmax),np.zeros(Nfmax)
    for frame in range(0, Nfmax):
        for ligne in Lignes:
            l = ligne.split()
            if (int(l[7]) + int(l[6])) == 0 and int(l[5]) == frame:
                if '"Saignement"' in l:
                    M[int(l[2]):int(l[4]) + 1, int(l[1]):int(l[3]) + 1] = 100
                elif '"Ulceration"' in l:
                    M[int(l[2]):int(l[4]) + 1, int(l[1]):int(l[3]) + 1] = 200

        M[M1 == 0] = 0
        per_bleed[frame]=float(np.sum(M==100))/n_masque
        per_ulcer[frame]=float(np.sum(M==200))/n_masque

    return per_bleed,per_ulcer, Nfmax
```

Same work done for data of percentage of inflammation in a given frame. For time saving, by the command *pickle.dump*, we make a copy of data information to computation complexity reduction after running the function *Data_perc_annotations*, as result of iteration we get two vectors *y_b_base* and *y_u_base* for every patient that contain for every frame the percentage of abnormal pixels labeled as bleeding and ulcer respectively. The dataset is splited into two parts, part 1 (video number 1 to 39) and part 2 (video 40 to 49).

Code chunk 52: python (part 4)

```
#with open('/users/alali/Cartographie-Lesions-DrAnnotation-and-Auto/Estim-EDO1st-\
# Percentage-Inflammation-AxeColon-DrAnnotation/'
#          'Data-BleedUlcer-Perc-DrAnnotation-vid40a49.pickle', 'wb') as f:
#     pickle.dump([y_b_base,y_u_base], f)

with open('/users/alali/Cartographie-Lesions-DrAnnotation-and-Auto/Estim-EDO1st-\
Percentage-Inflammation-AxeColon-DrAnnotation/'
         'Data-BleedUlcer-Perc-DrAnnotation-vid1a39.pickle', 'rb') as f:
          y_b_base,y_u_base=pickle.load(f)

#to extract the data information for every patient we split the list of arrays y_b_base
#verify according to number of video frames
# Case of bleeding
Data_Perc_b = [ [individualArray] for individualArray in y_b_base]
Data_Perc_b=Data_Perc_b[0][0]
print "Array vid 1= "+str(len(Data_Perc_b[0]))
print "Array vid 2= "+str(len(Data_Perc_b[1]))
print "Array vid 3= "+str(len(Data_Perc_b[2]))
# ulcer
Data_Perc_u = [ [individualArray] for individualArray in y_u_base ]
Data_Perc_u=Data_Perc_u[0][0]
print Data_Perc_b[0][100],Data_Perc_b[0][200]
```

Interpret with python2

```
Array vid 1= 812
Array vid 2= 378
Array vid 3= 2058
0.19890480232362412 0.4604880040398397
```

As results from the above two code chunk, we can see that the frames number 100 and 200 of video number 1, both present 2 box/annotations of bleeding but with diffrent boxes size hence with different inflammed area.

A.5 Test for chap

Code chunk 53: python (part 5)

```
print Data_Count_b[8][850],Data_Count_u[8][850]
print Data_Perc_b[8][850],Data_Perc_u[8][850]

print Data_Count_b[22][364],Data_Count_u[22][364]
print Data_Perc_b[22][364],Data_Perc_u[22][364]
```

Interpret with python2

```
5.0 1.0
0.7298445877498579 0.10278104274101606
1.0 1.0
0.4124588616710282 0.21212362642413693
```

Appendix B

Maps of lesions for Vatic patients

In this section, we aim to present the lesions' maps generated by the proposed technique in chapter 7 for all the patients of the Vatic database, for which we have the set of annotations for bleeding and ulcers.

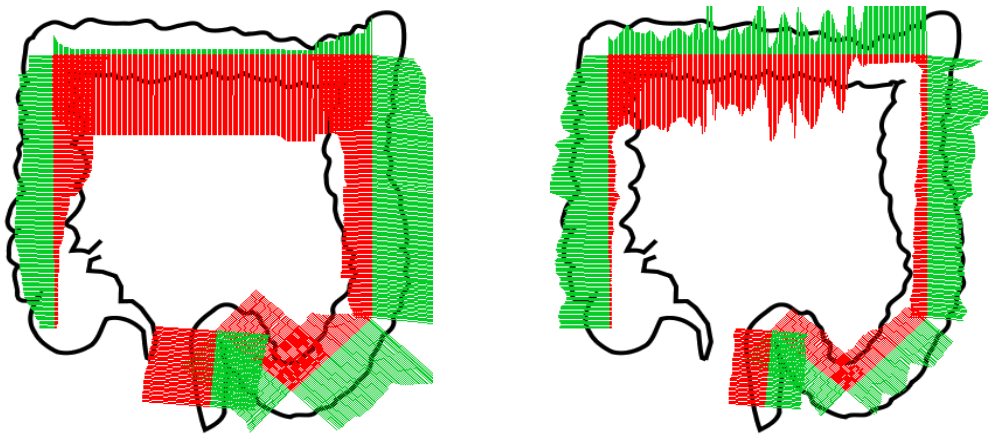


Figure B.1: Distribution of the UC lesions for the patient 1. Medical annotations on the left, and automatic detection on the right.

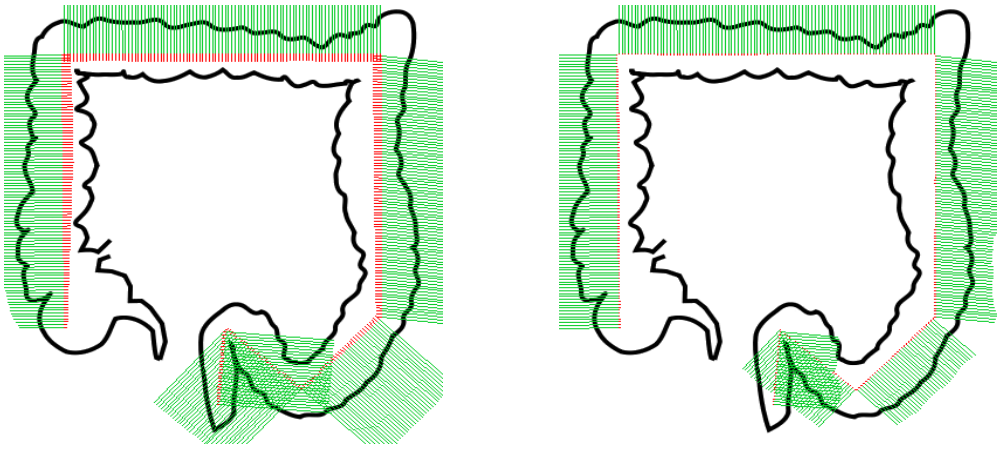


Figure B.2: Distribution of the UC lesions for the patient 2. Medical annotations on the left, and automatic detection on the right.

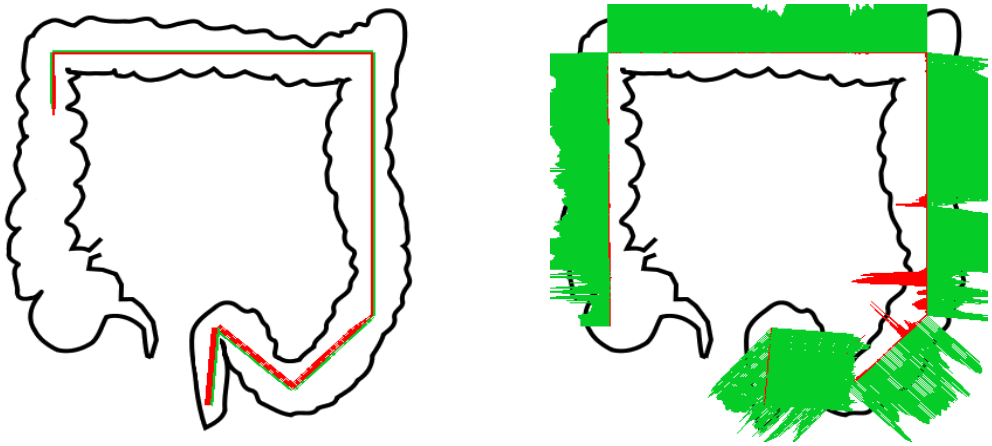


Figure B.3: Distribution of the UC lesions for the patient 3. Medical annotations on the left, and automatic detection on the right.

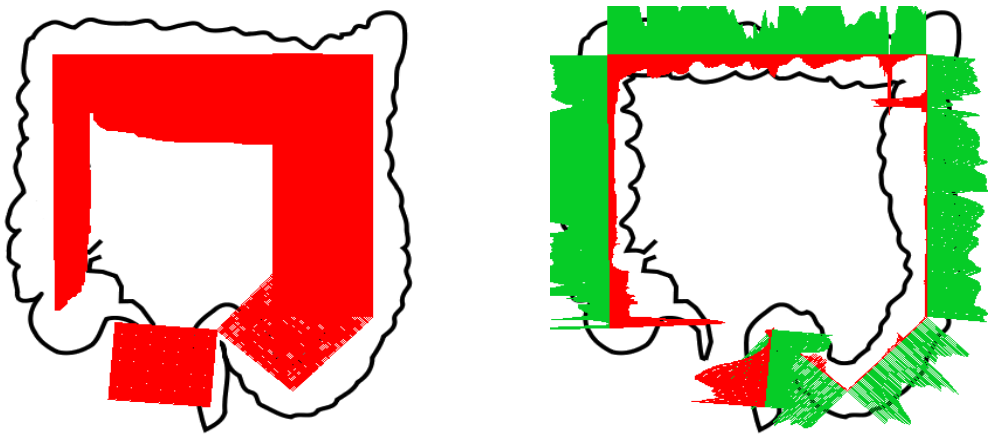


Figure B.4: Distribution of the UC lesions for the patient 5. Medical annotations on the left, and automatic detection on the right.

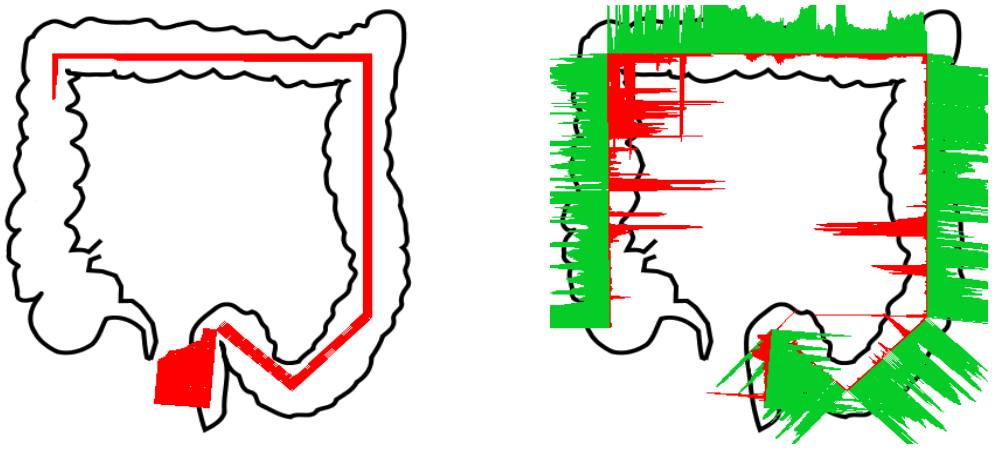


Figure B.5: Distribution of the UC lesions for the patient 6. Medical annotations on the left, and automatic detection on the right.

