



University of Coimbra  
Faculty of Sciences and Technology  
Department of Physics

Protocol and Support Infrastructure for the creation  
of an annotated database of images by Confocal  
Endomicroscopy

Aniana da Rosa de Brito

Coimbra, 2011

# Protocol and Support Infrastructure for the creation of an annotated database of images by Confocal Endomicroscopy

Advisor: Prof. Dr. João P. Barreto (ISR-FCTUC)

Co-Advisor: Prof. Dr. Pedro Figueiredo (HUC-FMUC)

Thesis for obtaining the degree of Integrated Master in  
Biomedical Engineering

Committee:

Prof. Dr. António Miguel Lino Santos Morgado

Prof. Dr. Paulo José Monteiro Peixoto

Prof. Dr. João Pedro de Almeida Barreto

Prof. Dr. Pedro Manuel Narra Figueiredo

September 2011



# Acknowledgements

To do this project and complete my course I had help from several persons. So I want to take this opportunity to express my gratitude for all these people.

Foremost, I want to thank my advisor Prof. DR. João Pedro Barreto, for the support and aid along all this work and constant guidance so dedicated where he always find time for discussions.

I want to thank my co-advisor Prof. Dr. Pedro Figueiredo and Dr. Paulo Freire, for the support in medical component which enabled to build the data base.

I want to thank all my lab friends to be always ready to help me in any difficulty. Special thanks for Melo for the proofreading of this thesis that it allows improved this document substantially.

Thanks to all my friends and colleague for all good times that we shared together during these years.

A grateful acknowledgement goes to my boyfriend Aricson Cruz for helping me to overcome the most difficult moments. Thanks for standing by my side when I most needed.

As the last is the first, the big acknowledgement goes to my family especially to my parents Olívio Brito and Cecília da Rosa who support me in every moment of my life. Special thanks for my grandmother Domingas Brito who sadly died this year. *“Vovô un ta lembra d bo para sempre”*.



# Abstract

Confocal laser endomicroscopy (CEM) is a new diagnosis technique that enables the histological examination of suspicious tissues in real-time during an ongoing endoscopy. The main advantage of the technique is that it avoids tissue biopsy for lab analysis, providing the doctor with the means to make an immediate diagnosis.

Two main difficulties can be found in endomicroscopy exam: (i) the simultaneous execution of endoscopy and histology exam is very difficult to accomplish in practice and requiring the doctor to go through a very long training period; (ii) this technique only recently was adapted for in vivo histological examination, so the image taxonomy and interpretation is not well defined yet.

The overall goal of this line of research is to develop a computational system that applies pattern recognition techniques for assisting the practitioner during the procedure by performing automatic diagnosis from the CEM images. The main goal of this study is to develop software for collecting and labeling CEM images to build a database, and use this data to develop basic algorithms for detection and segmentation of the cellular structures.

In this thesis we describe the image acquisition protocols and the application developed to acquire the data obtained by endomicroscopy. Regarding the data analysis, a statistical and semantic analysis of images was performed. The statistical analyses of database show that some characteristics like the number and shape of the crypts, allow distinguishing the different classes of the image. The results obtained from semantic analyses, done through texture analyses, show that it does not allow distinguishing the image classes.

As demonstrated the statistical analyses the number and shape of the crypts are the parameters that enable to distinguish the classes. Therefore in this project we segment and detect the crypts. Relatively to the crypts segmentation results, we concluded that the symmetry energy work well in image of normal tissue and image with light inflammation. So it is needed to take into consideration much more information and features.

**Keywords:** Confocal microscopy, endomicroscopy, application, database, segmentation.

# Contents

<b>Introduction .....</b>	<b>1</b>
1.1: A new approach for diagnosis .....	1
1.2: Diagnosis using CEM .....	2
1.3: Objectives of this thesis .....	5
1.4: Related work .....	6
1.5: Overview of the thesis .....	7
<b>The system of Confocal Endomicroscopy.....</b>	<b>8</b>
2.1: Basics of confocal microscopy .....	8
2.2: Confocal Endomicroscopy System Used.....	10
2.3: Collection of data.....	13
<b>Building the database .....</b>	<b>15</b>
3.1 Objectives of the database .....	15
3.2 Design of the data base .....	16
3.3 The tagging application .....	18
3.3.1 Main graphical screens interface.....	21
Schematic summary of the application screens. ....	23
3.4 Future Developments. ....	24
<b>Database Analysis.....</b>	<b>25</b>
4.1 Statistics of the Database .....	25
4.2 Discussion.....	30
4.3 Image Semantic.....	32
4.4 Discussion.....	35
<b>Detection and Segmentation of Crypts .....</b>	<b>37</b>
5.1 How do the symmetries and anti-symmetries work?.....	37
5.2 Algorithm to measure the Symmetry .....	39
5.3 Results.....	41

<b>Conclusion and Future Work.....</b>	<b>48</b>
6.1: Conclusion .....	48
6.2: Future Work.....	49
<b>A Protocol.....</b>	<b>50</b>
<b>B Code to create the Tables of data base.....</b>	<b>60</b>
<b>C Code to process the image in Database.....</b>	<b>63</b>
<b>Bibliography.....</b>	<b>66</b>



# List of Tables

**Table 1:** Types of images without artifacts acquired by endomicroscopy.

**Table 2:** Comparison between endmicroscopic image and image acquired from biopsy analysis.

**Table 3:** Main specifications of the Pentax CEM.

**Table 4:** Parameters that characterize the image in the database.

**Table 5:** Some examples of different classes of image in the database.

**Table 6:** Example of some image obtained from energy of symmetry (column pair) and brightness image (odd column).

# List of Figure

**Figure 1:** Endomicroscopic devices.

**Figure 2:** Horizontal cut of the bowel.

**Figure 3:** Motion artifacts (left) and artifacts due to air bubbles and fecal content (right).

**Figure 4:** Wide-field vs. confocal microscopy, illustrating the principal difference in volumes illumination.

**Figure 5:** Basic principle of confocal microscopy

**Figure 6:** Scheme of confocal laser endomicroscopy. The confocal laser microscope is integrated into the distal end of a conventional endoscope **A**. The blue laser light **B** crosses the mucosal layer at the surface of the intestine till reaching the deep tissue, **C**. **D** is the stack of CEM images that capture the reflection of the laser light in a contrast substance that has been previously injected. Each image of the stack corresponds to an optical horizontal section of the mucosal layer.

**Figure 7: (a):** The confocal laser endoscope with confocal imaging window signaled  
**(b):** Two additional buttons used to control the endoscope during the image acquisition.  
**(c):** Image acquisition screen.

**Figure 8:** Schematic representation of the plan to build the data base.

**Figure 9:** Relational Diagram of the database.

**Figure 10:** Physical Diagram of the database.

**Figure 11:** Schematic representation of the application screen that allows inserting the image in the database – Nova Imagen screen. There are predefined fields already filled.

**Figure 12:** Schematic representation of the application screen that allows annotating the images – Anotação screens.

**Figure 13:** Schematic representation of the application screens.

**Figure 14:** Types of crypts shapes.

**Figure 15:** Types of lamina propria shapes.

**Figure 16:** Types of lamina propria brightness.

**Figure 17:** Types of vases shape.

**Figure 18:** Types of goblet cells.

**Figure 19:** The tree image types in which we use the statistical techniques: the original image, the gradient image magnitude and the gradient image angle.

**Figure 20:** Representation graph of the mean intensity and entropy in original image, gradient image magnitude and gradient image angle for the different classes.

**Figure 21:** Representation graph of the fractal dimension and second order moments in original image, gradient image magnitude and gradient image angle for the different classes.

**Figure 22:** Representation graph of the mean and skewness of the histogram in original image, gradient image magnitude and gradient image angle for the different classes.

**Figure 23:** Representation graph of the energy and contrast of co-occurrence matrices in original image, gradient image magnitude and gradient image angle for the different classes.

**Figure 24:** Representation graph of the skewness in original image, gradient image magnitude and gradient image angle for the different classes.

**Figure 25:** Schematic representation of Fourier series of square wave and of triangular wave.

**Figure 26:** Schematic representation of Gabor pair.

**Figure 27:** Schematic representation of the evolution wavelet amplitude.

**Figure 28:** Schematic representation of the frequency evolution in different scales.

**Figure 29:** Results acquired by Hough transform applied in the symmetry image. The detected circles are marked with red color.

**Figure 30:** Results acquired by Mean Shift applied in image of the symmetry. The clusters centers are marked with blue circle.

**Figure 31:** Results acquired by Active Contour applied in symmetry image. The contours of the crypts are marked in green color. The ellipse fitting was applied on the symmetry image. The ellipses of the crypts are marked in green color.

**Figure 32:** Example of ellipse fitting to detect crypts in image with severe inflammation (left) and image with high probability of neoplasia (right).

**Figure 33:** Image Acquisition screens.

**Figure 34:** Start-up screens.

**Figure 35:** Procedure results search criteria screens.

**Figure 36:** Review procedure results screens.

**Figure 37:** Schematic representation of the first application screen – *Login* screens

**Figure 38:** Schematic representation of the first application screen – *Novo Registro* screens.

**Figure 39:** Schematic representation of the third application screen – *Processos* screens.

**Figure 40:** Schematic representation of the fourth application screen – *Novo Processo* screens.

**Figure 41:** Schematic representation of the fifth application screen – *Exame* screens.

**Figure 42:** Schematic representation of the sixth application screen – *Exames* screens.

**Figure 43:** Schematic representation of the seventh application screen – *Imagens* screens.

**Figure 44:** Schematic representation of the eighth application screen – *Nova Imagem* screens.

**Figure 45:** Schematic representation of the ninth application screen – *Lista de Imagem* screens.

**Figure 46:** Schematic representation of the tenth application screen – *Imagem* screens.

**Figure 47:** Schematic representation of the eleventh application screen – *Anotação* screens.

# List of Graphic

**Graphic 1:** Bar graph of the number of image on data base separated by types.

**Graphic 2:** Representation graph of crypts number distribution in different classes of images.

**Graphic 3:** Representation graph of crypts shape distribution in different classes of images.

**Graphic 4:** Representation graph of lamina propria shape distribution in different classes of images.

**Graphic 5:** Representation graph of lamina propria brightness distribution in different classes of images.

**Graphic 6:** Representation graph of vases types distribution in different classes of images.

**Graphic 7:** Representation graph of goblet cell types distribution in different classes of images.

**Graphic 8:** Graph representation of the percentage of detected crypts acquired by local maximum, Active Contour and ellipse fitting applied in symmetry images.

**Graphic 9:** Graph representation of the percentage of detected contours acquired by local maxima, Active Contour and ellipse fitting applied in symmetry image.

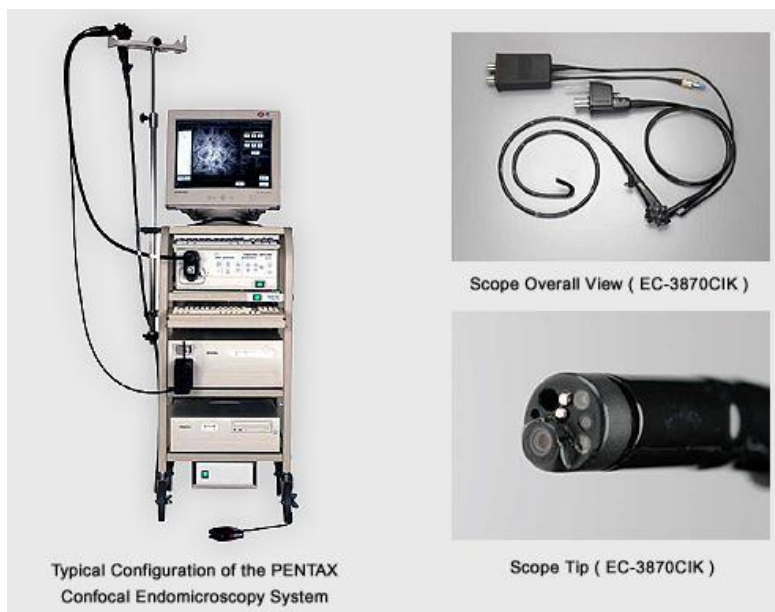


# Chapter 1

## Introduction

### 1.1: A new approach for diagnosis

Confocal microscopy is a technique for histological examination that was invented in the late 1950s. This technique won popularity later at time because the technical limitation did not permit its implementation. It was adapted for in vivo histological examination very recently. The principle of CEM is based on the integration of a confocal laser microscope into the conventional endoscope [1].



**Figure 1:** Endomicroscopic devices.

Confocal laser endomicroscopy is a new diagnostic tool that allows performing histologic examination in vivo during the endoscopy in real-time. Different diseases can be diagnosed immediately using this method. The main advantage of the technique is that it avoids tissue biopsy for lab analysis, providing the doctor with the means to make an immediate diagnosis.

## CHAPTER 1: INTRODUCTION

The early detection of the tumor is essential to its treatment and the analysis in real time of the mucosa of the gastrointestinal (GI) tract is fundamental for preventing the cancer. Although there are techniques that enable the visualization of the mucosa (e.g. chromoendoscopy) it is still necessary to gather the tissue specimen to subsequent laboratory analysis, making the procedure slow and aggressive for the patient. The endomicroscopy opened a new path for the diagnosis of different diseases of the GI-tract because the detection of mucosal lesions can be observed in real time during the ongoing procedure.

Actually there are two confocal endomicroscopy systems available. One of these systems was developed by Pentax (Tokyo, Japan). It consists in the introduction of a confocal laser microscope into the distal end of a conventional endoscope. The main features of this system will be described in the next chapter. The other system was developed by Mauna Kea Technologies (Paris, France), and consists in inserting a catheter probe with a semiconductor laser through the instrument channel of a conventional endoscope.

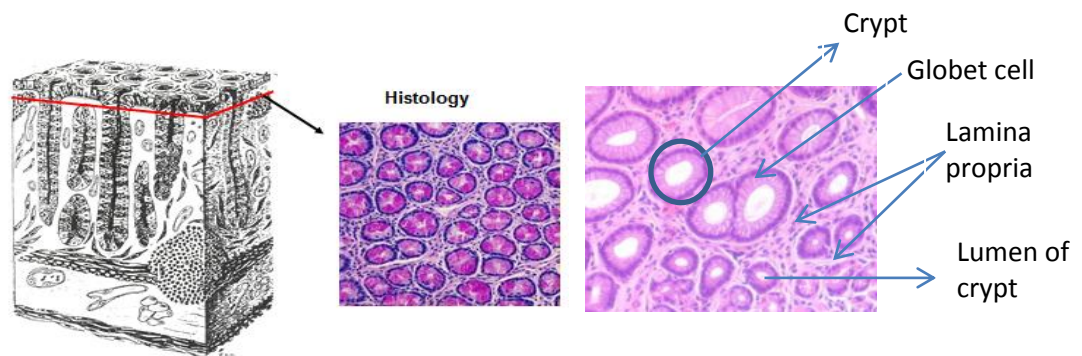
### **1.2: Diagnosis using CEM**

The doctor uses endomicroscopy exam to identify the specific alteration existent in gastrointestinal tract. There are several diseases that affect the gastrointestinal tract but in this project we will focus on the inferior gastrointestinal tract, more specifically in the inflammatory bowel disease (IBD).

IBD, as the own name indicates, is a group of inflammatory conditions of the bowel wall and represents the most severe disease of the GI. Crohn's disease and ulcerative colitis are the main types of IBD. Ulcerative colitis affects only the colon and the rectum, and Crohn disease can reach any segment of the GI from the mouth to the anus [2]. Actually in Portugal, more than 14.000 people have these diseases and in Europe this value reaches to one million being that the general tendency is that these diseases increase every year [3]. Until then the diagnosis of these diseases was generally performed by colonoscopy accompanied by biopsy in case of lesions suspicion. Nowadays, the endomicroscopy allows the early detection and improve the treatment of these diseases.



## CHAPTER 1: INTRODUCTION

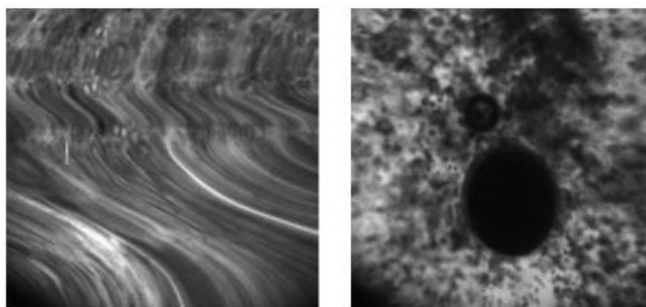


**Figure 2:** Horizontal cut of the bowel.

To detect the GI-tract alterations it is essential to know the histology of this tissue. Figure 2 shows a histological image of the bowel acquired through horizontal cut. The apparent circles are crypts with the lumen at the center. Around the lumen there are goblet cells and epithelia cells. The space between the crypts is the lamina propria.

In the endomicroscopy we can categorize the obtained images in four types: image of healthy tissue, tissue with inflammation, tissue with neoplasia and image artifacts. A thorough knowledge of the normal and abnormal microscopic architecture of gastrointestinal tract is necessary to do the interpretation of the confocal images.

A good image interpretation and the correct diagnosis depend on the quality of the image itself. Consequently it is essential to have the best image quality at acquisition time. For this to happen, the image artifacts need to be minimized. There are two general types of artifacts, which are motion and image artifacts.



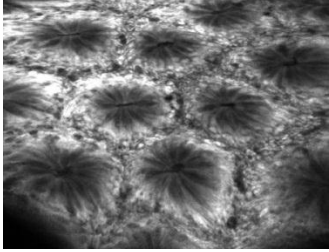
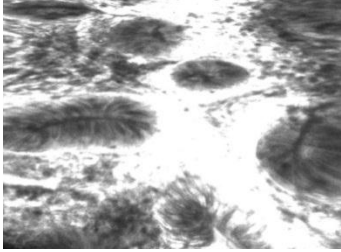
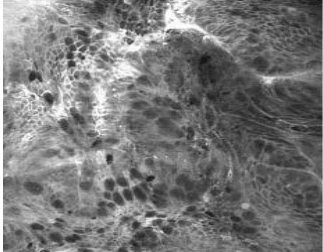
**Figure 3:** Motion artifacts (left) and artifacts due to air bubbles and fecal content (right).

Thus, it is of great importance keeping the endoscope in a stable position. This is mandatory to obtain a good quality of image; washing the mucosal surface removing the waste of the faces; minimizing patient movement.

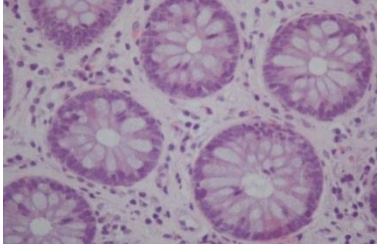
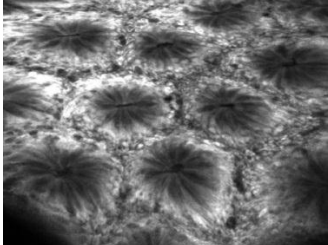
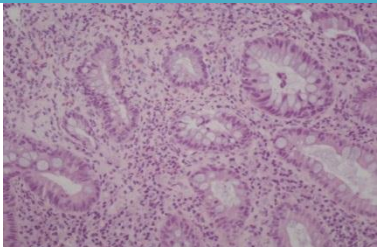
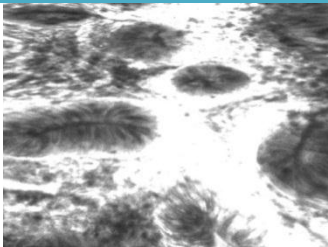
## CHAPTER 1: INTRODUCTION

Beyond the image with artifacts previously presented in the endomicroscopy exam we can obtain three more types of images as mentioned previously: image of healthy tissue, image of tissue with inflammation and image of tissue with neoplasia.

**Table 1:** Types of images without artifacts acquired by endomicroscopy.

Image of healthy tissue	Image of tissue with inflammation	Image of tissue with neoplasia
		
<b>Features</b>		
<ul style="list-style-type: none"> <li>• Many crypts surrounded by goblet cells.</li> <li>• The crypts openness appears like black holes with circular or slit form.</li> </ul>	<ul style="list-style-type: none"> <li>• Generally has few crypts with irregular structure, but it cannot have crypts.</li> <li>• The lamina propria is increased and bright.</li> </ul>	<ul style="list-style-type: none"> <li>• Have crypts with Ridged-lined structure or don't have crypts .</li> <li>• Have isolated black cells.</li> </ul>

**Table 2:** Comparison between endmicroscopic image and image acquired from biopsy analysis.

Image obtained from biopsy	Endomicroscopic image
	
	

## CHAPTER 1: INTRODUCTION

The left image of table 2 is an image obtained from the biopsy analysis and the image on the right was obtained by endomicroscopy. As it can be seen, the endomicroscopic image presents good results compared with the biopsy image. They clearly present all structures with well-defined forms.

The execution time of endomicroscopy is greater than in other conventional endoscopy, which brings slight discomfort to the doctor. The simultaneous execution of endoscopy and a microscopic histology exam is very difficult to accomplish in practice. The doctors are required to go through a very long training period to be able to do it. This technique is highly operator-dependent in the sense that the outcome is very influenced by the skills and experience of the endoscopist. It is only possible to make endomicroscopic diagnosis with a basic knowledge of normal and pathologic histology because the physician must distinguish a normal tissue from the abnormal tissue to make correct diagnosis. So the doctors who execute CEM require a fundamental knowledge of the normal and disease microscopic architecture of the gastrointestinal tract.

### **1.3: Objectives of this thesis**

The simultaneous execution of a regular endoscopy and a microscopic histological exam is very difficult because this technology is new and the taxonomy of the image diagnosis is yet to be defined. Computer assisted diagnosis can be beneficial so the main goal of this thesis is to develop tools to assist the doctor during the exam. This research project has three purposes in the long run:

(i) Procedures and software for collecting and labeling CEM images aiming to obtain a large scale database. The usefulness of such a database is twofold: obtain data for developing a classifier (learning by examples / data-driven methods), and provide medical education (namely by image content retrieval, this is a very new technique and the image taxonomy and interpretation is not well defined).

(ii) To develop basic algorithms for detection and segmentation of the cellular structures. The choice and extraction of image features to develop suitable global image descriptors that encode local image properties such as color, shape, and textures have to be addressed.

## CHAPTER 1: INTRODUCTION

(iii) To develop techniques that allows the automatic classification of the different types of image.

### **1.4: Related work**

Despite the endomicroscopy being a novel technique and its application in clinical diagnosis recent, it appealed the scientific community very quickly due its many advantages. Several studies have been performed in order to assist the training of physicians because the diagnosis does not have well-defined criteria. There are medical works aiming building a data base with annotations obtained from the correlation with the results of the biopsy to standardize the taxonomy of the images.

Johannes Gutenberg, University of Mainz Germany, developed the first endomicroscopy platform available on-line that teaches to identify the endomicroscopic image and distinguishes the various mucosal abnormalities through studies of several real clinical cases [4]. In [5], [6] Buchner et al. realized surveys about colorectal cancer presenting details about the acquisition protocol of the probe-based confocal laser endomicroscopy (pCLE) database and demonstrating that the confocal analysis of the crypts and vessels allow to distinguish neoplastic and non-neoplastic lesions. They also proved that CEM has a high diagnostic accuracy, reducing unnecessary treatments. The interpretation of images acquired by endomicroscopy can be quickly learned through an appropriate training and with the standardization of confocal images.

In [7], [8], [9] André et al. developed work that helps in image classification with the aim of making the content-based image retrieval. They used images obtained by a device different from ours with a smaller field of view (FOV), value 240  $\mu\text{m}$ . To solve the problem of small FOV they propose the utilization of mosaic images. These studies demonstrate that those endomicroscopic images have great intra-classes variability and small differences between classes. To improve these differences the physical process of acquisition was introduced and the Bag-of-Visual-Words method was utilized. Bag-of-Visual-Words was adjusted for endomicroscopic images, in other words, they used dense detector with multi-scale to obtain the local and global information densely distributed in image instead of sparse detectors to extract salient regions in the image. They concluded that the spatial organization of cell is discriminant. So they used the co-occurrence matrix to extract geometric measures.

## CHAPTER 1: INTRODUCTION

These works, although inspiring, concern small FOV CEM. We believe that no relevant literature has been provided in wide FOV CEM.

### **1.5: Overview of the thesis**

The present investigation project is organized in six chapters. The first chapters comprises the present chapter, where we present a description of confocal endomicroscopy technique, its extension for in vivo examination and understanding an image of CEM in Gastroenterology, the aims of this thesis and its structure.

In the next chapter the principles of confocal microscopy is explained in detail. As technologies incorporated in CEM are based in confocal microscopy, in this chapter we address the following issues: (i) knowing how the endomicroscopy system works; (ii) understanding the reasons for the endomicroscopy image has high quality and resolution; (iii) describing the feature of the device used and the criteria adopted in data acquisitions.

The objectives of the database will be presented in the third chapter. The design and technical implementation of the database is also shown. It will be demonstrated how the system is working now and its future developments.

The fourth chapter contain the statistics of the database analysis all the parameters that characterize the endomicroscopic image and image texture analysis, and the results discussion.

The research in pattern recognition techniques is done in the fifth chapter, where the technique to segment structure is presented.

Finally, in chapter six the results are presented and the conclusion of this investigation project.

## Chapter 2

### The system of Confocal Endomicroscopy

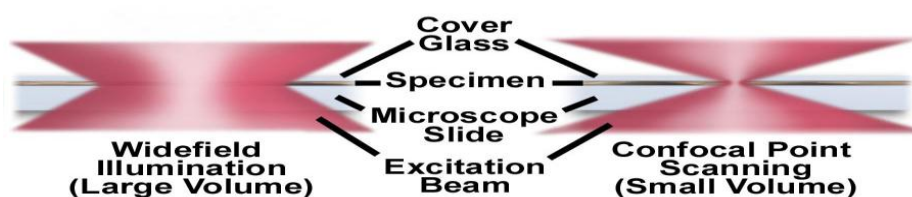
In this chapter we present the principle of confocal microscopy to help us understanding the technique of the confocal endomicroscopy system and with this enable to know (i) the main advantages of this technique; (ii) The reason of its high resolution and good image quality. To familiarize the read with the endomicroscopy equipment in the second and third section, the endomicroscopy system is presented in detail and the main characteristic of the hardware and software are described. We also explained how to export the data and manipulate third party software data and conventional platform.

#### 2.1: Basics of confocal microscopy

The principle of confocal image was patented by Marvin Minsky in 1957. The confocal microscopy aims to overcome some traditional wide-field fluorescence microscopes limitation.

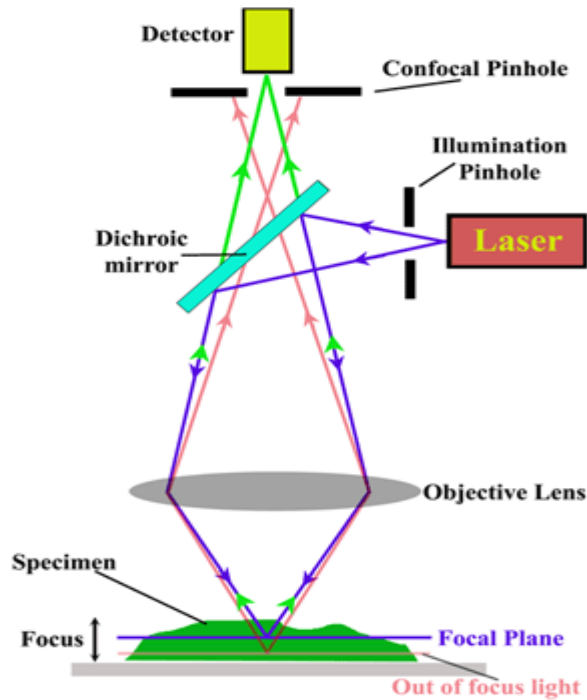
In a conventional fluorescence microscope all the specimen is lighted so the secondary fluorescence emitted by the sample, that is out-of-focus, contributes to the image formation, so this image has low resolution and quality. In contrast, the confocal microscopy uses focal laser illumination to scan one focal plane and confocal pinholes [10]. This system enables only the light emitted in focal plane reaches the detector and building the image.

Figure 4 shows the difference in illumination of the sample in confocal microscopy and traditional microscopy.



**Figure 4:** Wide-field vs. confocal microscopy, illustrating the principal difference in volumes illumination.

## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY



**Figure 5:** Basic principle of confocal microscopy

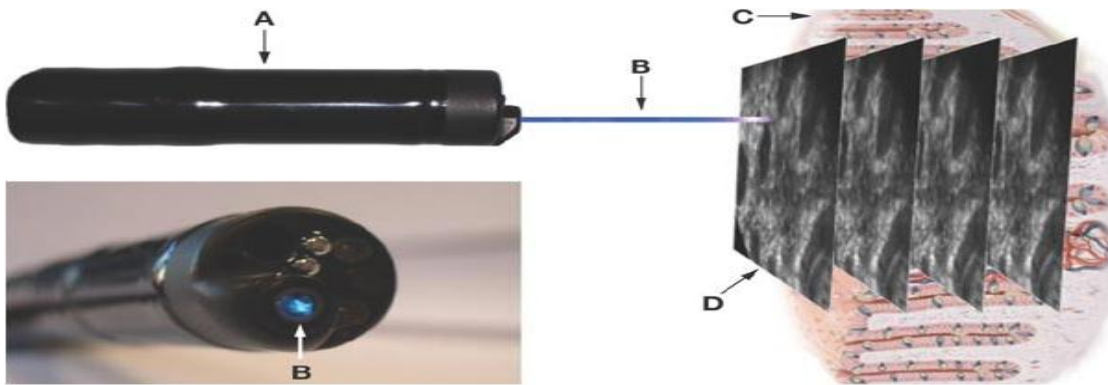
The principle of confocal microscopes is shown in Figure 5 and it consists of two steps. (i) The excitation of the sample that is done from coherent light sent by laser. The light emitted pass through the illumination pinhole, then it is reflected by dichromatic mirror and finally reaches the focal plane. (ii) The light emitted from the specimen uses the same focal plane to attain the detector. The out-of-focus light does not reach the detector because it is blocked by the confocal pinhole positioned in front of the detector [11]. As only the light emitted by the specimen in focal plane contributes to the image formation, the resolution is better than in a conventional microscope [12]. Nevertheless, the increase the resolution comes at the cost of a great reduction of light.

The increase of light intensity can be done in two ways. One way is to increase the intensity of the laser light, however high intensity of the laser causes other severe consequences like photo toxicity in the specimen. The other way is to increase exposure time, but this procedure makes the system incompatible with life cells observation. The other disadvantage of long exposure time is that it increases the possibility to occur the motion artifact. Because of these limitations generally confocal laser scanning systems are utilized on fixed specimens.

## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY

To overcome these limitations the spinning disk confocal systems was implemented. These systems consist of inserting the disk into the light path between the dichromatic mirror and lenses set. The spinning disk systems allow greater light transmission and a higher speed [13]. So this system enables to observe dynamic processes. However, this system has less resolution than laser scanning confocal systems.

In confocal laser endomicroscopy all the light coming from the focal point is detected and measured. This light is used to build the gray-scale image that represents an optical section of the examined specimen. Figure 6 illustrates the scheme of confocal laser endomicroscopy.