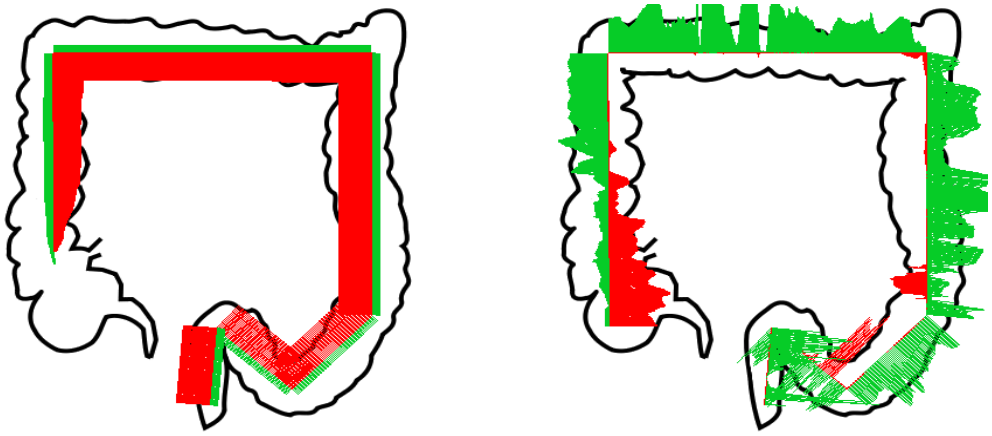


Figure B.6: Distribution of the UC lesions for the patient 7. Medical annotations on the left, and automatic detection on the right.

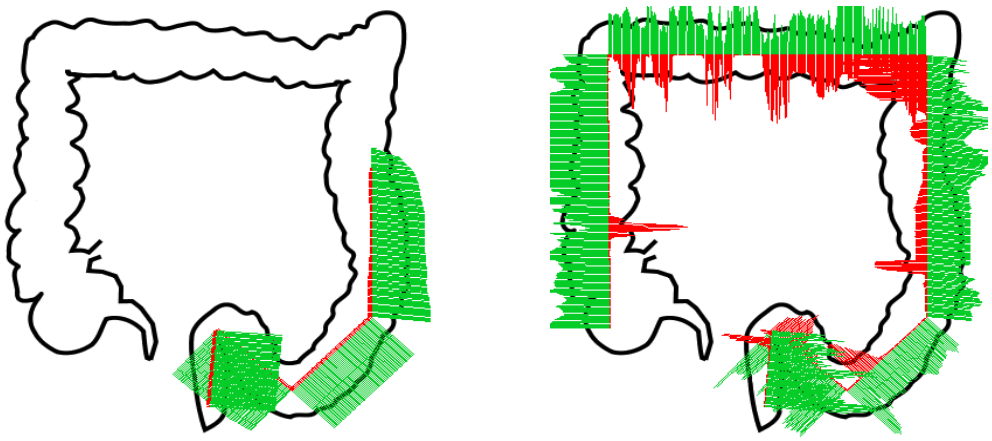


Figure B.7: Distribution of the UC lesions for the patient 9. Medical annotations on the left, and automatic detection on the right.

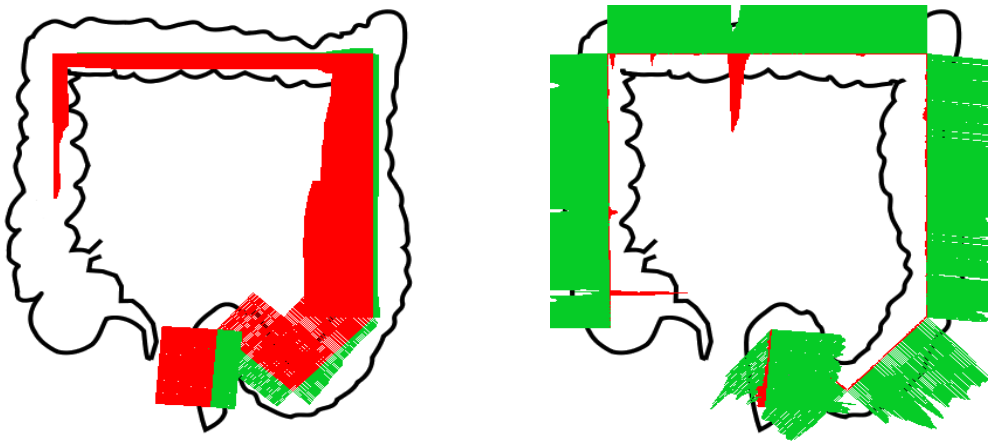


Figure B.8: Distribution of the UC lesions for the patient 10. Medical annotations on the left, and automatic detection on the right.

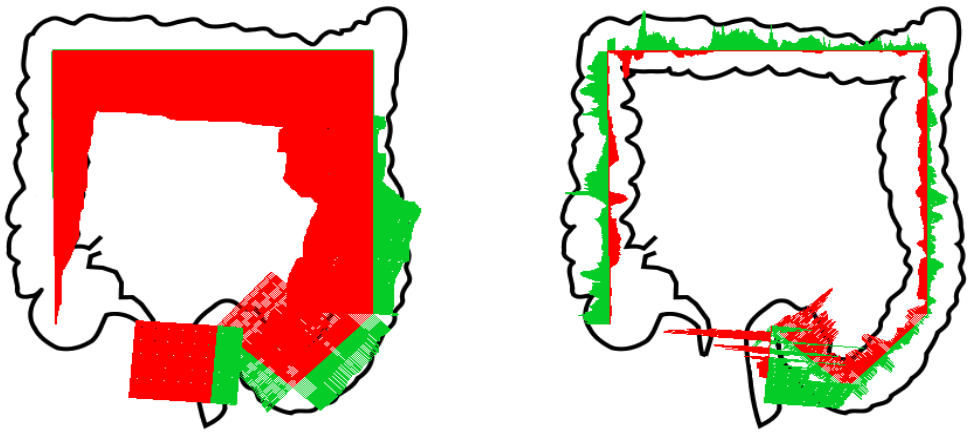


Figure B.9: Distribution of the UC lesions for the patient 11. Medical annotations on the left, and automatic detection on the right.

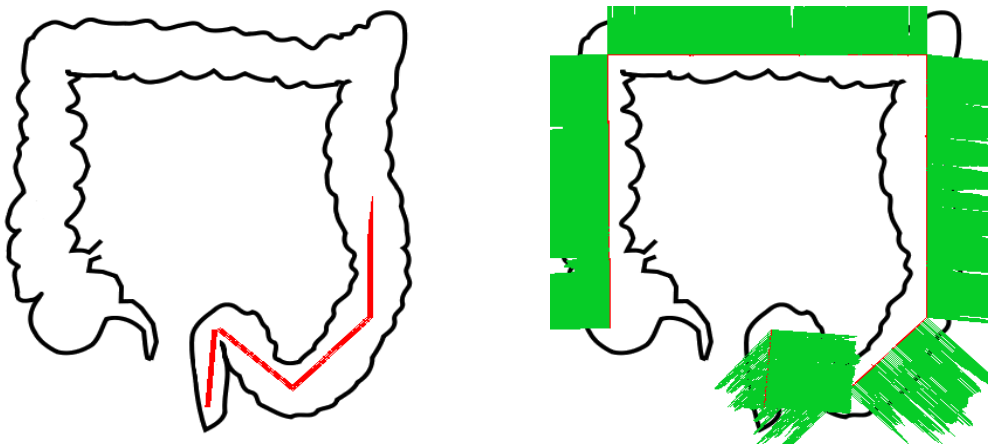


Figure B.10: Distribution of the UC lesions for the patient 12. Medical annotations on the left, and automatic detection on the right.

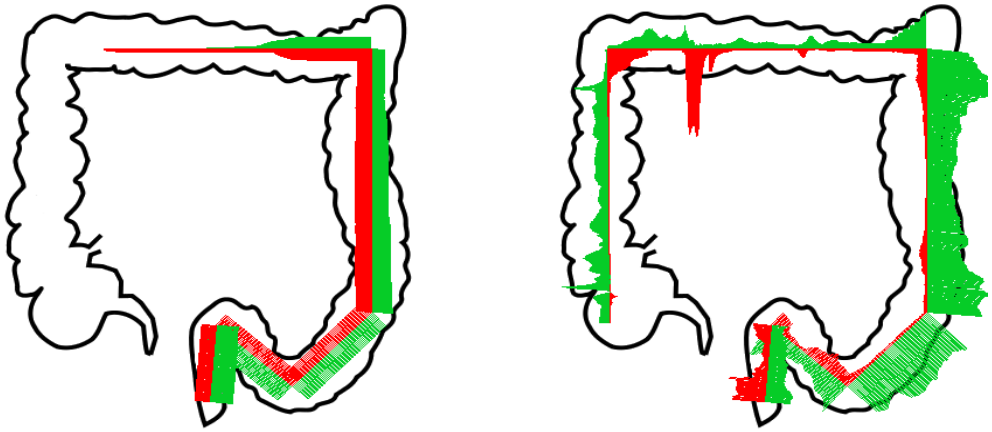


Figure B.11: Distribution of the UC lesions for the patient 13. Medical annotations on the left, and automatic detection on the right.

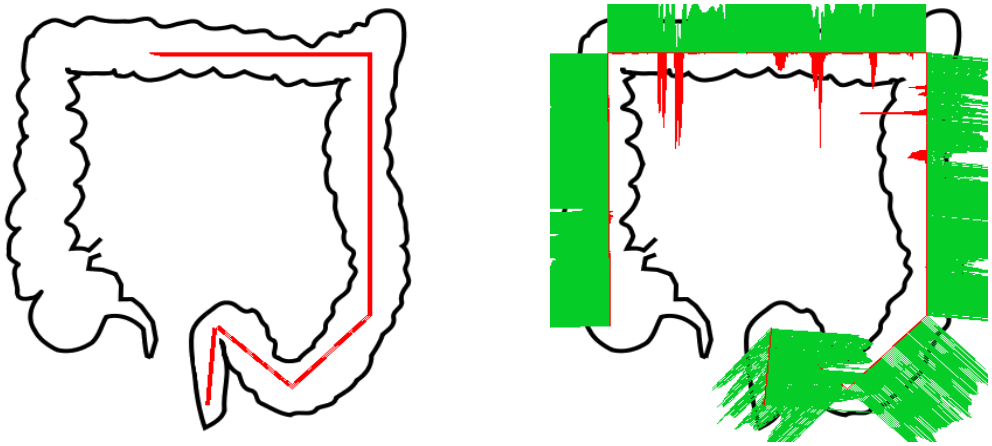


Figure B.12: Distribution of the UC lesions for the patient 14. Medical annotations on the left, and automatic detection on the right.