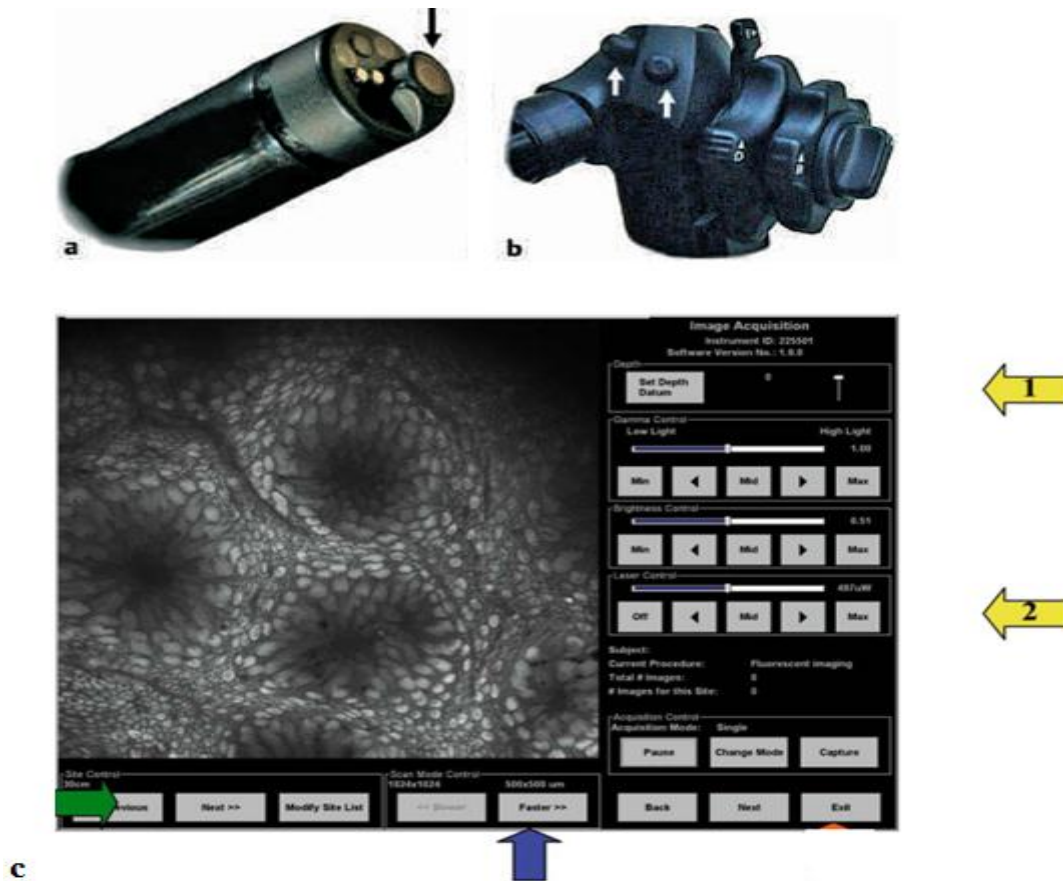
**Figure 6:** Scheme of confocal laser endomicroscopy. The confocal laser microscope is integrated into the distal end of a conventional endoscope *A*. The blue laser light *B* crosses the mucosal layer at the surface of the intestine till reaching the deep tissue, *C*. *D* is the stack of CEM images that capture the reflection of the laser light in a contrast substance that has been previously injected. Each image of the stack corresponds to an optical horizontal section of the mucosal layer.

### 2.2: Confocal Endomicroscopy System Used

Pentax, in cooperation with Australia's Optiscan, developed a confocal endomicroscopy System, ISC-1000. This system integrates the optical system for confocal observations in a normal endoscope. The confocal microscopy laser is incorporated into the distal end of a conventional endoscope [14]. As the endoscope and confocal microscopy laser is joined into the device, we can make the endomicroscopy exam and endoscopy exam with same device and at the same time.



## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY



**Figure 7:** (a): The confocal laser endoscope with confocal imaging window signaled (b): Two additional buttons used to control the endoscope during the image acquisition. (c): Image acquisition screen.

The endomicroscopy devices use two monitors, one to display the endomicroscopy video and other for image acquisition. The monitors operate simultaneously. Besides these components, there is a video processor and a confocal control unit. The endoscope contains an air and water jet nozzles, two light guides, an auxiliary water jet channel used to inject the contrast agent, an instrument channel that allowing guided biopsies and a confocal imaging window. Figure 7 (a) illustrates these components. There are two additional buttons located on the back of the endoscope (illustrated with white arrow in figure 7 (b) one of these allows to change the focal plane position, moving it to home position located approximately 10  $\mu\text{m}$  deep from the surface, and the other allows to change the scanning direction (up or down) depending on the acquisition direction [15].

Before the procedure we can manage some options regarding the acquisition of the image. We can adjust some features like gamma, brightness, scan mode and laser

## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY

control through the buttons positioned on the right side and at the bottom of the figure 7(c).

The field marked with the yellow arrow 1, figure 7(c) indicates the depth of the focal plane and the scanning direction. The yellow arrow 2 shows the laser control function that, as the name suggest, it allows controlling the laser power and changes the brightness of the image during the acquisition process. The blue arrow signals the scan mode control through which we can modify the resolution of the image. There are also two options, “faster” and “slower”, to increase or decrease the speed of the scan. The green arrow marks the buttons that enable to choose the site where the acquired image is saved.

**Table 3:** Main specifications of the Pentax CEM

Features	Values
Excitation Wavelength	488 nm
Range of depth	0–250 $\mu\text{m}$
Resolution	1024 x 512 pixels or 1024 x 1024 pixels
Range of laser power	0 to 1000 $\mu\text{W}$
Field of view	475 x 475 $\mu\text{m}$
frame rate	1.2 images/s at a resolution of 1024 x 512 pixels or 0.7 images/s at a resolution of 1024 x 1024 pixels

CEM is an image technique that uses the focal illuminations to scan one image. CEM allows observing only the area of interest since surface to a depth of 250  $\mu\text{m}$ . The current imaging depth is restricted, which limits the submucosal layer visualization.

To obtain an endomicroscopic image it is necessary to use a contrast agent injected intravenously. The potential agents of contrast are fluorescein, acriflavine and tetracycline. Typically acriflavine hydrochloride (0.05% in saline) or fluorescein sodium (5 $\pm$  10 ml of a 10% solution) are used. Fluorescein is the most used because it is rapidly metabolized by the liver and eliminated by the kidneys. The acriflavine allows the visualization of the nucleus and this enable an early detention of the cancer [16]. Despite this benefit acriflavine is rarely used due the high risk to cause mutations.

The endomicroscopy device relies on the *Clinical Software* for imaging, having two operations: acquisition and the review mode. The first application allows introducing the following data: (i) information about patients and the doctor who makes the exam; (ii) A site description of where each set of the images was acquired to allow future identification; (iii) Observations made by the doctor during the procedure.

## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY

Furthermore, it can obtain images, review them during the ongoing procedure and store them. The review mode allows: (i) an image review from any patient; (ii) importing and exporting the stored procedures to other equipment carrying only by the patient's ID as identifying data. This application does not need to have a confocal endoscope connected or operating.

In each area of interest, from the surface to deeper parts of mucosal layer, it is possible to acquire images which are digitally saved in a specific file. The images can be saved in four different file formats: JPEG, TIFF, Optiscan and windows 24 bit bitmap. Moreover, during the procedure we can review the images immediately or later in order to compare the exam and to see the patient evolution.

The Software also enables transferring the information introduced and the image to other equipment. The exported procedure results are saved as compressed ZIP file that has, as default name, the ID automatically created by the device. The ZIP file has a XML document and a file with the image in a proprietary format (.bin). The XML file contains the patient information, doctor identification, procedure description, number of the image acquired and the local acquisition. The BIN file includes the images and its technical features like the laser power, depth and resolution. The images in .bin format cannot be opened directly, in other words, they cannot be viewed without prior processing. We need to use some tools like matlab and java to use the data. Another way to open these images is through the *ImageJ* software with the specific Plugins developed by Optiscan.

### **2.3: Collection of data**

This project consists of studying the patient with inflammatory bowel disease. Patients were examined in the Gastroenterology Service at the Hospital of the University of Coimbra. This exam was performed using the confocal endomicroscopy System, ISC-1000. The agents of contrast used were fluorescein.

To acquire images with high quality it is essential to have the bowel prepared and the mucosa clean. These conditions avoid the faeces contribution in the image formation and consequently the changes in the diagnosis. Therefore, days before the exam, the patients prepare at home with a diet without waste and four liters of Klean

## CHAPTER 2: THE SYSTEM OF CONFOCAL ENDOMICROSCOPY

Prep (Polietilenoglicol). Usually this preparation is not enough so the mucosa is cleaned by a water jet during the exam.

The entire exams were performed by two experts in CEM. To introduce the images in the database it is necessary to obtain high quality images and export them out of the clinical software. To do so, it is essential to adjust some parameters as already we explained in previous section.

After the exam conclusion the images are saved in the file to export. The data exportation must be done through the specific Review Mode because we do not want to export only the images but as much information as possible. The exam exported in this way is saved as compressed ZIP file. This ZIP has a XML document and a file with the image as previously explained. After extracting the data we will introduce and annotate the image in the database through the application developed throughout this project.

## Chapter 3

### Building the database

One goal of this project is developed software for collecting and labeling CEM images. This chapter introduces the support infrastructure to create a database for collecting and labeling the image obtained from confocal endomicroscopy. The aim of the database is presented. The second section tackles the design of the database and it shows the physical and relational diagram. In the third section we present the screen interface with user and the parameters used to characterize the images are explained in detail. The fourth section presents the improvements that can be made on the future application.

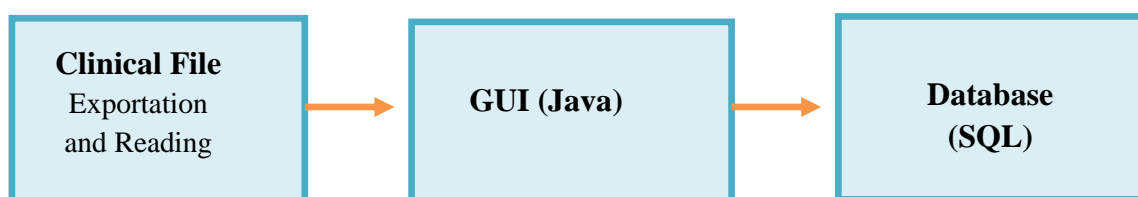
The developed application enables the doctors to autonomously label the images. This application is practical and easy to use, minimizing the change in the diary medical routine.

#### 3.1 Objectives of the database

The application developed to build the database is a program with a simple use that allows the doctors to store the process, exam and images of patients. The user can use this application to classify and to annotate each image in order to help in future training.

The application has many screens that allow information access and modification. As mentioned previously, the database has two purposes: (i) obtain data for developing a classifier; (ii) provide medical education.

Figure 8 shows a schematic of the data extraction into the database.



**Figure 8:** Schematic representation of the plan to build the data base.

## CHAPTER 3: BUILDING THE DATABASE

According to the problem specification we defined the following feature/requirements for the application:

1. **Two types of users:** Normal and Administrator;
2. **Log on:** Through the username and password;
3. **To make a new registration:** Only the Administrator can do this.
4. **To search Processes, Exams, Images and Annotations.**
5. **To introduce new Processes, Exams, Images and Annotations:** the user must fill all fields that characterize each entity.
6. **To delete and modify existing data.**

Before to create the application we defined the possible user types in order to ensure accesses security:

- **Normal Users:** they are already registered and they can do the login. They have access to all data but they can only edit and delete their own processes.
- **Administrators:** They are the only ones who can record the other users. They have access to all data and they can modify everything.

### 3.2 Design of the database

To fulfill the application requirements previously presented, the database was designed as follows: an entity “*Medico*” that refers the registered user, contains the doctor identification attribute. All the doctor relates to one or more patients and one patient can only has one physician responsible. So we created a relation 1xN between the entities “*Medico*” and “*Processo*”. The entity “*Processo*” has the following attribute: the patient’s name, the doctor who makes the exam, the examination tasks date and ID process. The ID process will be the primary key since it characterizes univocally this entity.

The entire process has one or many exams so a relation “*possui*” was established between the entities “*Processo*” and “*Exame*”. The entity “*Exame*” represents a summary of the exam and it has the following attributes: the examination date, the

## CHAPTER 3: BUILDING THE DATABASE

number of process that each patients has in the database, the diagnosis and observations done. The primary key of this entity is ID exam and the foreign key is ID process.

Each exam contains a set of images but an image can only belong to one process. Hence it is necessary to create a relation 1xN and an “Imagem” entity. This entity contains technical acquisition features and indicators of each image type. It has as the primary key the, ID image and the foreign key the ID exam.

The doctor has the possibility to make annotations in images. So more one entity “Anotacoes” was created that is related with “Imagem” entity. This relation is 1xN because an annotation can only belongs to one image but an image can have many annotations. This relation is not mandatory since an image cannot have annotations.

### Relational –Diagram

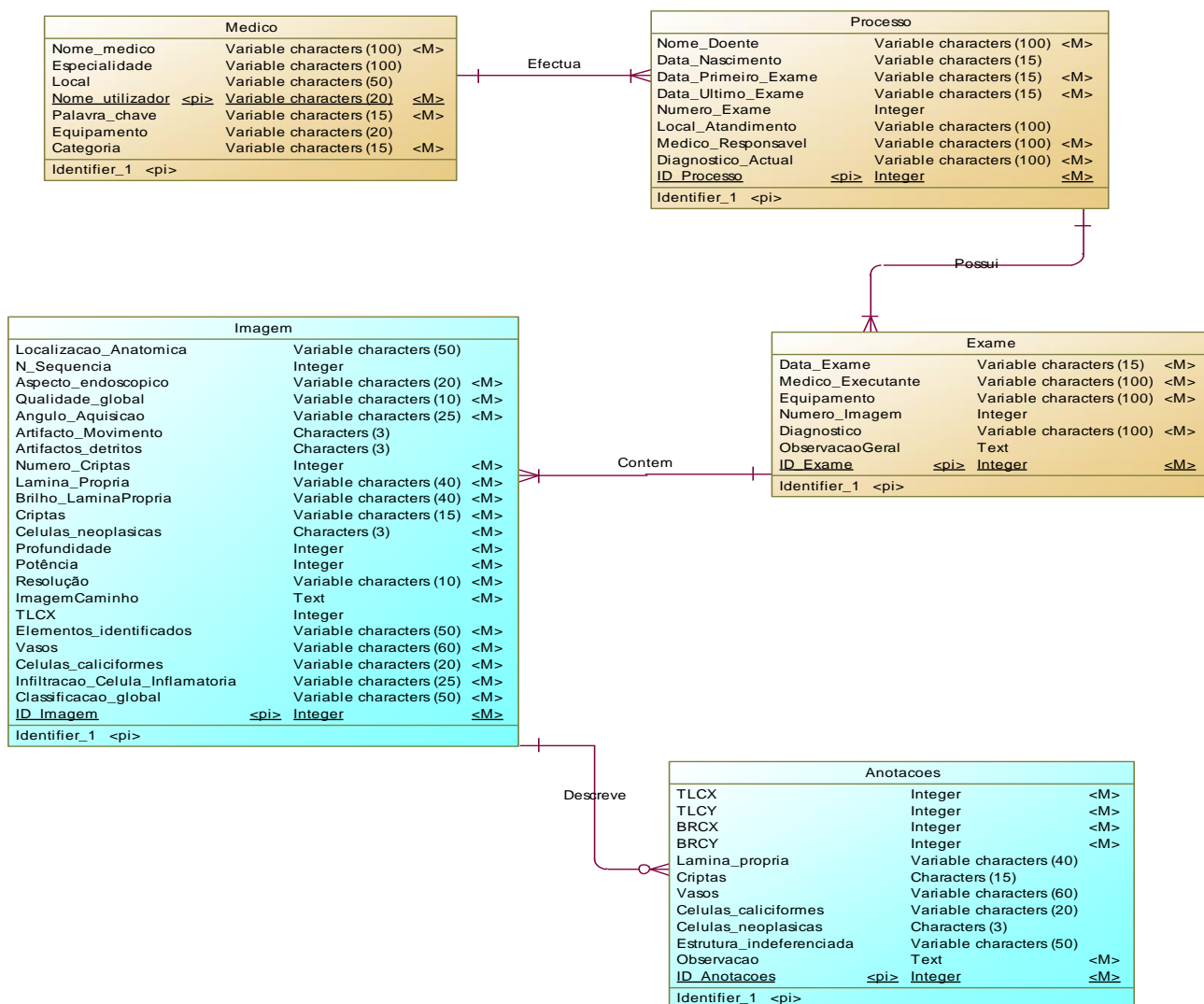


Figure 9: Relational Diagram of the database.

## CHAPTER 3: BUILDING THE DATABASE

### Physical -Diagram

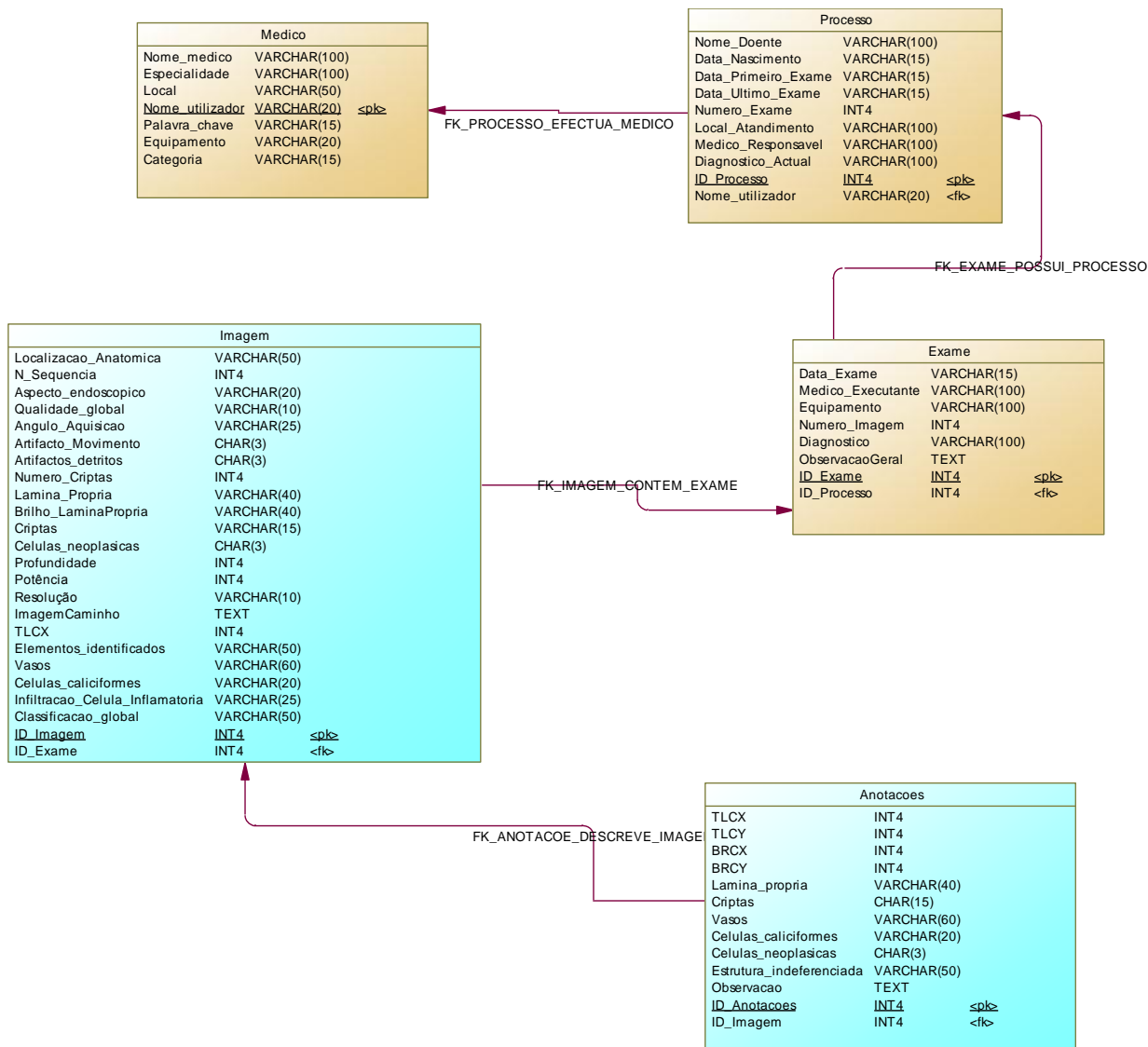


Figure 10: Physical Diagram of the database.

### 3.3 The tagging application

To build the database we used the PowerDesigner 15 from Sybase® to create the relational and physical diagram and PostgreSQL 8.4® to generate the entity and its attributes. In addition of these softwares we used the NetBeans IDE 6.9.1® to create the application in Java. The developed application consist essentially a set of graphical user interface (GUIs).



## CHAPTER 3: BUILDING THE DATABASE

The code was divided into three sets of class. One class allows accessing to the database to manipulate the data. In other words, it enables to introduce, delete and modify the information. There is another class that possesses all the attributes of each table and finally the class of the screens that control all the interface buttons.

The developed application allows: (i) the user to make a search, to update and to introduce the data in an efficient manner; (ii) ensuring the mandatory fields are filled.

The general purpose of the insertion process is to introduce the images obtained by examination. These images belong to a specific exam and process. The main aim of our application is inserting and to annotating the images, so we will give more attention to the parameters that characterize the images.

To extract technical acquisition features of the image we do the inverse engineering. The image is read directly of the BIN file. This file has a header that containing technical information of acquisition like: resolution, laser power and depth. The remaining information corresponds to the image data and the bit of initiation and termination. The image data is saved in a columnwise alignment.

To insert image information in the database we read the header of the BIN file to obtain the technical features. To extract the image we read another part of the file byte to byte. As the images are big (1024x1024 or 1024x512), and the database can have several hundreds of images, these are saved in a remote hand drive and only the paths are introduced in the database.

To insert an image in to the database it is necessary to introduce other parameters such as the attributes that characterize an endomicroscopic image. These parameters can allow the doctor to distinguish the different image types and make a better diagnosis. Therefore, these attributes were analyzed by a physician and chosen in order to obtain the maxima quantitative information, because the personal observations do not enable to do the statistical analysis of data. Table 4 shows these parameters in details.

## CHAPTER 3: BUILDING THE DATABASE

**Table 4:** Parameters that characterize the image in the database.

<b>Parameters</b>	<b>Options</b>
<b>Anatomic localization</b>	Colon Rectum
<b>Endoscopic aspect</b>	Normal Inflammation Circumscribed lesion
<b>Angle of acquisition</b>	Front Lightly oblique Very oblique
<b>Global quality</b>	Good Reasonable Bad
<b>Identified elements</b>	Global structure and individual cells Global structure without individual cells No global structure or individual cells
<b>Crypts shape</b>	Regular Irregular Absents Ridged-lined
<b>Number of the crypts</b>	Indicates the crypts numbers that exist in the image.
<b>Vases</b>	Regular Lightly dilated Markedly dilated
<b>Goblet cells</b>	Normal number Depletion Absent
<b>Lamina propria shape</b>	Normal Lightly enlarged Markedly enlarged
<b>Brightness of the lamina propria</b>	Normal Lightly intense Intense Very intense
<b>Inflammatory cells Infiltration</b>	Impossible evaluate Absent Less than 50% superior to 50%
<b>Motion artifact</b>	Yes or not.
<b>Detritus artifact</b>	Yes or not.
<b>Neoplastic cells</b>	Yes or not.
<b>Global image classification</b>	Normal, light inflammation Severe inflammation Light probability of the neoplasia High probability of the neoplasia Blurred image

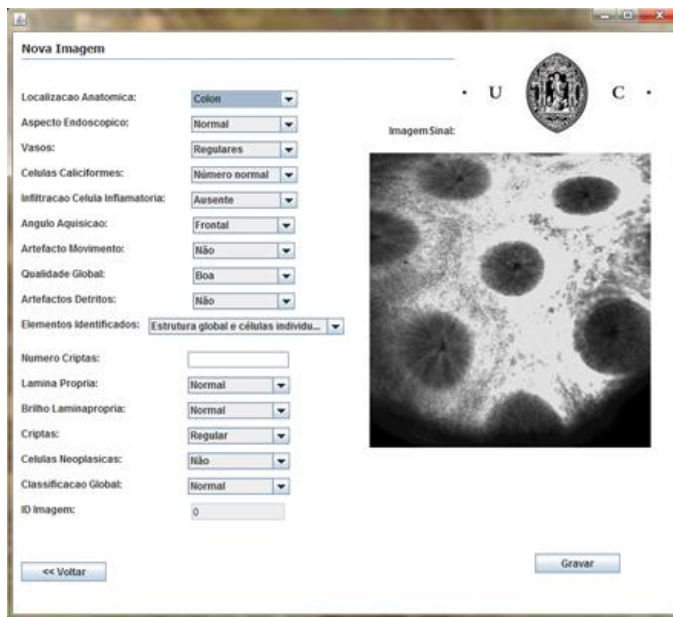
Beyond the parameter introduced by doctor there are the technical acquisition features that were read directly by the BIN file namely the laser power, resolution and depth. They are added automatically.

## CHAPTER 3: BUILDING THE DATABASE

There are two databases, one local database which is in the HUC and other remote database located in the ISR. The synchronization of these databases is made automatically through the primary keys comparison of each table. As we said previously we introduce the images path in database, then the images are transferred via SFTP. To simplify the synchronization we use the relative path. So the image file must be saved in the same local with the executable of the application.

### 3.3.1 Main graphical screens interface

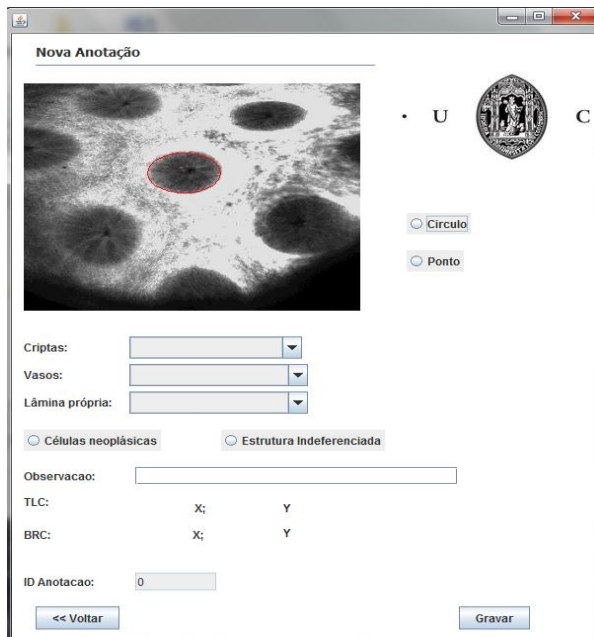
We give more emphasis in storage and annotation of the images. Figure 11 represents the window of the application developed that allows introducing the images on the database. To insert the image in database it is necessary to upload the image file of an exam.



**Figure 11:** Schematic representation of the application screen that allows inserting the image in the database – Nova Imagem screen. There are predefined fields already filled.

The doctor can select an area of the image that is relevant and then annotate. There are predefined fields to make note for example the crypts, vases, lamina propria, neoplastic cells.

## CHAPTER 3: BUILDING THE DATABASE



**Figure 12:** Schematic representation of the application screen that allows annotating the images – Anotação screens.

# CHAPTER 3: BUILDING THE DATABASE

## Schematic summary of the application screens.

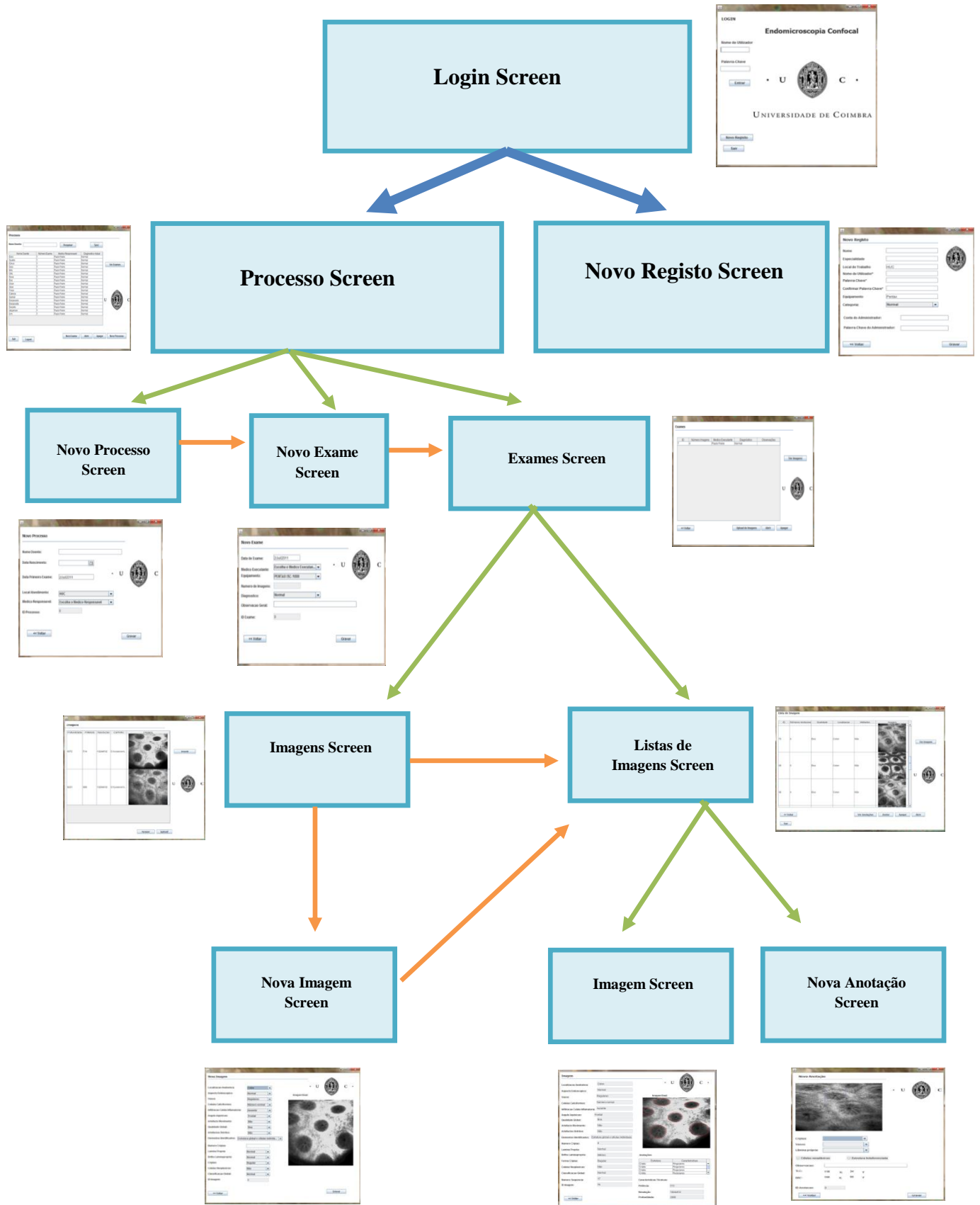


Figure 13: Schematic representation of the application screens.

### **3.4 Future Developments.**

Our program was developed for just one user, there is only a local database and just one doctor can access at the same time. In case there are several experts to do the annotation this application is limited.

The database modifications is made always in one sense, that is, the information is introduced and modified only in the local database and then we make the synchronization to the remote database existing in the ISR.

As future work we intend to modify the database to support the concurrent access of multiple local clients installed in the computers of different medical users, becoming a more flexible and embracing application.

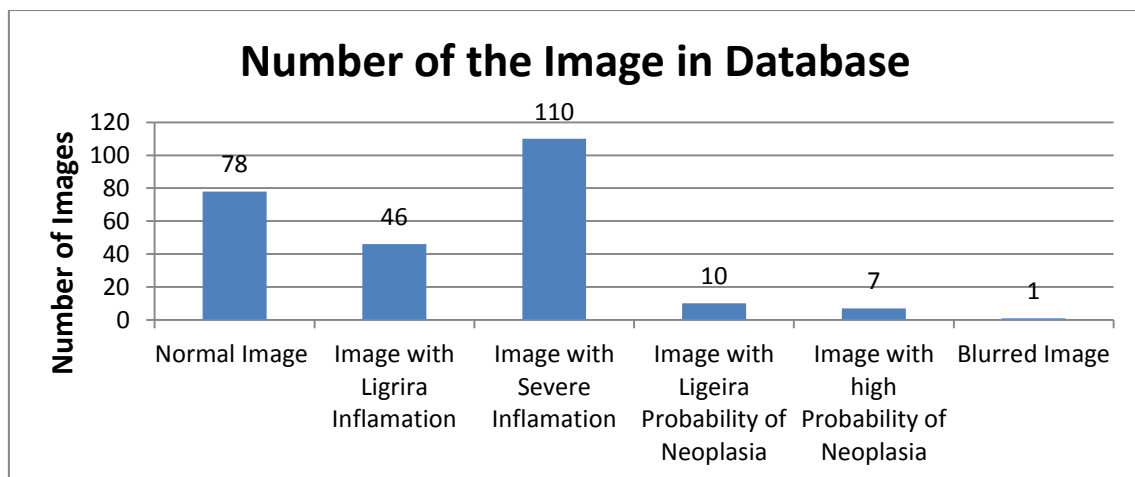
## Chapter 4

### Database Analysis

To know the influence that each parameter has in each image type is fundamental in the distinction of the classes. This information will help in the construction of the descriptors. Therefore in this chapter we analyze the database statistics, present the texture analysis and discuss the results.