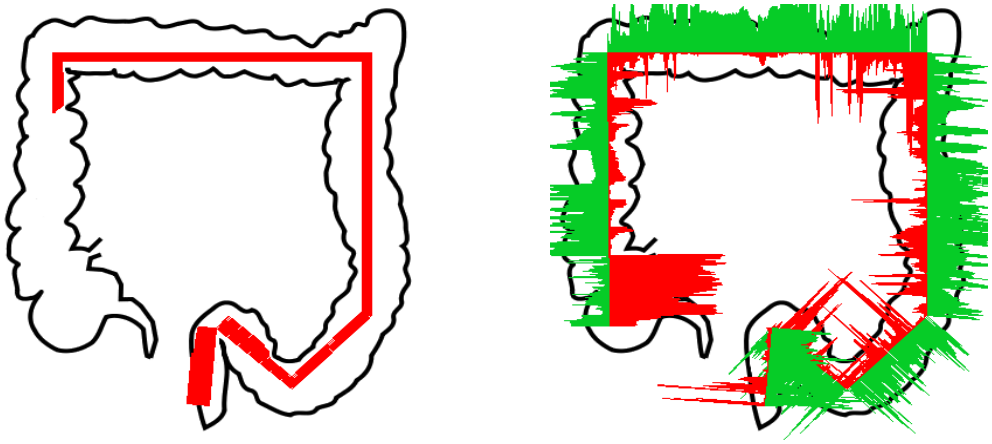


Figure B.13: Distribution of the UC lesions for the patient 15. Medical annotations on the left, and automatic detection on the right.

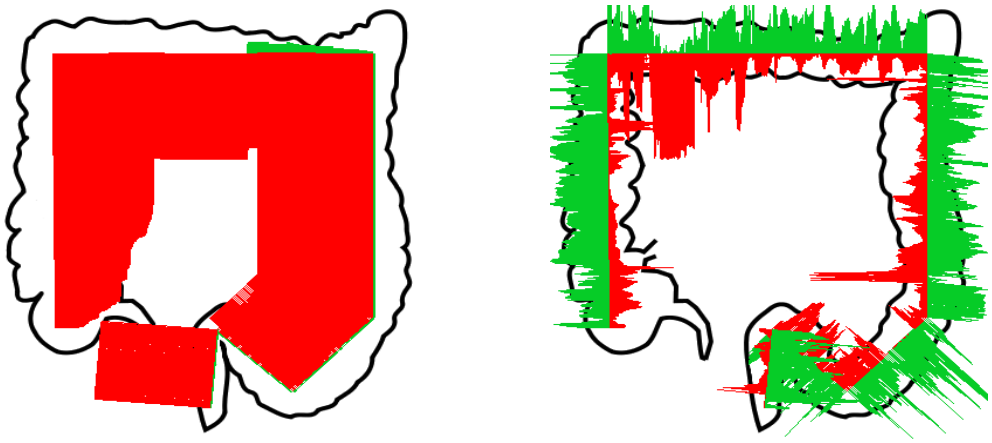


Figure B.14: Distribution of the UC lesions for the patient 16. Medical annotations on the left, and automatic detection on the right.

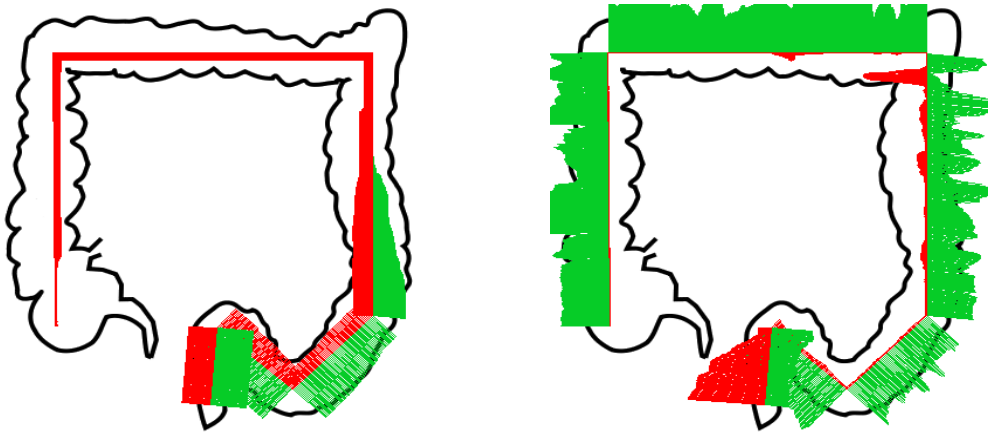


Figure B.15: Distribution of the UC lesions for the patient 17. Medical annotations on the left, and automatic detection on the right.

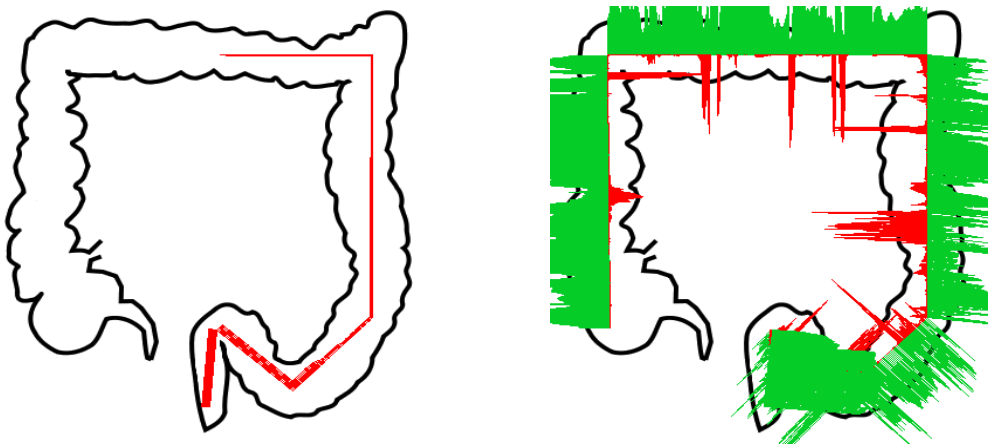


Figure B.16: Distribution of the UC lesions for the patient 19. Medical annotations on the left, and automatic detection on the right.

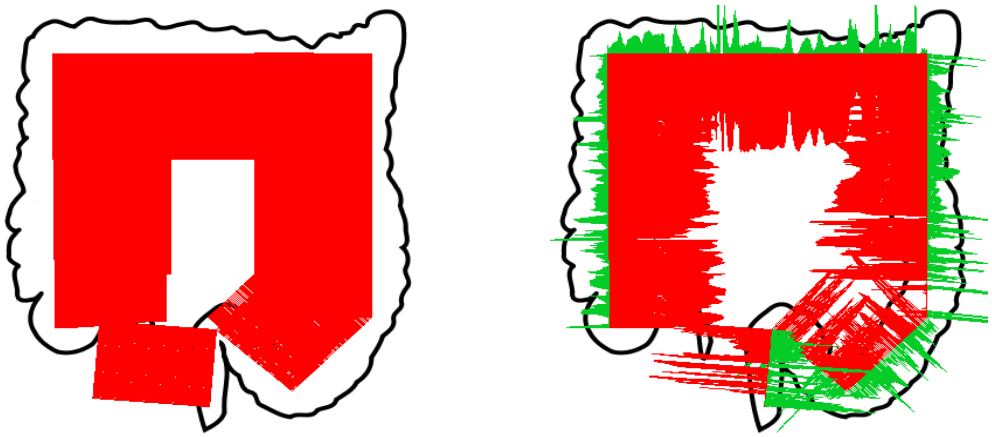


Figure B.17: Distribution of the UC lesions for the patient 21. Medical annotations on the left, and automatic detection on the right.

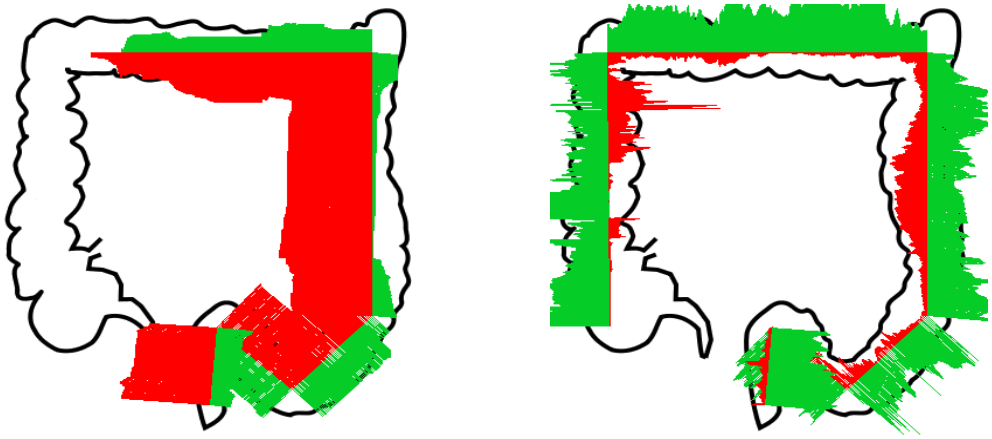


Figure B.18: Distribution of the UC lesions for the patient 23. Medical annotations on the left, and automatic detection on the right.

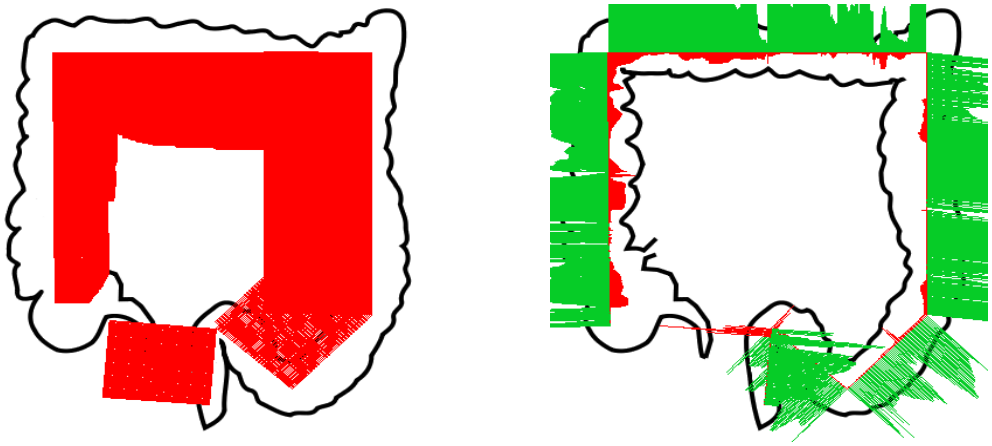


Figure B.19: Distribution of the UC lesions for the patient 24. Medical annotations on the left, and automatic detection on the right.