#### 4.1 Statistics of the Database

The database has 14 patient processes, 14 exam and 252 images. These images are distributed in six categories: normal image, image with a light inflammation, image with a severe inflammation, image with low neoplasia probability, image with high neoplasia probability and blurred image.

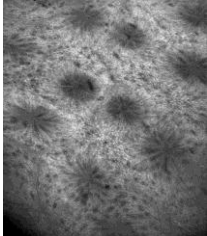
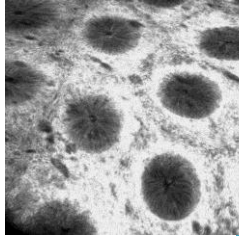
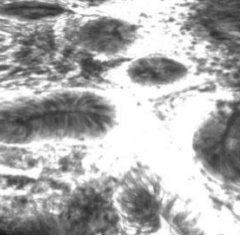
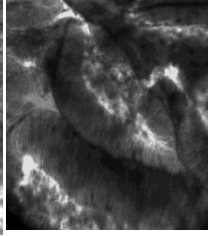
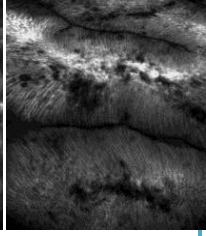
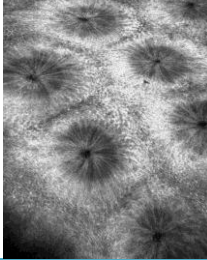
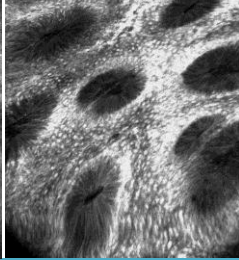
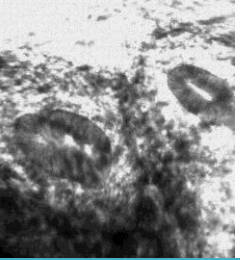
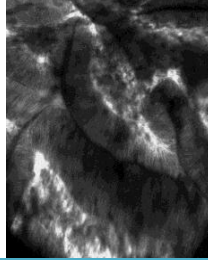
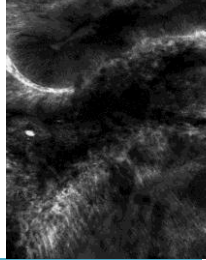
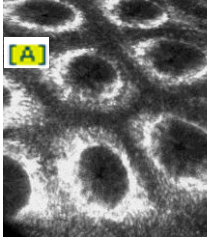
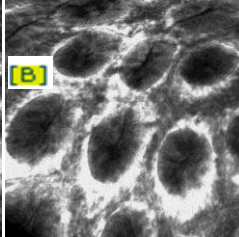

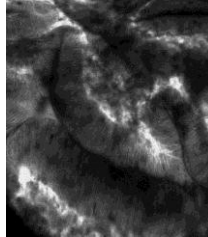
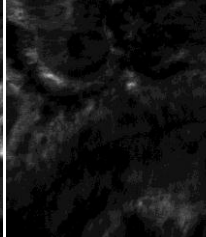
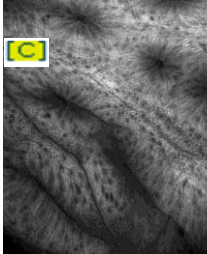
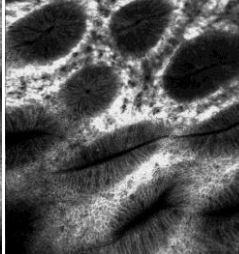
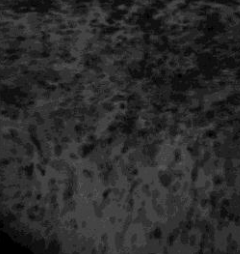
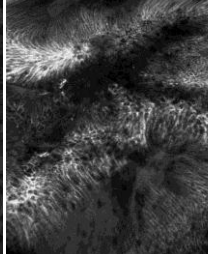
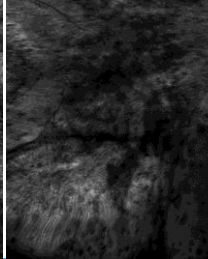


**Graphic 1:** Bar graph of the number of image on database separated by types.

As we mention previously there are three diagnostic kinds obtained by endomicroscopy: normal tissue, tissue with inflammation and tissue with neoplasia. But the images were divided in five classes due the great intra-classes variability and small differences between classes. Table 5 contains some example of each images type existent in the database.

## CHAPTER 4: DATABASE ANALYSIS

**Table 5:** Some examples of different classes of image in the database.

Image of healthy tissue	Image of tissue with Inflammation		Image of tissue with Neoplasia	
Normal Image	Image with light Inflammation	Image with Severe Inflammation	Image with low Neoplasia Probability	Image with high Neoplasia Probability
				
				
				
				

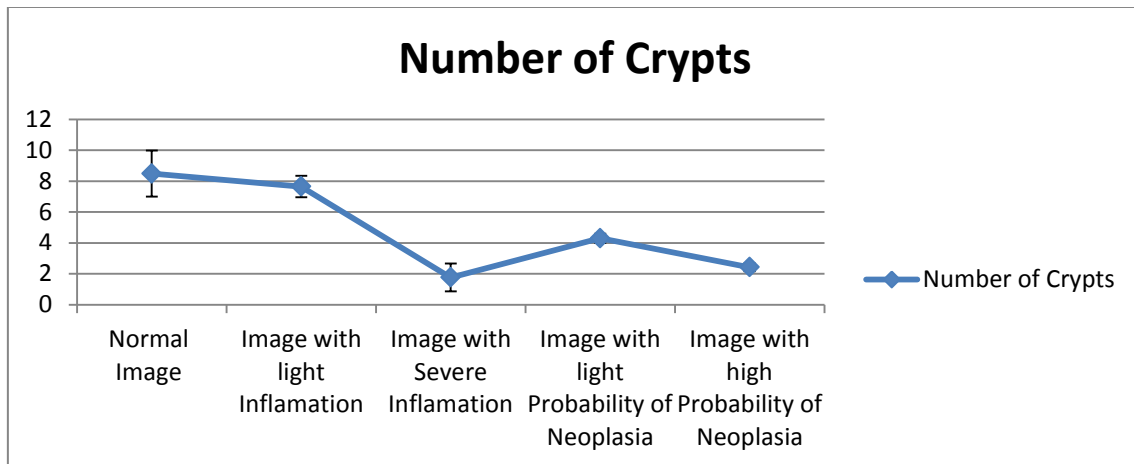
It is very difficult to differentiate the classes due the high similarity between the images. For example the images, marked with the letter *A* and *B* belong to different classes but they are identical. They only differ in lumen being the image *A* has a round lumen and the image *B* lumen is slightly elongate. The image marked with *C* is a normal image with feature similar to the image with light inflammation because the image was acquired in inclined endoscopy position.

Performing the statistical analysis of the database we obtain the results presented in graphic 2 until graphic 7 and figure 14 until 18. All quantitative parameters that



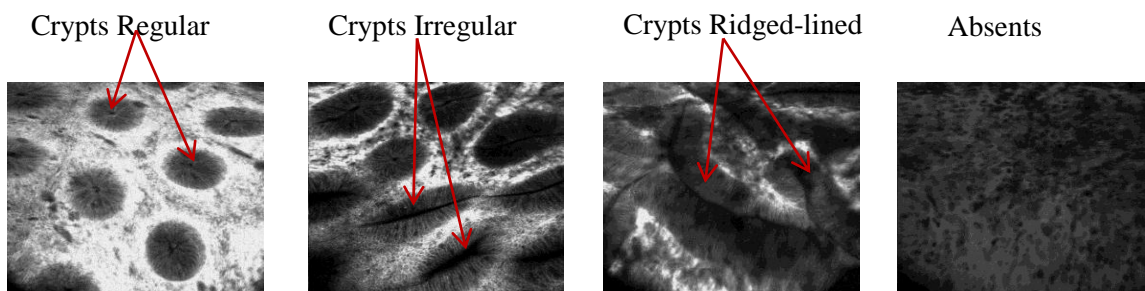
## CHAPTER 4: DATABASE ANALYSIS

characterize the image and distinguish one class from another were evaluated. We made this evaluation in all classes with clinical interest.

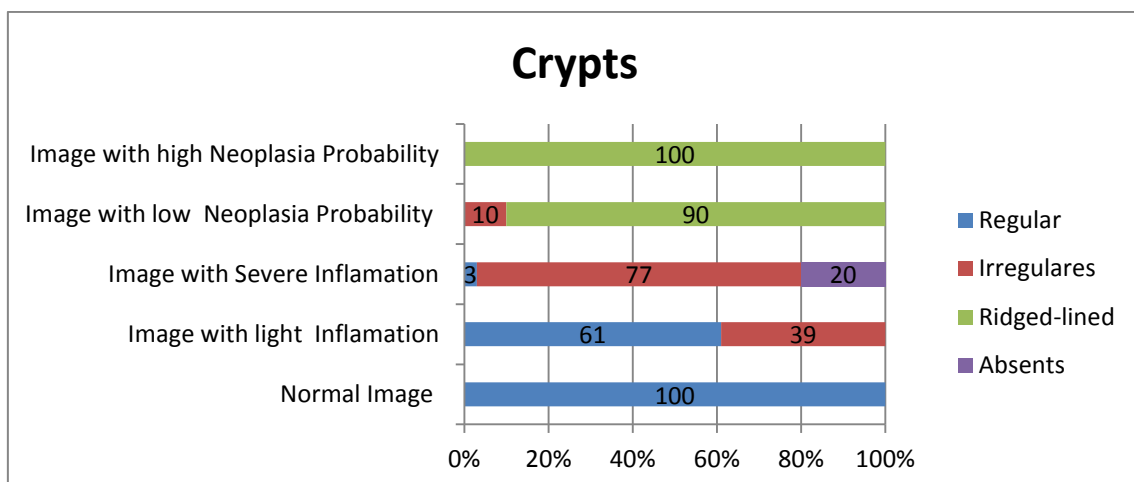


**Graphic 2:** Representation graph of crypts number distribution in different classes of images.

The crypts shapes are another attribute that characterize the image. The following figure has the four crypts shapes.



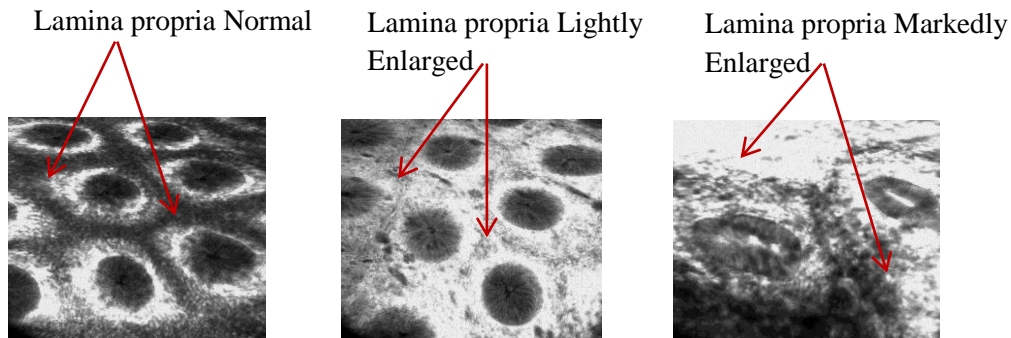
**Figure 14:** Types of crypts shapes.



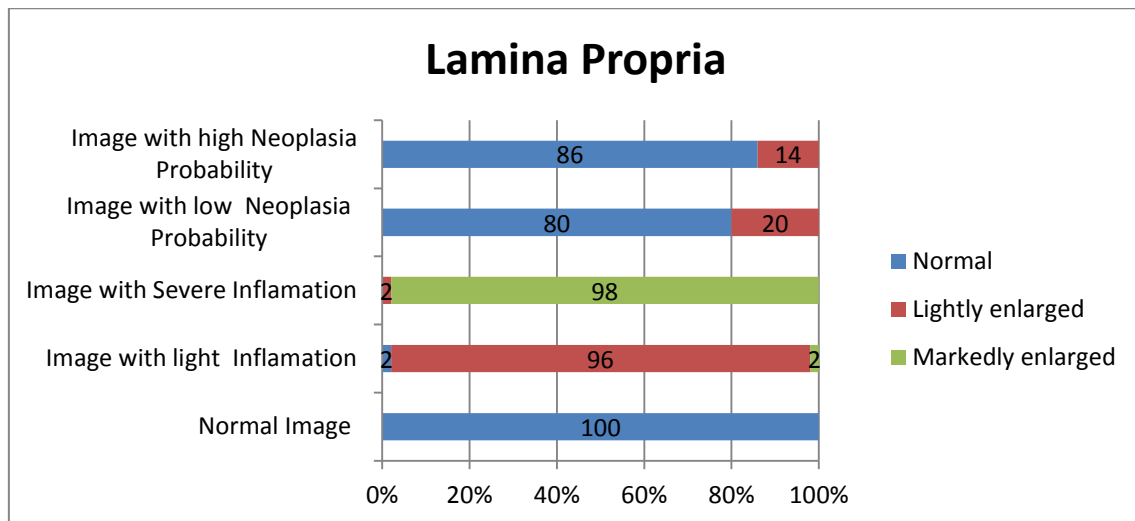
**Graphic 3:** Representation graph of crypts shape distribution in different classes of images.

## CHAPTER 4: DATABASE ANALYSIS

To characterize the lamina propria we use two parameters: shape and brightness. The lamina propria shape can be: normal, lightly enlarged and markedly enlarged.

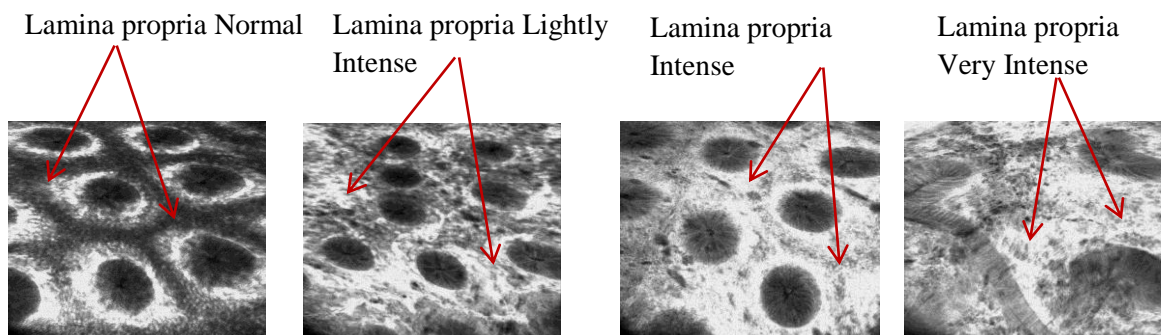


**Figure 15:** Types of lamina propria shapes.

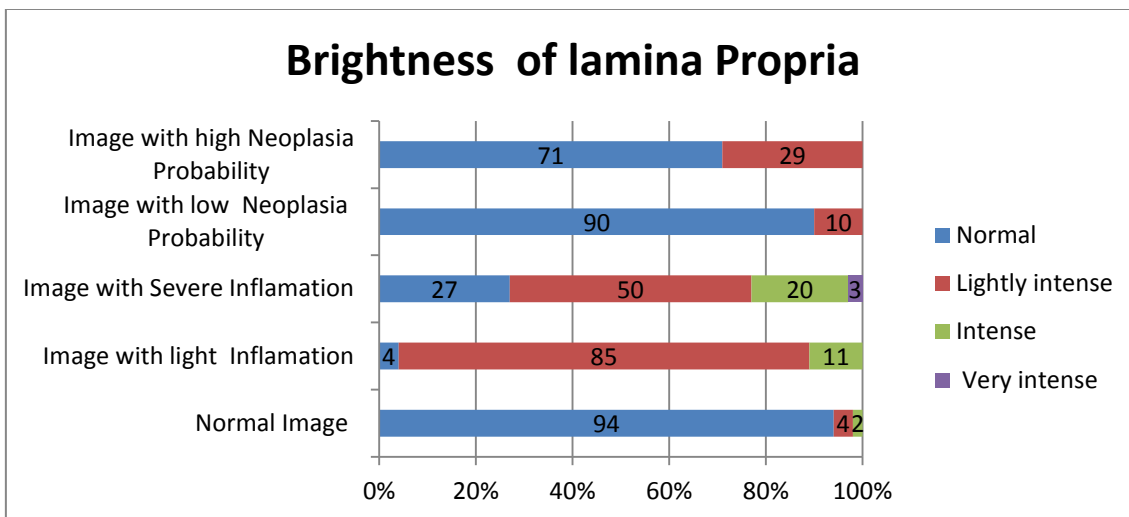


**Graphic 4:** Representation graph of lamina propria shape distribution in different classes of images.

The following figure contains the brightness kinds of the lamina propria.

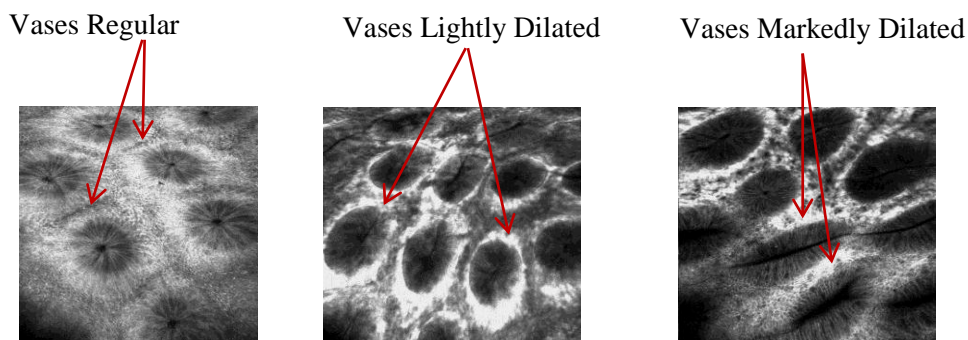


**Figure 16:** Types of lamina propria brightness.

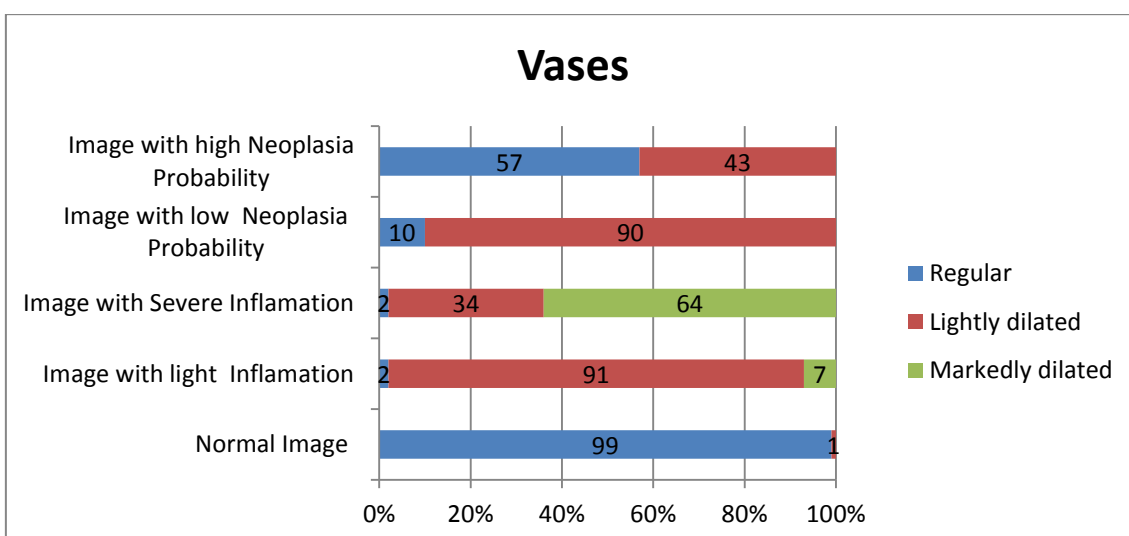


**Graphic 5:** Representation graph of lamina propria brightness distribution in different classes of images.

Other elements that characterize the endoscopic images are the vases, so we evaluate the vases shape.



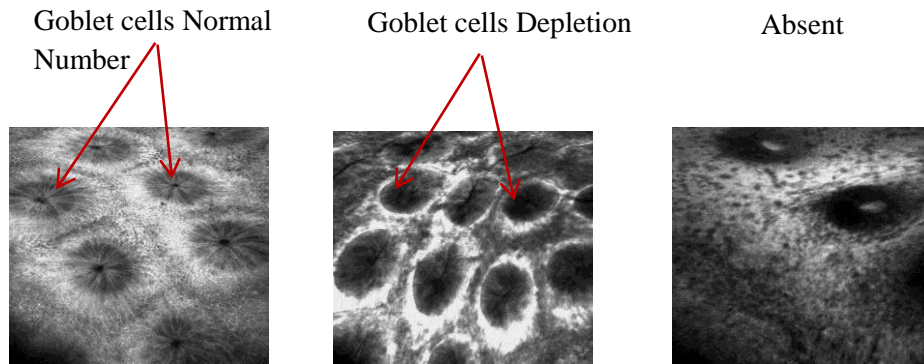
**Figure 17:** Types of vases shape.



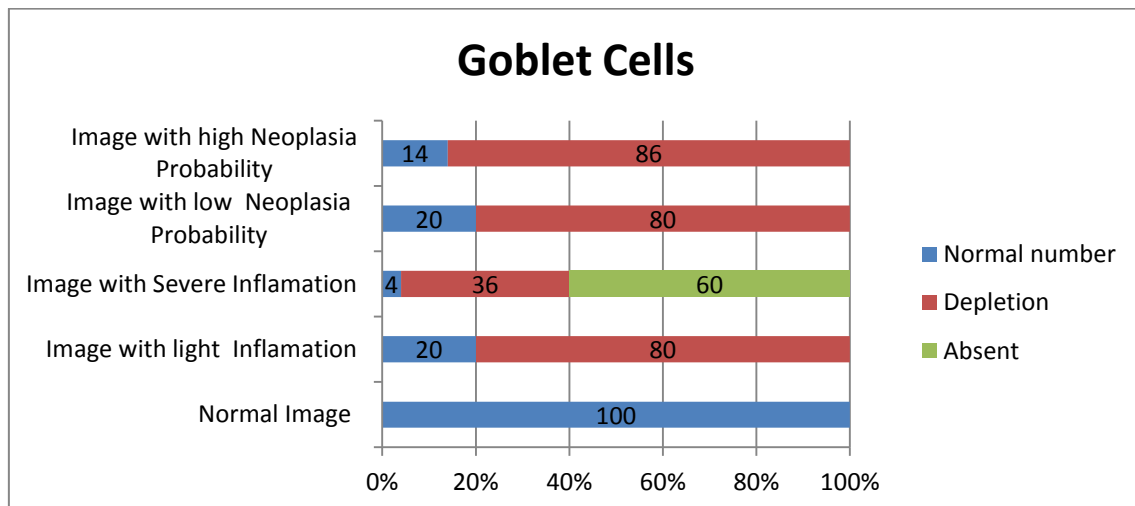
**Graphic 6:** Representation graph of vases types distribution in different classes of images.

## CHAPTER 4: DATABASE ANALYSIS

The crypts of endomicroscopic image are surrounded by goblet cells, for this reason we examine the different kinds of these cells.



**Figure 18:** Types of goblet cells.



**Graphic 7:** Representation graph of goblet cell types distribution in different classes of images.

## 4.2 Discussion

Considering the image features obtained by confocal endomicroscopy we expect that structure like crypts, lamina propria and vases are discriminant in the distinction between the classes.

Analyzing the statistic results of the database there is great variation in the crypts number in images of healthy tissue and in images with severe inflammation. But in others classes this variation is not significant as shown the Graphic 2. The average numbers of the crypts obtained were 8.5 for normal images, 7.7 for images with light

## CHAPTER 4: DATABASE ANALYSIS

inflammation, 1.8 for images with severe inflammation, 4.3 for images with light neoplasia probability and 2.4 to image with high neoplasia probability. Considering the tree diagnosis types obtained by the endomicroscopy this analysis of the results show that there are classes with small difference (mean 8.5 and 7.7) and class where images have great variability in crypts number (mean 7.7 and 1.8). This agrees with the result presented in table 5. Despite this variability, in some class we can consider the crypts number as a discriminant parameter for the characterization of the endomicroscopic image.

For the data provided by the crypts type, the graphic shows that the crypts of the normal image present regular shape and the crypts of the image with high neoplasia probability is ridged-lined. For images with light inflammation, the crypts shapes are irregulars (77%) or there are not crypts (20%). The image with light neoplasia probability, the shape crypts are the following: ridged-lined (90%) or irregulars (10%).

Examining the parameters results that characterize the lamina propria (graphic 4 and 5), we concludes that the lamina propria of the normal image is normal (100%). the image with inflammation has generally a modified lamina propria. In images with severe inflammation, this alteration is lightly enlarged (96%) and in images with severe inflammation it is markedly enlarged (98%). Most images with the neoplasia have normal lamina propria (close to 80%) and the others have lamina propria lightly enlarged. Studying the brightness lamina propria, the results present the same evolution of the shape lamina propria, in other words, the normal image has a normal brightness (94%). In images with neoplasia, the brightness is normal or lightly intense. In images with severe inflammation the only difference is that they can have very intense brightness.

Graph 6 represents image vases and we observe that the normal image has regular vases (99%). The majority percentages of images with light inflammation and light neoplasia probability have vases lightly dilated (close to 90%). The majority percentages of images with severe inflammation have vases markedly dilated. But this class has also a significant number of images with vases lightly dilated (33%).

By the Graph 7 analysis we observed that the normal image has a normal number of goblet cells and the absent feature of the goblet cells appears only in the image with sever inflammation. Beyond this feature, these images have a significant amount of the goblet cells depletion (36%). The other image types generally have goblet

## CHAPTER 4: DATABASE ANALYSIS

cells depletion. But there are a small percentage of images with a normal goblet cells number.

The analyses of the results show that the images with many and regular crypts and normal lamina propria have high probability to be a normal image and an image with many regular crypts but with lamina propria lightly enlarged it has high probability to be an image with light inflammation. There is a high probability of an image belonging to class images with light inflammation where it has a few crypts with irregular shape and lamina propria markedly enlarged. The images with few crypts and ridged-lined shape have high probability to be an image with neoplasia. The images with light and high neoplasia probability are difficult to distinguish because they have similar characteristic.

The lamina propria is very difficult to characterize and differentiate computationally because its shape is difficult to detect and brightness is a difficult metric to measure. The parameter vases and goblet cell together with other parameter like crypts are discriminate, but these structures is very difficult to detect. So the parameter discriminant and easy to compute is the crypts, for this reason we will segment and detect the crypt.

### 4.3 Image Semantic

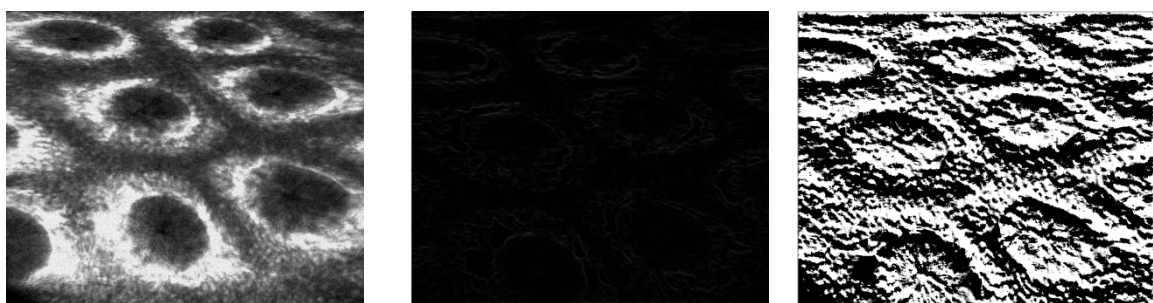
The texture is an attribute that characterize the surface and the image structure. Image texture applies a set of mathematical procedure that allows extracting the texture properties. Image texture encodes local and global image information such as color, shape, spatial arrangement, intensity, etc. We use the texture analysis to help us choosing the better statistical parameter that characterizes the endomicroscopic image.

There are two ways to make a texture analysis: statistical and syntactic approaches. The syntactic approaches work well for artificial textures because they consider the texture image as a set of the regular pattern. The statistical metrics allows evaluating the spatial distribution of intensity in an image region. The statistical metrics is more often used because the natural texture has irregular patterns and it is easy to compute. Techniques like Fourier transforms, convolution, co-occurrence matrix and fractals, can be used to obtain texture information [17].

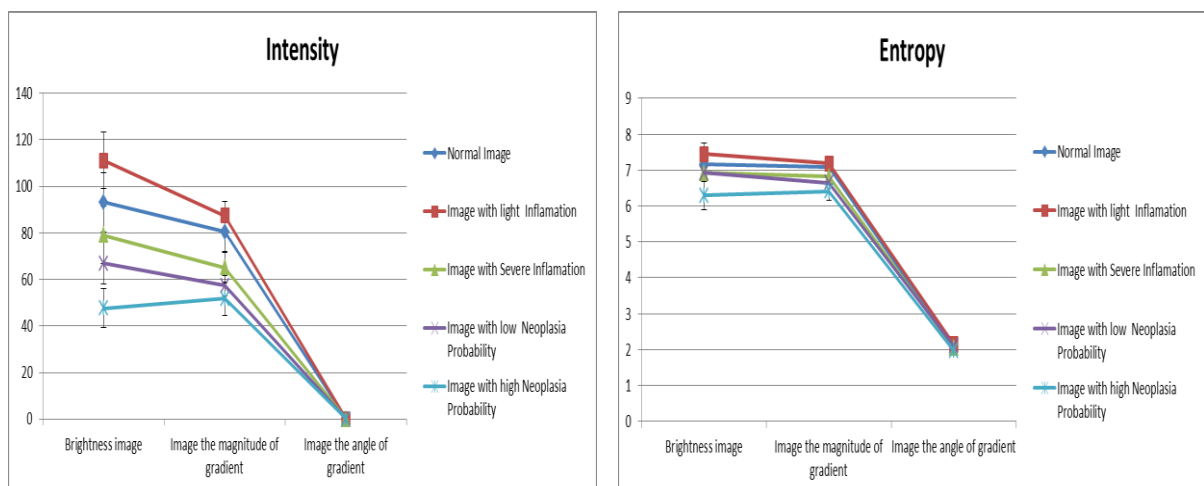
## CHAPTER 4: DATABASE ANALYSIS

In this project we use the statistical approaches with the following techniques: **entropy**, **pixels intensity**, the **second order moments**, the **fractal dimension**, the **skweness coefficient**, the **histogram analysis** and **co-occurrence matrices**. There are many features that can be extracted from gray level co-occurrence matrices but we will use only the energy and contrast. To describe the shape of the histogram we will use the mean and the skewness.

The statistical techniques are used on tree image types: the original image (brightness image), the gradient image magnitude and the gradient image angle.