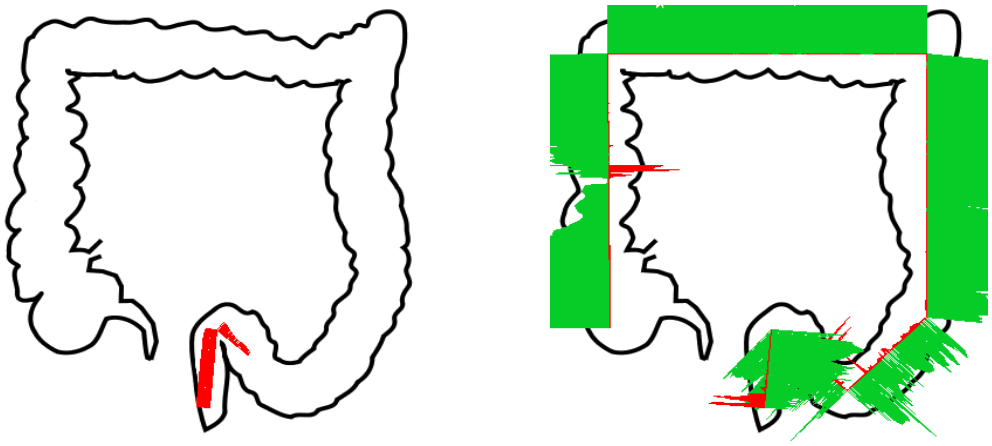


Figure B.20: Distribution of the UC lesions for the patient 27. Medical annotations on the left, and automatic detection on the right.

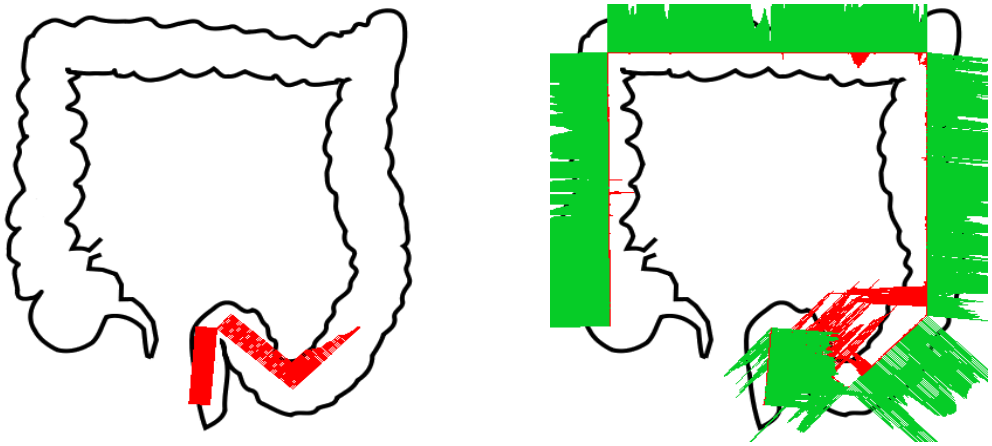


Figure B.21: Distribution of the UC lesions for the patient 29. Medical annotations on the left, and automatic detection on the right.

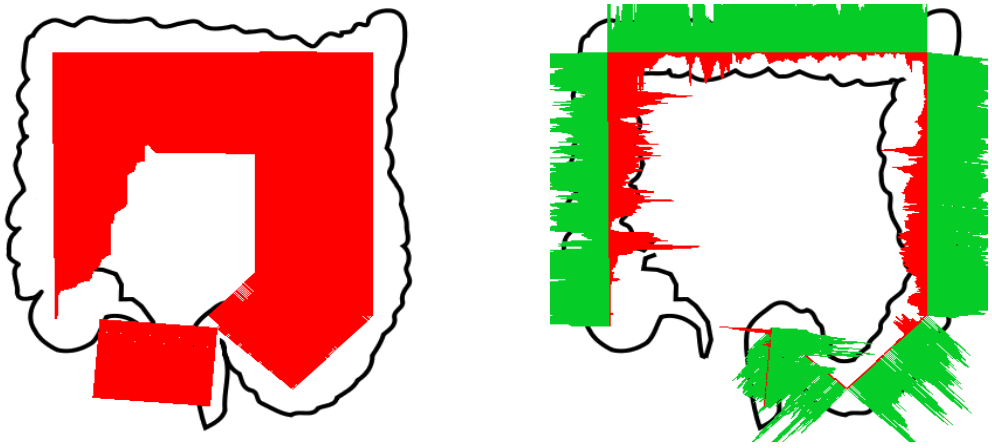


Figure B.22: Distribution of the UC lesions for the patient 31. Medical annotations on the left, and automatic detection on the right.

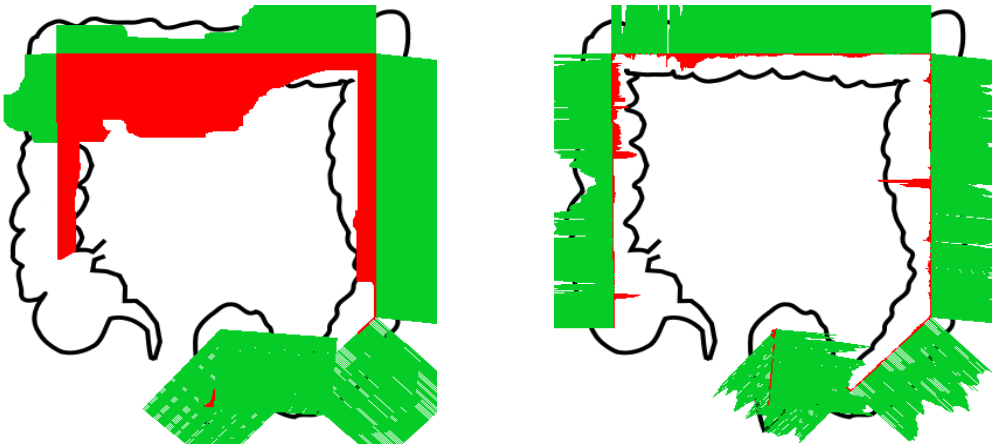


Figure B.23: Distribution of the UC lesions for the patient 32. Medical annotations on the left, and automatic detection on the right.

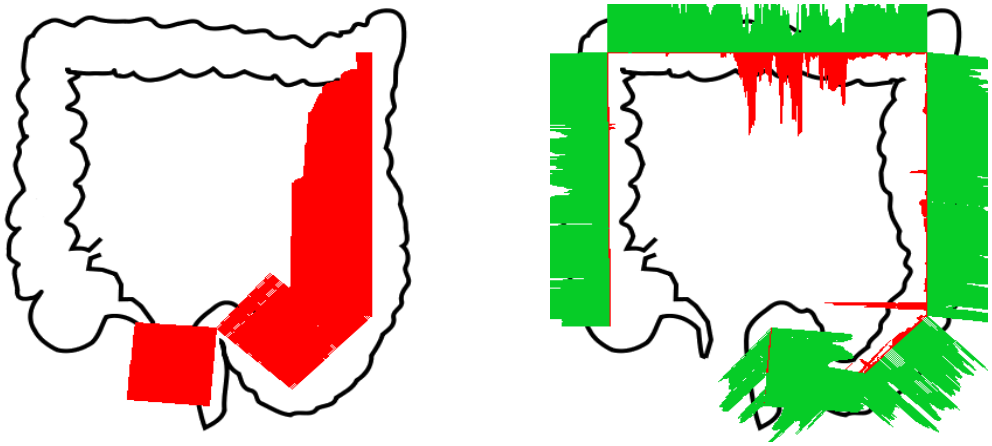


Figure B.24: Distribution of the UC lesions for the patient 33. Medical annotations on the left, and automatic detection on the right.

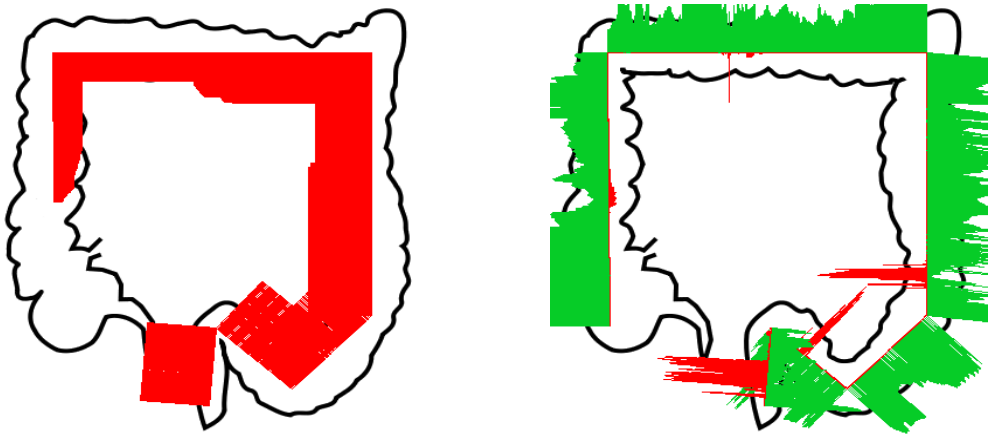


Figure B.25: Distribution of the UC lesions for the patient 34. Medical annotations on the left, and automatic detection on the right.

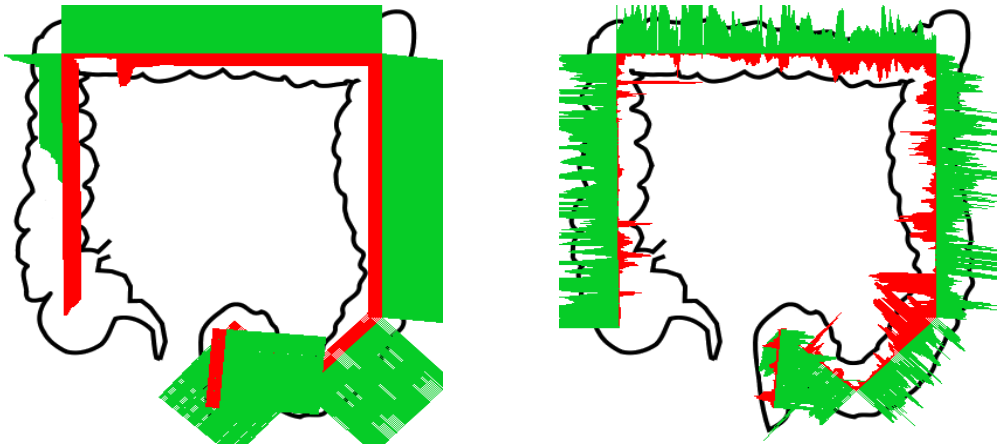


Figure B.26: Distribution of the UC lesions for the patient 37. Medical annotations on the left, and automatic detection on the right.

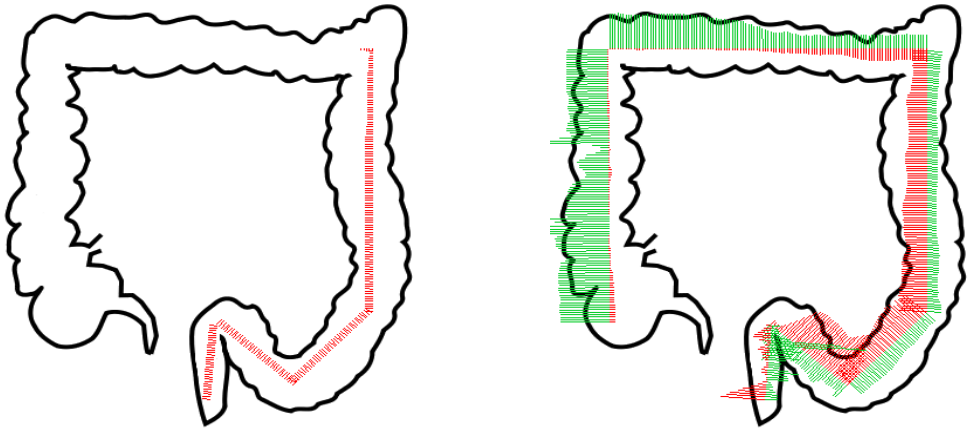


Figure B.27: Distribution of the UC lesions for the patient 38. Medical annotations on the left, and automatic detection on the right.