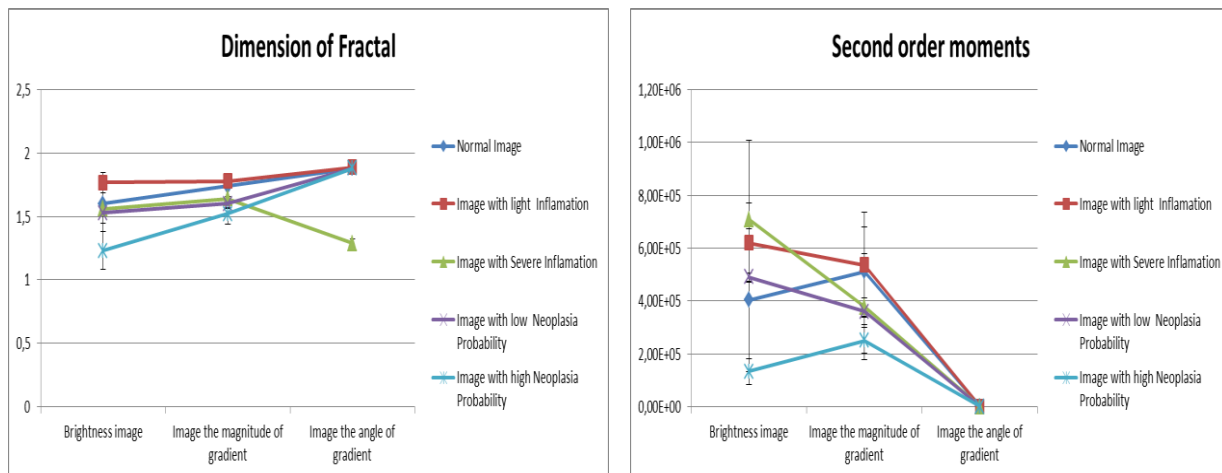


**Figure 19:** The tree image types in which we use the statistical techniques: the original image, the gradient image magnitude and the gradient image angle.

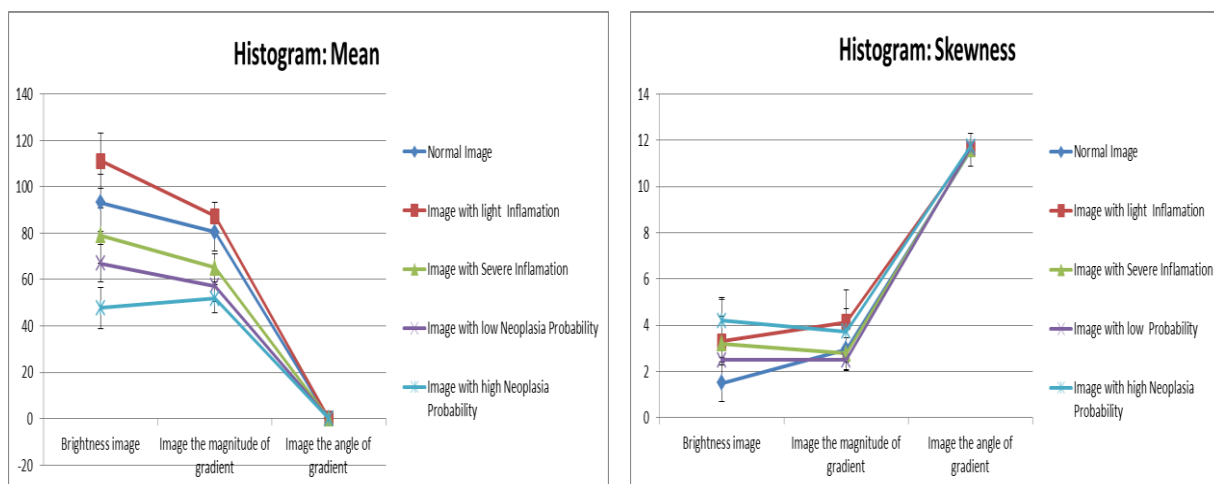


**Figure 20:** Representation graph of the mean intensity and entropy in original image, gradient image magnitude and gradient image angle for the different classes.

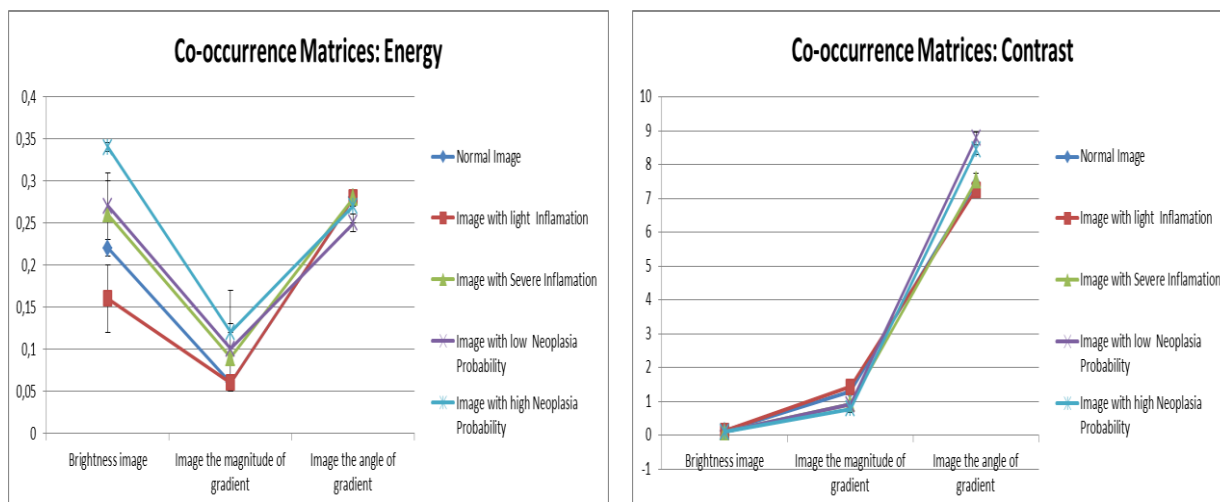
## CHAPTER 4: DATABASE ANALYSIS



**Figure 21:** Representation graph of the fractal dimension and second order moments in original image, gradient image magnitude and gradient image angle for the different classes.

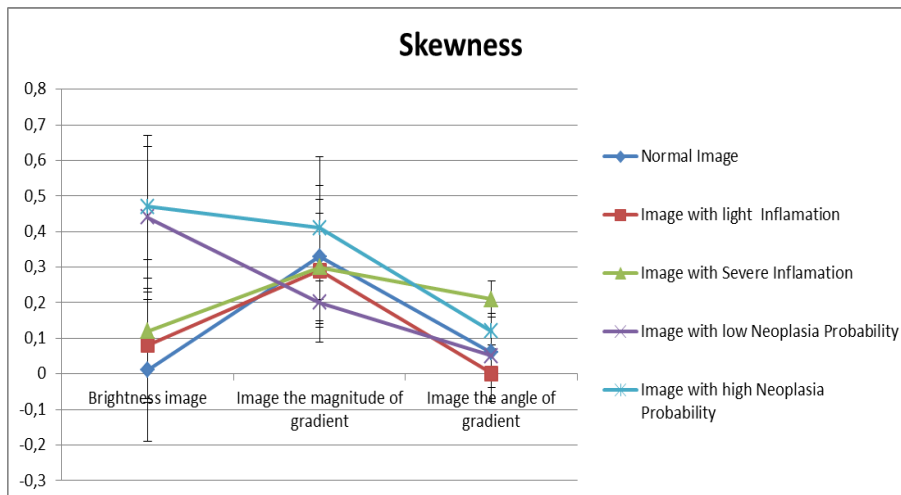


**Figure 22:** Representation graph of the mean and skewness of the histogram in original image, gradient image magnitude and gradient image angle for the different classes.



**Figure 23:** Representation graph of the energy and contrast of co-occurrence matrices in original image, gradient image magnitude and gradient image angle for the different classes.





**Figure 24:** Representation graph of the skewness in original image, gradient image magnitude and gradient image angle for the different classes.

#### 4.4 Discussion

In texture analysis we use the first and second order statistics approach. The first order texture estimates feature like average, entropy, skewness, histogram properties and it does not consider the interaction between image pixels. Contrarily, the second order texture takes into account the relation between the pixels and estimates properties like spatial gray level co-occurrence and second order moments etc. [18].

Analyzing the texture we observe that the intensity value of the pixels change significantly (it has a large error) for all classes. This is expected since the fluorescein concentration decrease during the acquisition image process.

Looking at the entropy results we can say that the endomicroscopy image is very unpredictable. The entropy mean does not alter much between the classes and the values are in the same range.

The moment allows evaluating the set shape points. The second order moment show how a set of point can be fitted by an ellipsoid. The results are in same order of magnitude for all classes and it has a high error.

Fractals are irregular geometric shape that can be fragmented. The fractal dimension allows knowing how rough and complex an object is. The values of fractal dimension acquired in each image do not present a significant variation between the classes.

## CHAPTER 4: DATABASE ANALYSIS

The skewness assesses the distribution asymmetry of the random variable. Its values can be positive or negative, or even undefined. Examining the skewness results we observe that the images with normal tissue and the image with inflammation are more symmetric than others. But this parameter is not discriminative enough.

To characterize the histogram shape we use the mean and skewness. The mean results are similar to the intensity results. The differences between image types are not significant for the two parameters.

The gray level co-occurrence evaluates image features related to second-order statistics. Relatively to the results, there is no considerable variation between the classes.

Analyzing the global results of the texture analysis we can say that the results obtained from texture analyses do not allow distinguishing the image classes. So it demonstrates that the different types of images have similar texture.

## Chapter 5

# Detection and Segmentation of Crypts

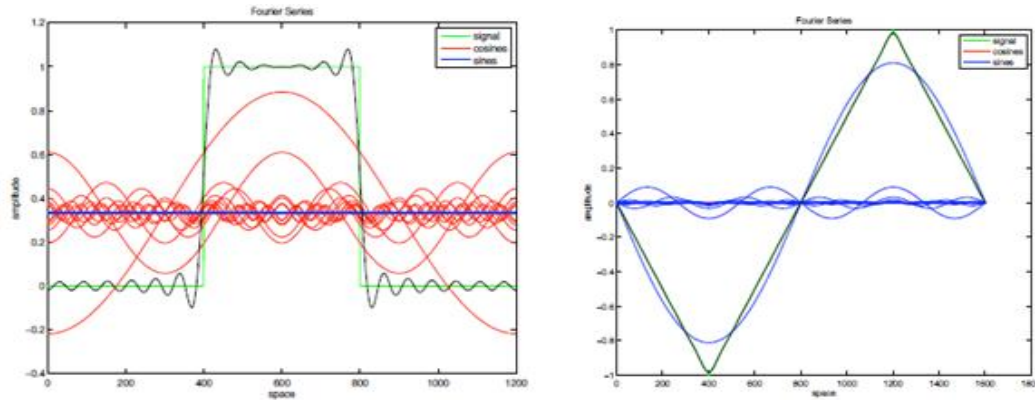
As demonstrated in the previous chapter the number and shape of the crypts are the parameters that enable to distinguish the classes. Therefore in this chapter we aim at segmenting and detecting the crypts.

Analyzing the endomicroscopic image we can see that they present high symmetry. So we use the symmetry energy to detect these structures. The first section explains how the symmetries and anti-symmetries work. The second section describes the algorithm used to measure the symmetry and the parameters of this algorithm are explained in details. In the last section we present the symmetry results obtained and method used to detect the crypts.

### 5.1 How do the symmetries and anti-symmetries work?

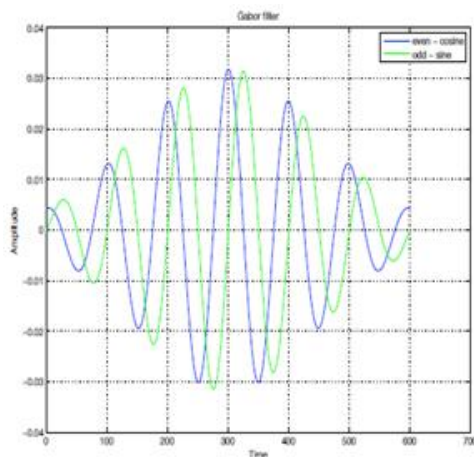
The endomicroscopy image exhibit high symmetry. Therefore to detect the crypts we use the symmetry energy. There can only be symmetry if there is periodicity in the image signal. For this reason we will analyze the signal in the frequency domain because it is easier to approach the periodicity in this domain than in the time domain. Examining Figure 10 it can be seen that the points of the Fourier series more symmetry are the maximum or minimum in their cycles. These points also correspond to the most symmetric points in their cycles. In the same logic the asymmetric points of the Fourier series correspond to the most asymmetry points in their cycles [19, 20].

## CHAPTER 5: DETECTION AND SEGMENTATION OF CRYPTS



**Figure 25:** Schematic representation of Fourier series of square wave and of triangular wave.

In this work, as proposed by Kovesi in [19], we use the Wavelet Transform to extract local frequency information. The Wavelet Transform is a bank of filters that work at various scales and allows analyzing the signal.



**Figure 26:** Schematic representation of Gabor pair.

From the results analysis obtained by signal convolution with two wavelets in quadrature, we verify that in the symmetric point the response of even-symmetric filters is much greater than the response of odd-symmetric filters. In the case of a point being asymmetry it happens the opposite, that is, the absolute value of the odd-symmetric filter output is much greater than the absolute value of the even-symmetric filter output. So, Kovesi [19] defined the measure of symmetry and asymmetry through the equation presented below.

The measure of symmetry is:

$$\begin{aligned} Sym(x) &= \frac{\sum_n [A_n(x) [|\cos(\phi_n(x))| - |\sin(\phi_n(x))|] - T]}{\sum_n A_n(x) + \varepsilon} \\ &= \frac{\sum_n [ |e_n(x)| - |o_n(x)| ] - T}{\sum_n A_n(x) + \varepsilon} \end{aligned}$$

The measure of asymmetry is defined as:

$$ASym(x) = \frac{\sum_n [ |o_n(x)| - |e_n(x)| ] - T}{\sum_n A_n(x) + \varepsilon}$$

Where  $A_n(x)$  is the amplitude of the wavelet and the index  $n$  specifies the scale of the filter. The parameters  $e_n(x)$  and  $o_n(x)$  correspond to the even-symmetric part and odd-symmetric part of the filter respectively. The term  $T$  is the compensation of the noise, and  $\varepsilon$  is a constant that avoid division by zero.

## 5.2 Algorithm to measure the Symmetry

As we want to detect the symmetry in different image orientation we use the code available on-line in [21] that enables to measure the symmetry from the local frequency information. The kovesi implementation receives the following parameters:

- *nscals*, the number of wavelet scales.
- *norient*, the number of filter orientations.
- *minWaveLength*, the wavelength of the smallest filter and this also represents the maximum center frequency once  $f_{c\max} = \frac{1}{\lambda_{\min}}$ .
- *mult*, the scaling factor between successive filter scales.

## CHAPTER 5: DETECTION AND SEGMENTATION OF CRYPTS

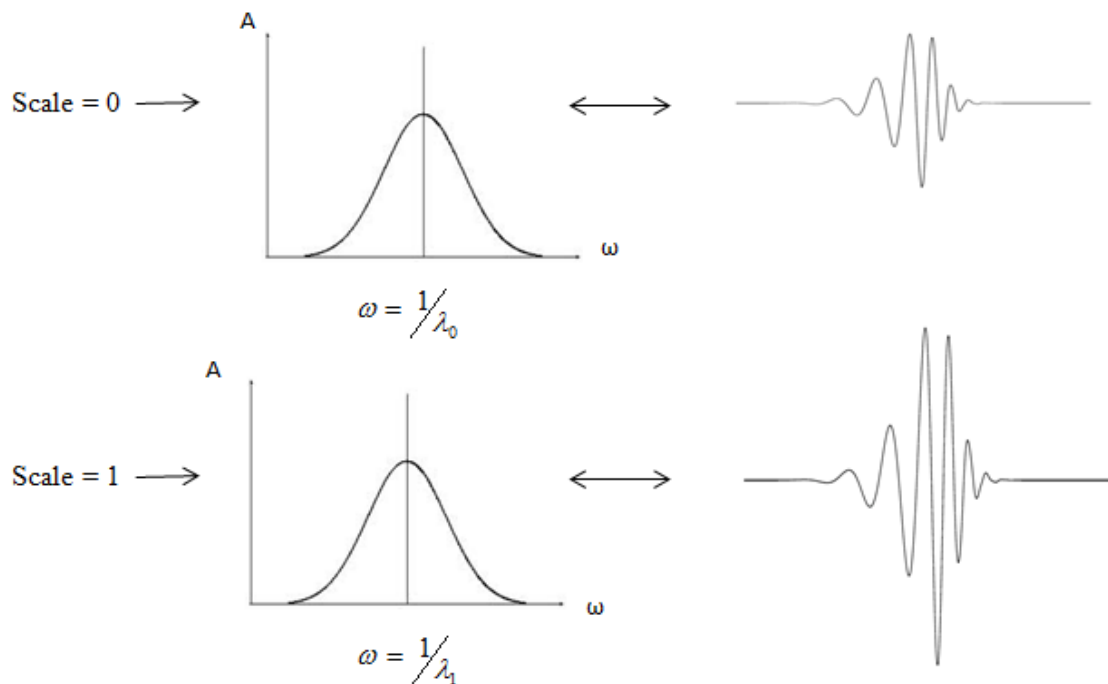
- *polarity*, this parameter allows choosing the feature of symmetry that we want to find, in other word, it is possible return only ‘bright’ symmetry points or only ‘dark’ symmetry points or yet all symmetry points (‘bright’ and ‘dark’).