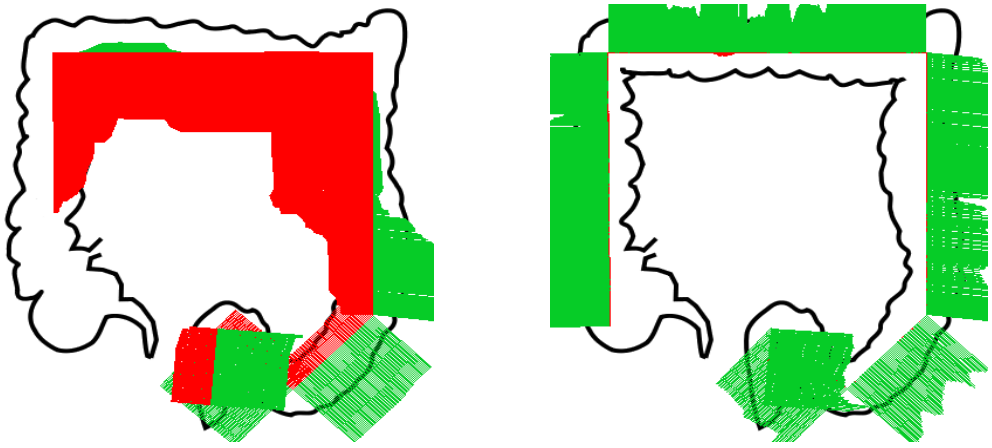


Figure B.28: Distribution of the UC lesions for the patient 39. Medical annotations on the left, and automatic detection on the right.

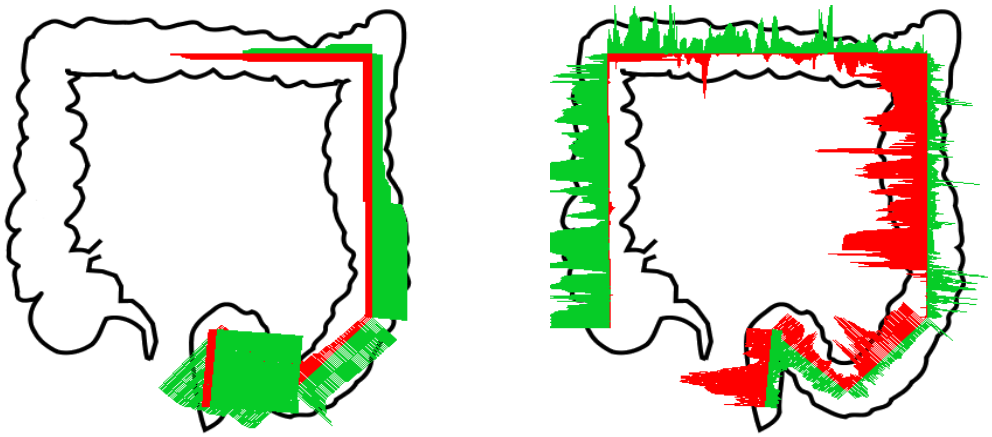


Figure B.29: Distribution of the UC lesions for the patient 40. Medical annotations on the left, and automatic detection on the right.

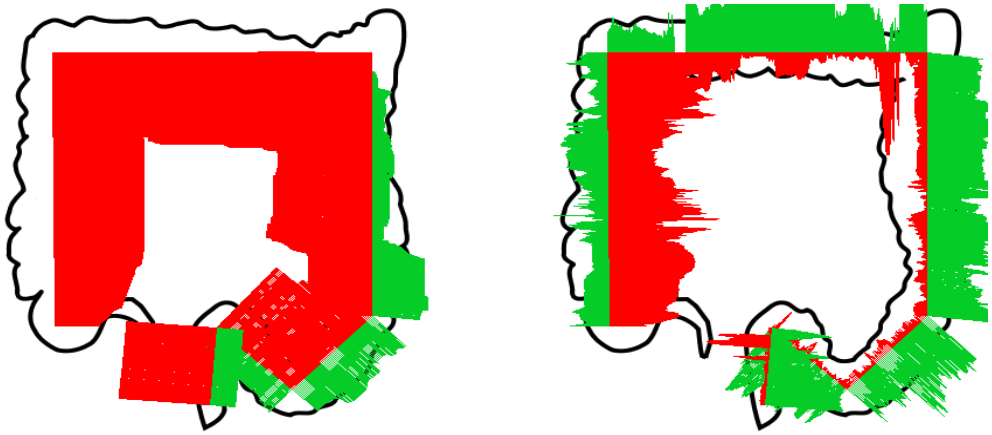


Figure B.30: Distribution of the UC lesions for the patient 41. Medical annotations on the left, and automatic detection on the right.

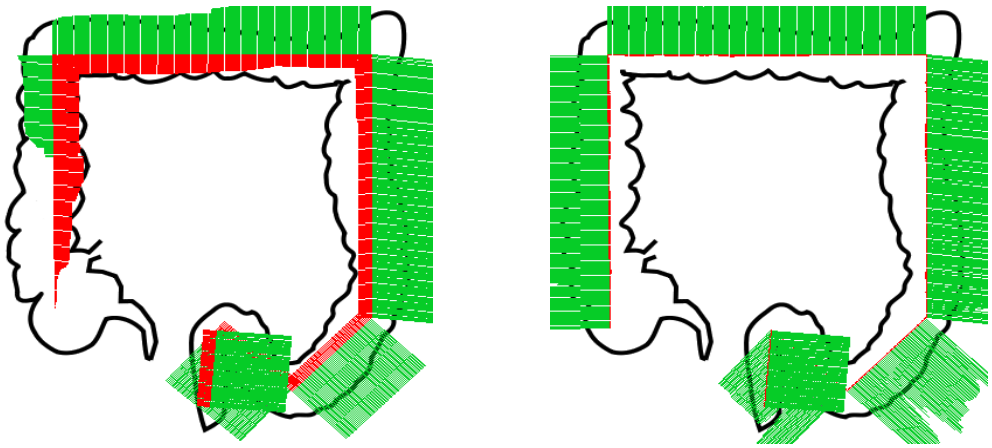


Figure B.31: Distribution of the UC lesions for the patient 42. Medical annotations on the left, and automatic detection on the right.

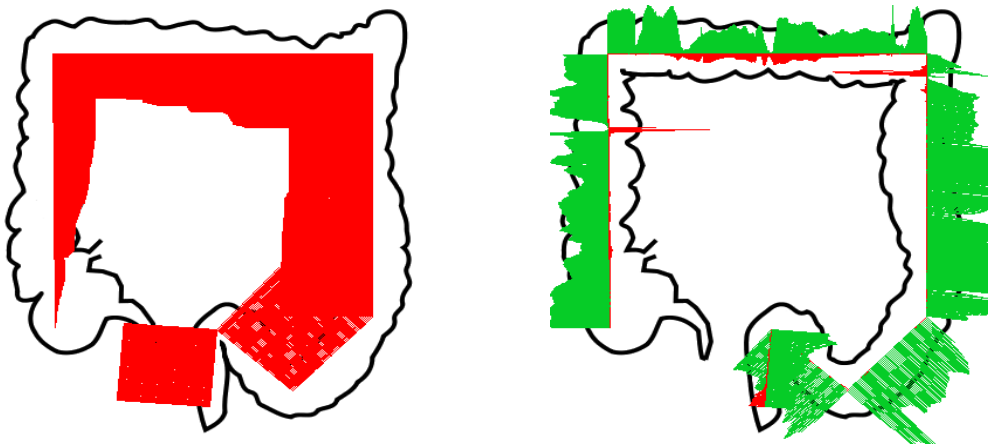


Figure B.32: Distribution of the UC lesions for the patient 43. Medical annotations on the left, and automatic detection on the right.

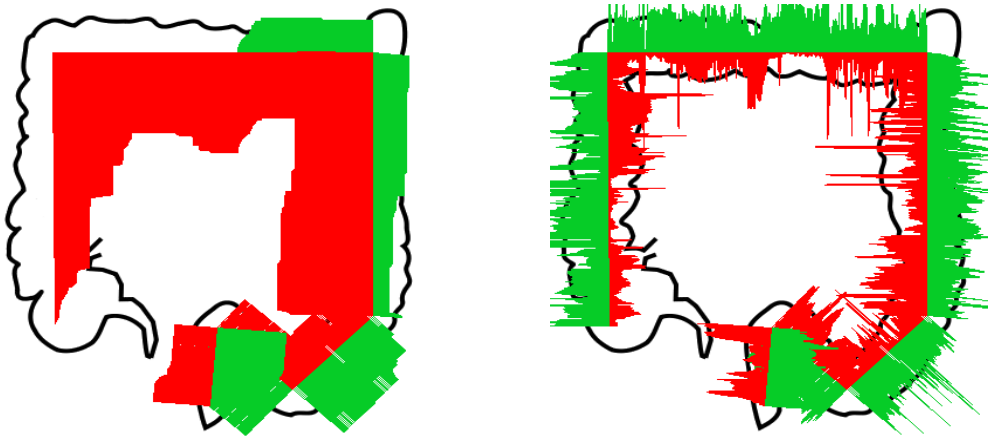


Figure B.33: Distribution of the UC lesions for the patient 44. Medical annotations on the left, and automatic detection on the right.

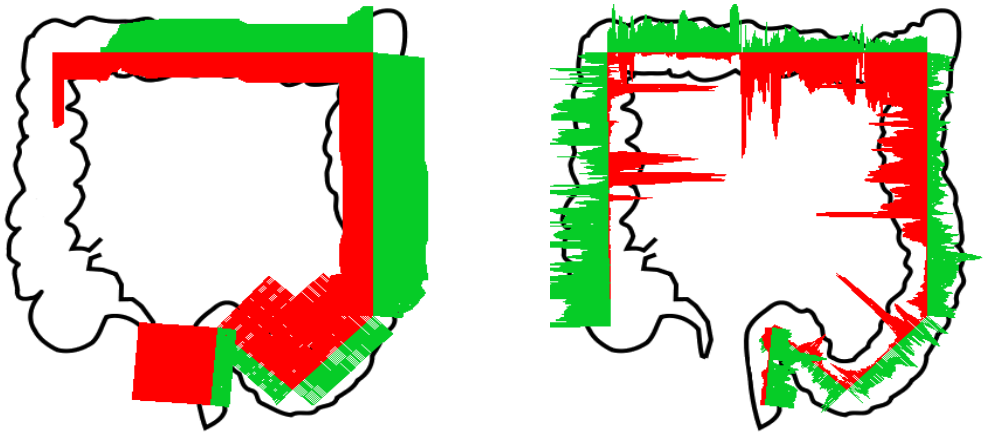


Figure B.34: Distribution of the UC lesions for the patient 46. Medical annotations on the left, and automatic detection on the right.

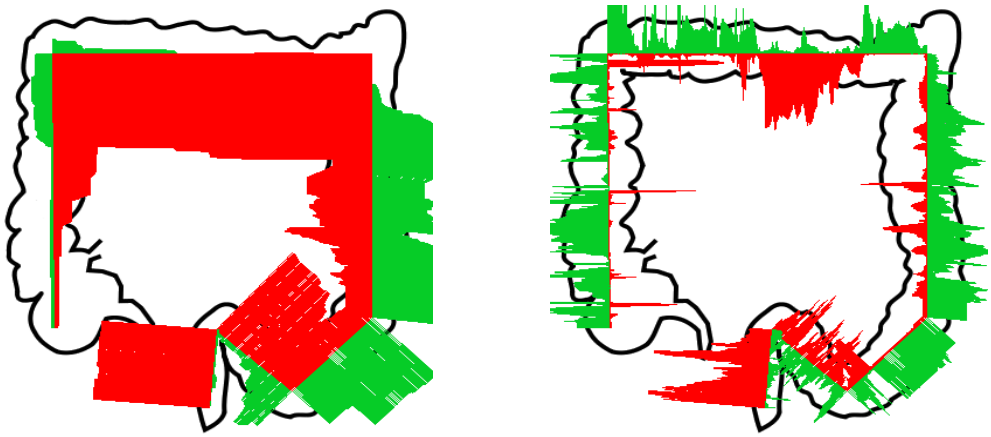


Figure B.35: Distribution of the UC lesions for the patient 47. Medical annotations on the left, and automatic detection on the right.

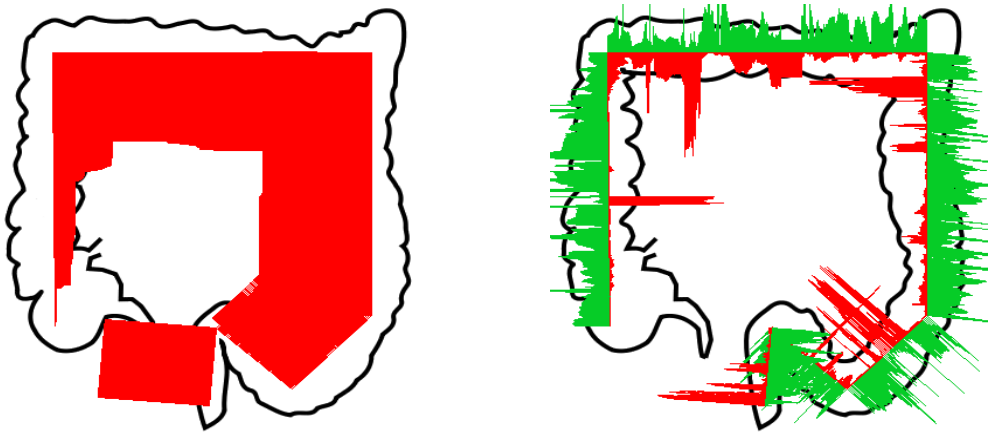


Figure B.36: Distribution of the UC lesions for the patient 48. Medical annotations on the left, and automatic detection on the right.

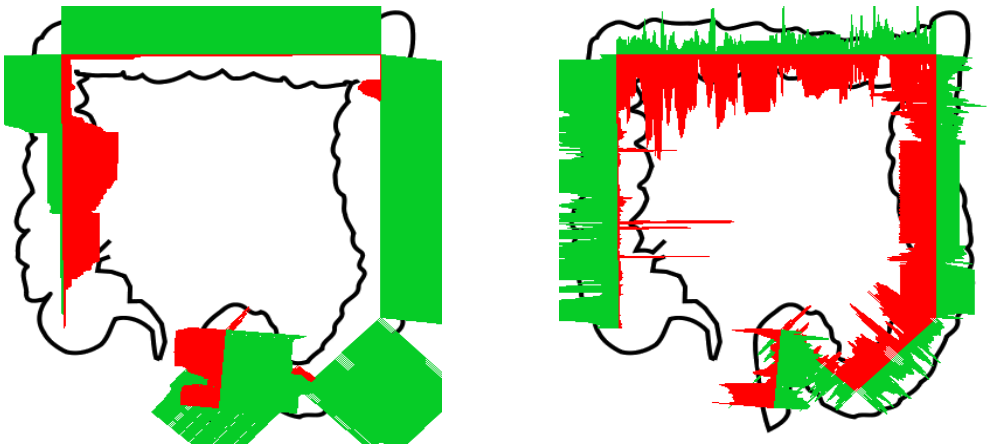


Figure B.37: Distribution of the UC lesions for the patient 49. Medical annotations on the left, and automatic detection on the right.

Bibliography

[Ach+12a] Radhakrishna Achanta et al. "SLIC superpixels compared to state-of-the-art superpixel methods". In: *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012), pp. 2274–2282.

[Ach+12b] Radhakrishna Achanta et al. "SLIC superpixels compared to state-of-the-art superpixel methods". In: *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012), pp. 2274–2282.

[AM17] Bader Saad Alshammari and Daoud Suleiman Mashat. "Numerical study of Fisher's equation by finite difference schemes". In: *Applied Mathematics* 8.8 (2017), pp. 1100–1116.

[ARG15] Mary A Atia, Francisco C Ramirez, and Suryakanth R Gurudu. "Quality monitoring in colonoscopy: Time to act". In: *World journal of gastrointestinal endoscopy* 7.4 (2015), p. 328.

[AZ79] Mark J Ablowitz and Anthony Zeppetella. "Explicit solutions of Fisher's equation for a special wave speed". In: *Bulletin of Mathematical Biology* 41.6 (1979), pp. 835–840.

[Bál+18] Anita Bálint et al. "How disease extent can be included in the endoscopic activity index of ulcerative colitis: the panMayo score, a promising scoring system". In: *BMC gastroenterology* 18.1 (2018), p. 7.

[Bar+85] GI Barenblatt et al. "The mathematical theory of combustion and explosions". In: *Consultants Bureau* (1985).

[Bas+12] Md Rubel Basar et al. "Ingestible wireless capsule technology: A review of development and future indication". In: *International Journal of Antennas and Propagation* 2012 (2012).

[BCLJ64a] JH Baron, AM Connell, and JE Lennard-Jones. "Variation between observers in describing mucosal appearances in proctocolitis". In: *British Medical Journal* 1.5375 (1964), p. 89.

[BCLJ64b] JH Baron, AM Connell, and JE Lennard-Jones. "Variation between observers in describing mucosal appearances in proctocolitis". In: *British medical journal* 1.5375 (1964), p. 89.

[BD15] Éric Brunet and Bernard Derrida. "An exactly solvable travelling wave equation in the Fisher–KPP class". In: *Journal of Statistical Physics* 161.4 (2015), pp. 801–820.

[Bia+19] Livia Biancone et al. "Cancer Risk in Inflammatory Bowel Disease: A 6-Year Prospective Multicenter Nested Case–Control IG-IBD Study". In: *Inflammatory Bowel Diseases* 26.3 (July 2019), pp. 450–459. ISSN: 1078-0998. DOI: 10.1093/ibd/izz155. eprint: https://academic.oup.com/ibdjournal/article-pdf/26/3/450/32436956/izz155.pdf. URL: https://doi.org/10.1093/ibd/izz155.

[BKL11] Giorgos Bamias, Garyfallia Kaltsa, and Spiros D Ladas. "Cytokines in the pathogenesis of ulcerative colitis". In: *Discovery medicine* 11.60 (2011), pp. 459–467.

[BT11] Nick H Barton and Michael Turelli. "Spatial waves of advance with bistable dynamics: cytoplasmic and genetic analogues of Allee effects". In: *The American Naturalist* 178.3 (2011), E48–E75.

[Buc+18] Sarah Buckley et al. "Endoscopic Appendectomy in a Patient with Malrotation and Previous Ladd Procedure: 1703". In: *Official journal of the American College of Gastroenterology | ACG* 113 (2018), S971–S972.

[CGGCA21] Miguel Camba-Gómez, Oreste Gualillo, and Javier Conde-Aranda. "New Perspectives in the Study of Intestinal Inflammation: Focus on the Resolution of Inflammation". In: *International Journal of Molecular Sciences* 22.5 (2021), p. 2605.

[Che+21] Tao Chen et al. "Transcolonic endoscopic appendectomy: a novel natural orifice transluminal endoscopic surgery (NOTES) technique for the sessile serrated lesions involving the appendiceal orifice". In: *Gut* (2021). ISSN: 0017-5749. DOI: 10.1136/gutjnl-2020-323018. eprint: https://gut.bmj.com/content/early/2021/01/21/gutjnl-2020-323018.full.pdf. URL: https://gut.bmj.com/content/early/2021/01/21/gutjnl-2020-323018.

[Chi+18] Nai-Hsuan Chien et al. "Cold snare polypectomy vs cold forceps biopsy in endoscopic treatment of colonic small and diminutive polyps—Effectiveness and safety in the real world". In: *Advances in Digestive Medicine* 5.4 (2018), pp. 127–131. DOI: https://doi.org/10.1002/aid2.13093. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aid2.13093. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/aid2.13093.

[CLT04] Martyn J Carter, Alan J Lobo, and Simon PL Travis. "Guidelines for the management of inflammatory bowel disease in adults". In: *Gut* 53.suppl 5 (2004), pp. v1–v16.

[CMD11] Gastone Ciuti, Arianna Menciassi, and Paolo Dario. "Capsule Endoscopy: From Current Achievements to Open Challenges". In: *IEEE reviews in biomedical engineering* 4 (Jan. 2011), pp. 59–72. DOI: 10.1109/RBME.2011.2171182.

[CMH15] Shannon Chang, Lisa Malter, and David Hudesman. "Disease monitoring in inflammatory bowel disease". In: *World Journal of Gastroenterology: WJG* 21.40 (2015), p. 11246.

[CMN17] Sharmista Chatterjee, Christine MV Moore, and Moheb M Nasr. "An overview of the role of mathematical models in implementation of quality by design paradigm for drug development and manufacture". In: *Comprehensive Quality by Design for Pharmaceutical Product Development and Manufacture* 1 (2017).

[Cos+11] Jacques Cosnes et al. "Epidemiology and natural history of inflammatory bowel diseases". In: *Gastroenterology* 140.6 (2011), pp. 1785–1794.

[Dap+04] Marco Daperno et al. "Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD". In: *Gastrointestinal endoscopy* 60.4 (2004), pp. 505–512.

[Dav+21] Vivek Pravin Dave et al. "Intraocular endoscopy: A review". In: *Indian Journal of Ophthalmology* 69.1 (2021), p. 14.

[Dee+18] Farah Deeba et al. "Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection". In: *Biomedical Signal Processing and Control* 40 (2018), pp. 415–424.

[Dee20] Parakkal Deepak. *Reversing the Tide: Improving the Recruitment of Patients With Inflammatory Bowel Disease in Clinical Trials in the United States*. 2020.

[DF11] Silvio Danese and Claudio Fiocchi. "Ulcerative Colitis". In: *New England Journal of Medicine* 365.18 (2011). PMID: 22047562, pp. 1713–1725. DOI: 10.1056/NEJMra1102942. eprint: https://doi.org/10.1056/NEJMra1102942. URL: https://doi.org/10.1056/NEJMra1102942.

[D'h+07] Geert D'haens et al. "A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis". In: *Gastroenterology* 132.2 (2007), pp. 763–786.

[DLLA04] Thomas De Lange, Stig Larsen, and Lars Aabakken. "Inter-observer agreement in the assessment of endoscopic findings in ulcerative colitis". In: *BMC gastroenterology* 4.1 (2004), pp. 1–10.

[Eri02] Bradley J Erickson. "Irreversible compression of medical images". In: *Journal of Digital Imaging* 15.1 (2002), pp. 5–14.

[Fan+18] Shanhui Fan et al. "Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images". In: *Physics in Medicine & Biology* 63.16 (2018), p. 165001.

[Fea+05] Brian G Feagan et al. "Treatment of ulcerative colitis with a humanized antibody to the $\alpha 4\beta 7$ integrin". In: *New England Journal of Medicine* 352.24 (2005), pp. 2499–2507.

[FMG11] Yanan Fu, Mrinal Mandal, and Gencheng Guo. "Bleeding region detection in WCE images based on color features and neural network". In: *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE. 2011, pp. 1–4.

[FMM80] RICHARD G FARMER, WILLIAM M MICHENER, and EDWARD A MORTIMER. "Studies of family history among patients

with inflammatory bowel disease". In: *Clinics in gastroenterology* 9.2 (1980), pp. 271–278.

[Fu+14] Y. Fu et al. "Computer-Aided Bleeding Detection in WCE Video". In: *IEEE Journal of Biomedical and Health Informatics* 18.2 (2014), pp. 636–642.

[Gau+19] Anchal Gaur et al. "CHALLENGES IN DEVELOPMENT OF COLON DRUG DELIVERY SYSTEM WITH RESPECT TO DISSOLUTION STUDIES". In: *WORLD JOURNAL OF PHARMACY AND PHARMACEUTICAL SCIENCES* 8 (July 2019), pp. 1465–1491. DOI: 10.20959/wjpps20196-14051.

[Gaw+17] Beata Gawrońska et al. "Markers of inflammation and influence of nitric oxide on platelet activation in the course of ulcerative colitis". In: *Oncotarget* 8.40 (2017), p. 68108.

[GDS16] Tonmoy Ghosh, Antara Das, and Rosni Sayed. "Automatic small intestinal ulcer detection in capsule endoscopy images". In: *International Journal of Scientific and Engineering Research* 7.10 (2016), pp. 737–741.

[GFW18] Tonmoy Ghosh, Shaikh Anowarul Fattah, and Khan A Wahid. "CHOBS: color histogram of block statistics for automatic bleeding detection in wireless capsule endoscopy video". In: *IEEE journal of translational engineering in health and medicine* 6 (2018), pp. 1–12.

[Gho+14a] Tonmoy Ghosh et al. "A statistical feature based novel method to detect bleeding in wireless capsule endoscopy images". In: *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*. IEEE. 2014, pp. 1–4.

[Gho+14b] Tonmoy Ghosh et al. "An automatic bleeding detection scheme in wireless capsule endoscopy based on histogram of an RGB-indexed image". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2014, pp. 4683–4686.

[Gho+15] Tonmoy Ghosh et al. "An automatic bleeding detection technique in wireless capsule endoscopy from region of interest". In: *2015 IEEE International Conference on Digital Signal Processing (DSP)*. IEEE. 2015, pp. 1293–1297.

[Gho+18] Tonmoy Ghosh et al. "Cluster based statistical feature extraction method for automatic bleeding detection in wireless capsule endoscopy video". In: *Computers in biology and medicine* 94 (2018), pp. 41–54.

[GL+15] Yago González-Lama et al. "Impact of endoscopic monitoring in postoperative Crohn's disease patients already receiving pharmacological prevention of recurrence". In: *Revista Española de Enfermedades Digestivas* 107.10 (2015), pp. 586–590.

[Goe18] Martin Goetz. "Endoscopic surveillance in inflammatory bowel disease". In: *Visceral medicine* 34.1 (2018), pp. 66–71.

[Gom+86] P Gomes et al. "Relationship between disease activity indices and colonoscopic findings in patients with colonic inflammatory bowel disease." In: *Gut* 27.1 (1986), pp. 92–95.

[GQ18] Anna Geyer and Ronald Quirchmayr. "Traveling wave solutions of a highly nonlinear shallow water equation". In: *Discrete & Continuous Dynamical Systems-A* 38.3 (2018), p. 1567.

[HJ18] Xiaochuan Hu and Sophia R.-J. Jang. "Dynamics of tumor–CD4+–cytokine–host cells interactions with treatments". In: *Applied Mathematics and Computation* 321 (2018), pp. 700–720. ISSN: 0096-3003. DOI: https://doi.org/10.1016/j.amc.2017.11.009. URL: https://www.sciencedirect.com/science/article/pii/S0096300317307828.

[Idd+00] Gavriel Iddan et al. "Wireless capsule endoscopy". In: *Nature* 405.6785 (2000), pp. 417–417.

[JSW13] Simon Jäger, Eduard F Stange, and Jan Wehkamp. "Inflammatory bowel disease: an impaired barrier disease". In: *Langenbeck's archives of surgery* 398.1 (2013), pp. 1–12.

[Ken38] Maurice G Kendall. "A new measure of rank correlation". In: *Biometrika* 30.1/2 (1938), pp. 81–93.

[KFR18] Amit Kumar Kundu, Shaikh Anowarul Fattah, and Mamshad Nayeem Rizve. "An automatic bleeding frame and region detection scheme for wireless capsule endoscopy videos based on interplane intensity variation profile in normalized RGB color space". In: *Journal of healthcare engineering* 2018 (2018).

[KG15] Nimisha Elsa Koshy and Varun P Gopi. "A new method for ulcer detection in endoscopic images". In: *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. IEEE. 2015, pp. 1725–1729.

[Kir01] Joseph B Kirsner. "Historical origins of current IBD concepts". In: *World journal of gastroenterology* 7.2 (2001), p. 175.

[Kol37] Andrei N Kolmogorov. "Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique". In: *Bull. Univ. Moskow, Ser. Internat., Sec. A* 1 (1937), pp. 1–25.

[Kri08] KK Kristiansen. "Reaction Diffusion Models in Mathematical Biology Technical University of Denmark". In: *The Department of Mathematics, Lyngby, Denmark* (2008).

[KS04] Asher Kornbluth and David B Sachar. "Ulcerative colitis practice guidelines in adults (update): American College of Gastroenterology, Practice Parameters Committee". In: *American Journal of Gastroenterology* 99.7 (2004), pp. 1371–1385.

[Kum+] R Vijaya Kumar et al. "Grey level to RGB using YCbCr color space Technique". In: *International Journal of Computer Applications* 975 (), p. 8887.

[Kut11] Christina Kuttler. "Reaction-diffusion equations with applications". In: *Internet seminar*. 2011.

[LAF13a] Wing-Cheong Lo, Razvan I Arsenescu, and Avner Friedman. "Mathematical model of the roles of T cells in inflammatory bowel disease". In: *Bulletin of mathematical biology* 75.9 (2013), pp. 1417–1433.

[LAF13b] Wing-Cheong Lo, Razvan I Arsenescu, and Avner Friedman. "Mathematical model of the roles of T cells in inflammatory bowel disease". In: *Bulletin of mathematical biology* 75.9 (2013), pp. 1417–1433.

[Lai+09] Edwin J Lai et al. "The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research". In: *Gastrointestinal endoscopy* 69.3 (2009), pp. 620–625.

[Lan+07] Robert C Langan et al. "Ulcerative colitis: diagnosis and treatment". In: *American family physician* 76.9 (2007), pp. 1323–1330.

[Lew+08] James D Lewis et al. "Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis". In: *Inflammatory bowel diseases* 14.12 (2008), pp. 1660–1666.

[LKM19] Joon Seop Lee, Eun Soo Kim, and Won Moon. "Chronological review of endoscopic indices in inflammatory bowel disease". In: *Clinical endoscopy* 52.2 (2019), p. 129.

[LL16] Ji Min Lee and Kang-Moon Lee. "Endoscopic diagnosis and differentiation of inflammatory bowel disease". In: *Clinical endoscopy* 49.4 (2016), p. 370.

[LM01] Thomas Leung and Jitendra Malik. "Representing and recognizing the visual appearance of materials using three-dimensional textons". In: *International journal of computer vision* 43.1 (2001), pp. 29–44.

[Lo+16] Wing-Cheong Lo et al. "Inflammatory bowel disease: How effective is TNF-α suppression?" In: *PloS one* 11.11 (2016), e0165782.

[LPF20] Jimmy K. Limdi, Michael Picco, and Francis A. Farraye. "A review of endoscopic scoring systems and their importance in a treat-to-target approach in inflammatory bowel disease (with videos)". In: *Gastrointestinal Endoscopy* 91.4 (2020), pp. 733–745. ISSN: 0016-5107. DOI: https://doi.org/10.1016/j.gie.2019.11.032. URL: https://www.sciencedirect.com/science/article/pii/S0016510719324824.

[Mai+15] Shreya Maiti et al. "Mathematical modeling of pro-and anti-inflammatory signaling in macrophages". In: *Processes* 3.1 (2015), pp. 1–18.

[Mak+20] Wing Yan Mak et al. "The epidemiology of inflammatory bowel disease: east meets west". In: *Journal of gastroenterology and hepatology* 35.3 (2020), pp. 380–389.

[Mar06] John K Marshall. "Reporting ileocolonoscopy in inflammatory bowel disease". In: *Canadian Journal of Gastroenterology* 20.6 (2006), pp. 399–400.

[Mat+04] K Matsuoka et al. "T-bet upregulation and subsequent interleukin 12 stimulation are essential for induction of Th1 mediated immunopathology in Crohn's disease". In: *Gut* 53.9 (2004), pp. 1303–1308.

[Met96] Ralph Metson. "Holmium: YAG Laser Endoscopic Sinus Surgery; A Randomized, Controlled Study". In: *The Laryngoscope* 106.S77 (1996), pp. 1–18.

[MF96] Wen-Xiu Ma and Benno Fuchssteiner. "Explicit and exact solutions to a Kolmogorov-Petrovskii-Piskunov equation". In: *International Journal of Non-Linear Mechanics* 31.3 (1996), pp. 329–338.

[MJ+17] Ignacio Marín-Jiménez et al. "Management of the psychological impact of inflammatory bowel disease: perspective of doctors and patients—the ENMENTE Project". In: *Inflammatory bowel diseases* 23.9 (2017), pp. 1492–1498.

[ML04] Helen Moore and Natasha K Li. "A mathematical model for chronic myelogenous leukemia (CML) and T cell interaction". In: *Journal of theoretical biology* 227.4 (2004), pp. 513–523.

[MM89] Jean-Yves Mary and Robert Modigliani. "Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID)." In: *Gut* 30.7 (1989), pp. 983–989.

[Mon+87] U Monsen et al. "Prevalence of inflammatory bowel disease among relatives of patients with ulcerative colitis". In: *Scandinavian journal of gastroenterology* 22.2 (1987), pp. 214–218.

[MWLJ13] Jerome L Myers, Arnold D Well, and Robert F Lorch Jr. *Research design and statistical analysis*. Routledge, 2013.

[MZ05] Shiwang Ma and Xingfu Zou. "Existence, uniqueness and stability of travelling waves in a discrete reaction–diffusion monostable equation with delay". In: *Journal of Differential Equations* 217.1 (2005), pp. 54–87.

[Nag+10] Makoto Naganuma et al. "Novel endoscopic activity index is useful for choosing treatment in severe active ulcerative colitis patients". In: *Journal of gastroenterology* 45.9 (2010), pp. 936–943.

[Naw+14] Ruwan Nawarathna et al. "Abnormal image detection in endoscopy videos using a filter bank and local binary patterns". In: *Neurocomputing* 144 (2014), pp. 70–91.

[OEI18] Maxwell Scale Uwadia Osagie, Osatohanmwen Enagbonma, and Amanda Iriagbonse Inyang. "Structural Dynamics and Evolution of Capsule Endoscopy (Pill Camera) Technology in Gastroenterologist Assertion". In: *arXiv preprint arXiv:1804.01171* (2018).

[Ord+12] Ingrid Ordás et al. "Ulcerative colitis". In: *The Lancet* 380.9853 (2012), pp. 1606–1619. ISSN: 0140-6736. DOI: https://doi.org/10.1016/S0140-6736(12)60150-0. URL: https://www.sciencedirect.com/science/article/pii/S0140673612601500.

[Orl+98] F Orlandi et al. "Observer agreement in endoscopic assessment of ulcerative colitis." In: *Italian journal of gastroenterology and hepatology* 30.5 (1998), pp. 539–541.

[Pan+11] Guobing Pan et al. "Bleeding detection in wireless capsule endoscopy based on probabilistic neural network". In: *Journal of medical systems* 35.6 (2011), pp. 1477–1484.

[Pan+16] Julián Panés et al. "Central endoscopy reading in inflammatory bowel diseases". In: *Journal of Crohn's and Colitis* 10.suppl_2 (2016), S542–S547.

[PB+15] L Peyrin-Biroulet et al. "Selecting therapeutic targets in inflammatory bowel disease (STRIDE): determining therapeutic goals for treat-to-target". In: *Official journal of the American College of Gastroenterology | ACG* 110.9 (2015), pp. 1324–1338.

[Per15] Benoît Perthame. "Parabolic equations in biology". In: *Parabolic Equations in Biology*. Springer, 2015, pp. 1–21.

[Per+87] A Pera et al. "Colonoscopy in inflammatory bowel disease: diagnostic accuracy and proposal of an endoscopic score". In: *Gastroenterology* 92.1 (1987), pp. 181–185.

[Pig+15] Charly Pignon et al. "Flexible gastrointestinal endoscopy in ferrets (Mustela putorius furo)". In: *Veterinary Clinics: Exotic Animal Practice* 18.3 (2015), pp. 369–400.

[PJ16] Anna Park and Il Hyo Jung. "A MATHEMATICAL MODEL OF IMMUNE-MEDIATED DISORDER IN INFLAMMATORY BOWEL DISEASE". In: *East Asian mathematical journal* 32.1 (2016), pp. 139–152.

[PKH20] RJ Porter, R Kalla, and GT Ho. "Ulcerative colitis: Recent advances in the understanding of disease pathogenesis [version 1; peer review: 2 approved]". In: *F1000Research* 9.294 (2020). DOI: 10.12688/f1000research.20805.1.

[Pog+19] Konstantin Pogorelov et al. "Bleeding detection in wireless capsule endoscopy videos—Color versus texture features". In: *Journal of applied clinical medical physics* 20.8 (2019), pp. 141–154.

[Pro+18] Fay Probert et al. "Plasma nuclear magnetic resonance metabolomics discriminates between high and low endoscopic activity and predicts progression in a prospective cohort of patients with ulcerative colitis". In: *Journal of Crohn's and Colitis* 12.11 (2018), pp. 1326–1337.

[Pub] *The capsule endoscopy website. [Online]. Available: http://www.capsuleendoscopy.org.* 2014. URL: "http://www.capsuleendoscopy.org.".

[RA12] Rajaratnam Rameshshanker and Naila Arebi. "Endoscopy in inflammatory bowel disease when and why". In: *World journal of gastrointestinal endoscopy* 4.6 (2012), p. 201.

[Rac89] Dob Rachmilewitz. "Coated mesalazine (5-aminosalicylic acid) versus sulphasalazine in the treatment of active ulcerative colitis: a randomised trial." In: *Bmj* 298.6666 (1989), pp. 82–86.

[RBC10] Priti Kumar Roy, Jhumpa Bhadra, and Biplab Chattopadhyay. "Mathematical modeling on immunopathogenesis in chronic plaque of psoriasis: A theoretical study". In: *Proceedings of the World Congress on Engineering*. Vol. 1. 2010, pp. 550–555.

[Rex07] Douglas K Rex. "Colonoscopy withdrawal times and adenoma detection rates". In: *Gastroenterology & hepatology* 3.8 (2007), p. 609.

[Rut+05] Paul Rutgeerts et al. "Infliximab for induction and maintenance therapy for ulcerative colitis". In: *New England Journal of Medicine* 353.23 (2005), pp. 2462–2476.

[Sam+13] Sunil Samuel et al. "Validation of the ulcerative colitis colonoscopic index of severity and its correlation with disease activity measures". In: *Clinical Gastroenterology and Hepatology* 11.1 (2013), pp. 49–54.

[San02] Björn Sandstede. "Stability of travelling waves". In: *Handbook of dynamical systems*. Vol. 2. Elsevier, 2002, pp. 983–1055.

[SBW14] Sonu Sainju, Francis M Bui, and Khan A Wahid. "Automated bleeding detection in capsule endoscopy videos using statistical features and region growing". In: *Journal of medical systems* 38.4 (2014), p. 25.

[SM03] Endre Süli and David F Mayers. *An introduction to numerical analysis*. Cambridge university press, 2003.

[SM96] JA Sherratt and BP Marchant. "Nonsharp travelling wave fronts in the Fisher equation with degenerate nonlinear diffusion". In: *Applied Mathematics Letters* 9.5 (1996), pp. 33–38.

[Smo12] Joel Smoller. *Shock waves and reaction—diffusion equations*. Vol. 258. Springer Science & Business Media, 2012.

[Son16] Amnon Sonnenberg. "Sedation in colonoscopy". In: *Gastroenterology & hepatology* 12.5 (2016), p. 327.

[SP21] Clyde M Stauffer and Christopher Pfeifer. "Colonoscopy". In: *StatPearls [Internet]* (2021).

[Sti89] Stephen M. Stigler. "Francis Galton's Account of the Invention of Correlation". In: *Statistical Science* 4.2 (1989), pp. 73 –79. DOI: 10.1214/ss/1177012580. URL: https://doi.org/10.1214/ss/1177012580.

[Sum+15] Shipra Suman et al. "Optimum colour space selection for ulcerated regions using statistical analysis and classification of ulcerated frames from wce video footage". In: *International Conference on Neural Information Processing*. Springer. 2015, pp. 373–381.

[Sum+17a] Shipra Suman et al. "Detection and classification of bleeding region in WCE images using color feature". In: *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*. 2017, pp. 1–6.

[Sum+17b] Shipra Suman et al. "Feature selection and classification of ulcerated lesions using statistical analysis for WCE images". In: *Applied Sciences* 7.10 (2017), p. 1097.

[Sut+87] Lloyd R Sutherland et al. "5-Aminosalicylic acid enema in the treatment of distal ulcerative colitis, proctosigmoiditis, and proctitis". In: *Gastroenterology* 92.6 (1987), pp. 1894–1898.

[SW11] Claude Sammut and Geoffrey I Webb. *Encyclopedia of machine learning*. Springer Science & Business Media, 2011.

[Tra+12] Simon PL Travis et al. "Developing an instrument to assess the endoscopic severity of ulcerative colitis: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS)". In: *Gut* 61.4 (2012), pp. 535–542.

[Tra+13] Simon PL Travis et al. "Reliability and initial validation of the ulcerative colitis endoscopic index of severity". In: *Gastroenterology* 145.5 (2013), pp. 987–995.

[TT08] E Telakis and E Tsironi. "Indeterminate colitis-definition, diagnosis, characteristics and management". In: *Annals of Gastroenterology* (2008), pp. 173–180.

[Tur90] Alan Mathison Turing. "The chemical basis of morphogenesis". In: *Bulletin of mathematical biology* 52.1-2 (1990), pp. 153–197.

[TV20] Sergei Trofimchuk and Vitaly Volpert. "Traveling waves in delayed reaction-diffusion equations in biology [J]". In: *Mathematical Biosciences and Engineering* 17.6 (2020), pp. 6487–6514.

[TW55] Sidney C Truelove and LJ Witts. "Cortisone in ulcerative colitis". In: *British medical journal* 2.4947 (1955), p. 1041.

[Ura+14] Toshio Uraoka et al. "Cold polypectomy techniques for diminutive polyps in the colorectum". In: *Digestive Endoscopy* 26 (2014), pp. 98–103.

[Vas+18] Nadia Mohammed Vashist et al. "Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis". In: *Cochrane Database of Systematic Reviews* 1 (2018).

[VP09] Vitaly Volpert and Sergei Petrovskii. "Reaction–diffusion waves in biology". In: *Physics of life reviews* 6.4 (2009), pp. 267–310.

[VPR13] Carl Vondrick, Donald Patterson, and Deva Ramanan. "Efficiently scaling up crowdsourced video annotation". In: *International journal of computer vision* 101.1 (2013), pp. 184–204.

[VVV94] AI Volpert, Vi A Volpert, and Vl A Volpert. "Traveling waves solutions of parabolic systems. Transl. Math. Mono. 140". In: *Amer. Math. Soc., Providence* (1994).

[Wal+98] RS Walmsley et al. "A simple clinical colitis activity index". In: *Gut* 43.1 (1998), pp. 29–32.

[Wen+10] Katherine Wendelsdorf et al. "Model of colonic inflammation: immune modulatory mechanisms in inflammatory bowel disease". In: *Journal of theoretical biology* 264.4 (2010), pp. 1225–1239.

[YCS04] Andrew Yates, Robin Callard, and Jaroslav Stark. "Combining cytokine signalling with T-bet and GATA-3 regulation in Th1 and Th2 differentiation: a model for cellular decision-making". In: *Journal of theoretical biology* 231.2 (2004), pp. 181–196.

[Ye+15] Yulan Ye et al. "The epidemiology and risk factors of inflammatory bowel disease". In: *International journal of clinical and experimental medicine* 8.12 (2015), p. 22529.

[You50] William J Youden. "Index for rating diagnostic tests". In: *Cancer* 3.1 (1950), pp. 32–35.

[Yua+15] Yixuan Yuan et al. "Saliency based ulcer detection for wireless capsule endoscopy diagnosis". In: *IEEE transactions on medical imaging* 34.10 (2015), pp. 2046–2057.

[YWT+14] Jinn-Yi Yeh, Tai-Hsi Wu, Wei-Jun Tsai, et al. "Bleeding and ulcer detection using wireless capsule endoscopy images". In: *Journal of Software Engineering and Applications* 7.05 (2014), p. 422.