This implementation returns the followed values:

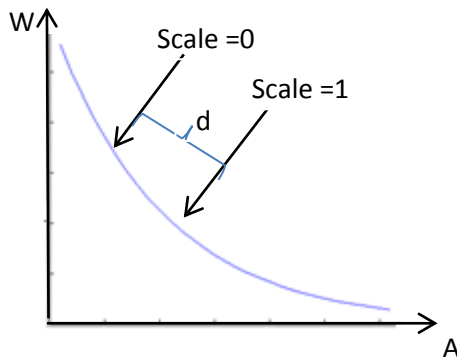
1. *phaseSym*, the phase symmetry image with values between 0 and 1.
2. *orientation*, the orientation in which local symmetry energy is a maxima.
3. *totalEnergy*, corresponds the unnormalised symmetry energy.

### Example:

Supposes that we choose the filter with minimal wavelength  $\lambda_0$ , with the parameter *n scales* equal to 1. So the image is convolved successively with the following filters:



**Figure 27:** Schematic representation of the evolution wavelet amplitude.



**Figure 28:** Schematic representation of the frequency evolution in different scales.

As shown in Figure 27, an increase in the scale number of the filter also increases the wavelet amplitude. The wavelength of each scale is the minimal wavelength multiplied by the parameter *mult* power of scales. So  $\lambda_1 = \lambda_0 \times mult^1$  and  $\lambda_2 = \lambda_1 \times mult = \lambda_0 \times mult^2$  and so for different scales.

Figure 28 shows the frequency evolution in different scales. The wavelength  $\lambda_0$  determines the initial frequency and the *mult* parameter defines the distance *d* (letter *d* in figure 28).

### 5.3 Results



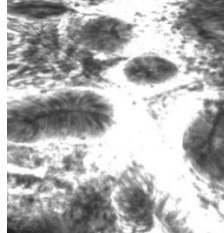
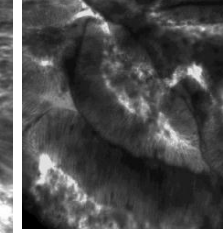
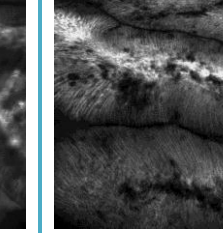
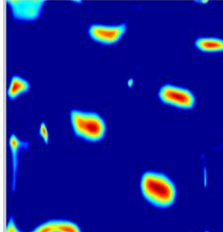
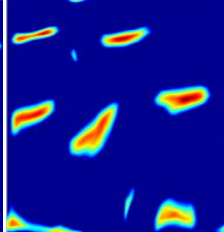
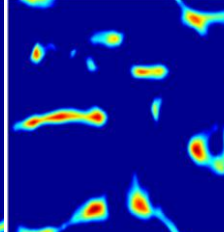
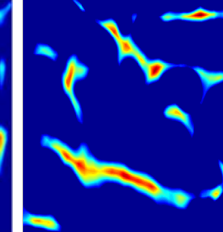
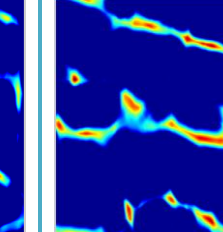
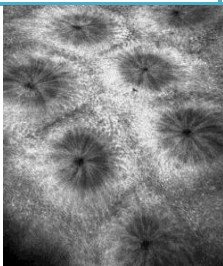
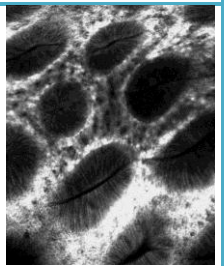

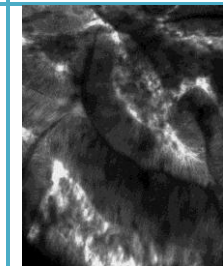
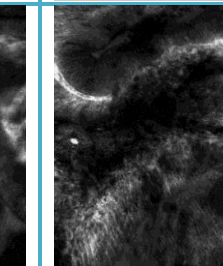
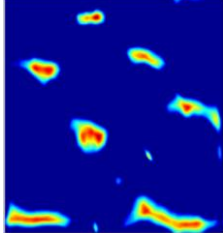
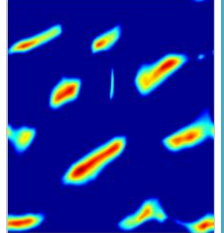
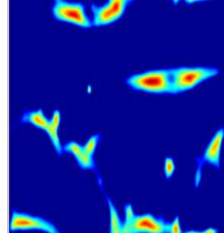
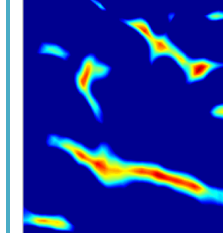
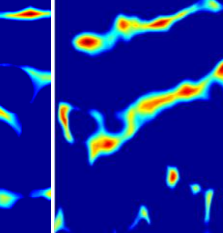
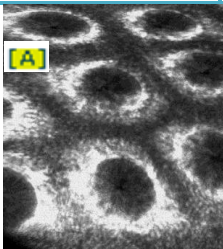
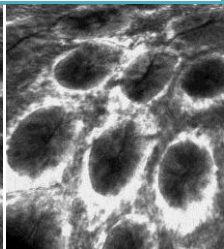
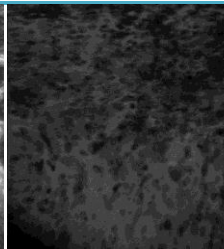
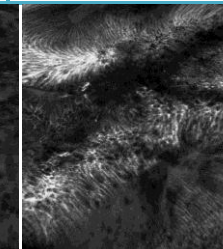
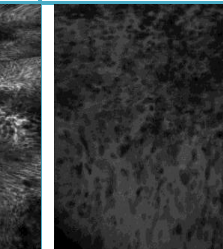
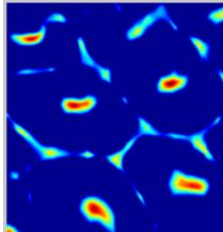
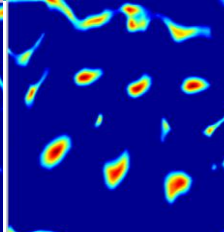
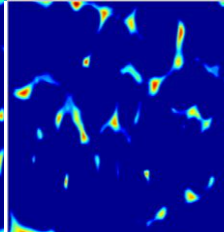
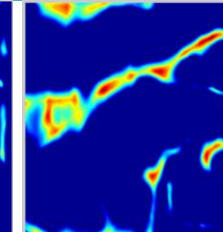
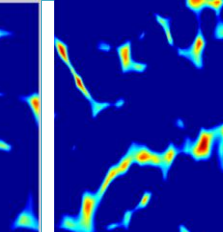
We use the Kovesi's matlab code to detect the symmetries in image. We detected the symmetry energy of the dark points.

Table 6 shows the symmetry result to different classes of image. By analyzing of results we can see that the crypts of normal images and images with light inflammation present a good energy of symmetry. In these images the crypts are well segmented because there is a great contrast between crypts and lamina propria. But there are cases where the share of lamina propria is dark, (e.g. the image marked with the letter *A* in table 6) in these cases the result is not favorable because the lamina propria is also segmented.

In the other classes we do not have a great contrast between crypts and lamina propria since the lamina propria is very dark or present dark points (both structures are segmented). Analyzing globally the results of symmetry we can conclude that the symmetry measured from the local frequency information enables to segment the crypts.

## CHAPTER 5: DETECTION AND SEGMENTATION OF CRYPTS

**Table 6:** Example of some image obtained from energy of symmetry (column pair) and brightness image (odd column).

Normal Image	Image with Light Inflammation	Image with Severe Inflammation	Image with Low Neoplasia Probability	Image with high Neoplasia Probability
				
				
				
				
				
				



## CHAPTER 5: DETECTION AND SEGMENTATION OF CRYPTS

To detect the crypts and its contour we tested various techniques like Hough Transform, Mean Shift, active contour and ellipse fitting. We apply these methods in image symmetry energy.

### Hough Transform

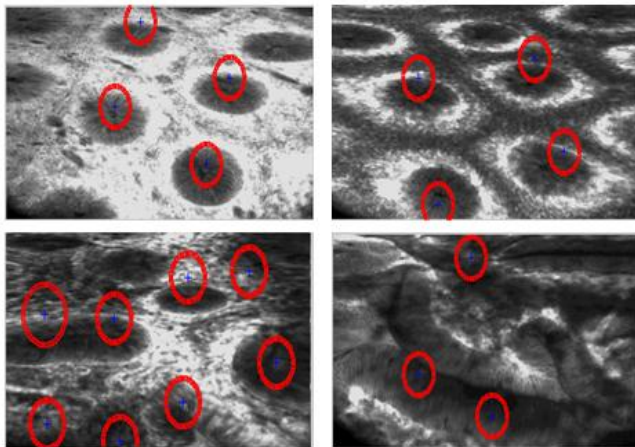
The Hough transform is a technique that allows finding object with specific shape within an image. The classical Hough transform consist in detecting lines in the image, but later the generalized Hough transform applied this idea to identify arbitrary parameterizable shapes, most frequently circles or ellipses. In this work we use the Hough transform for circles because the crypt shapes in normal image are circular.

A circle with center  $(a, b)$  and radius  $R$  can be described by the following equation:

$$x = a + R\cos(\theta)$$

$$y = b + R\sin(\theta)$$

The circle Hough transform determines the tree parameters  $a$ ,  $b$  and  $R$  in region that are probably circles in the image.

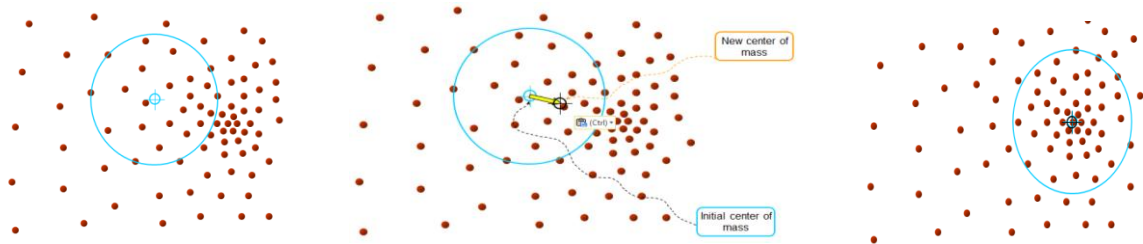


**Figure 29:** Results acquired by Hough transform applied in the symmetry image. The detected circles are marked with red color.

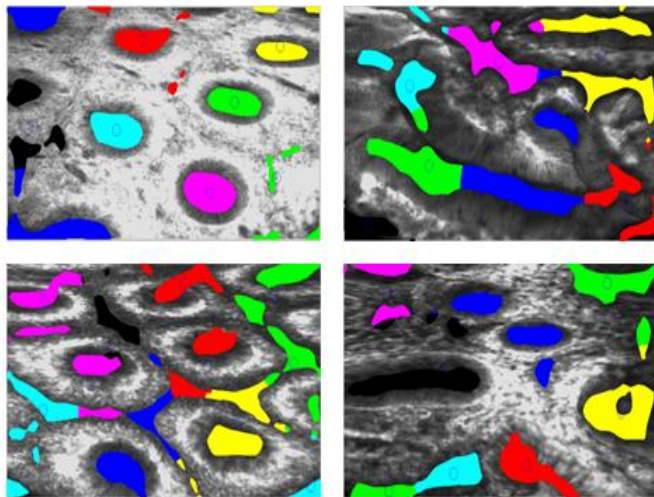
As we can see in Figure 29 the circle Hough transform presents bad results in crypts detention once it does not detect all the crypts in normal image and the crypts detected do not have the correct radius. In images with pathology it detects circles they are false positives.

**Mean Shift**

Mean Shift is an iterative procedure for finding the maximum in a set of data samples and it does not require that the clusters number is known a priori [22]. The principal idea behind Mean Shift is to locate the densest region and the mean shift procedure is the following:



1. Randomly select the region of interest (kernel).
2. Calculate the mass center of the data.
3. Move the kernel where to find the new mass center.
4. Repeat the steps 2 and 3 until convergence.



**Figure 30:** Results acquired by Mean Shift applied in image of the symmetry. The clusters centers are marked with blue circle.

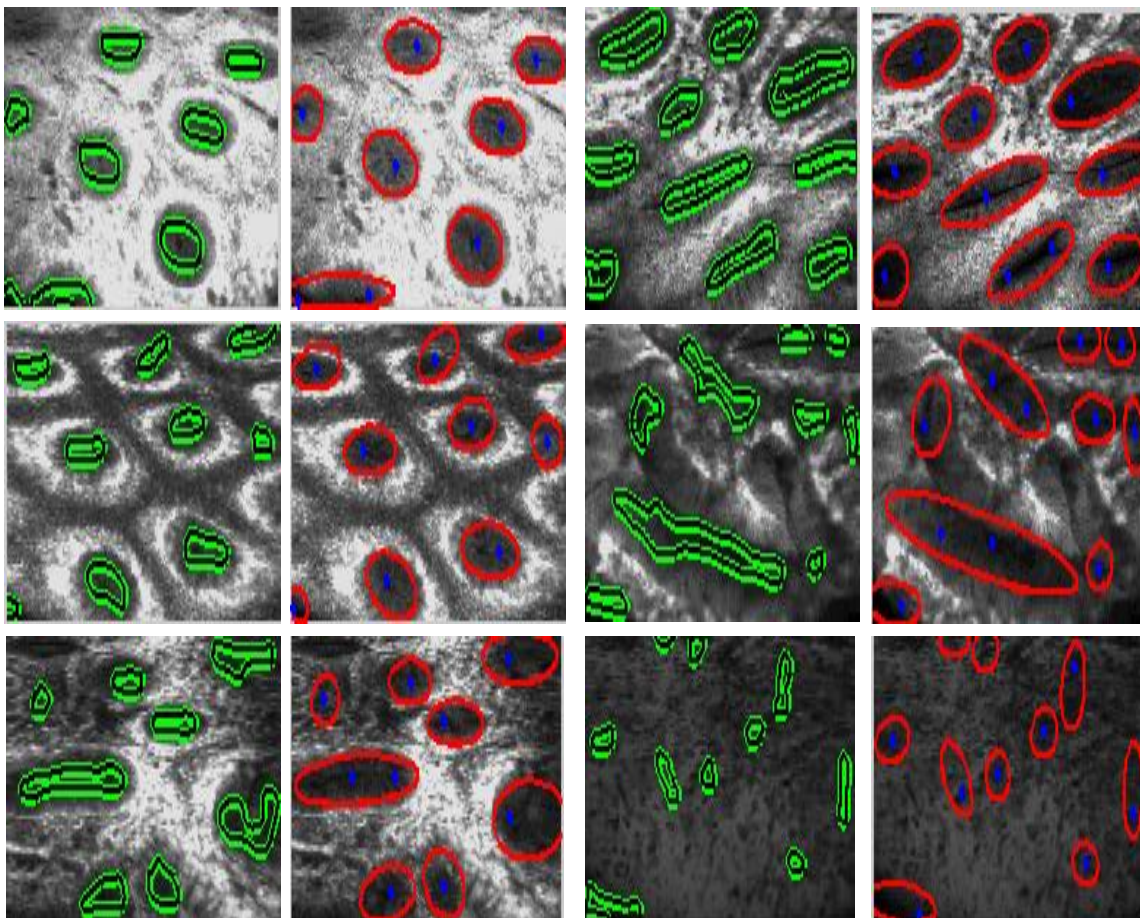
The results in Figure 30 shows that, for the majority of images the cluster centers (blue small circle) do not matches with the crypt centers. There are clusters composed

by more than one crypt and crypts fragmented in many clusters. By these reason we conclude that the mean shift does not allow detecting the crypts.

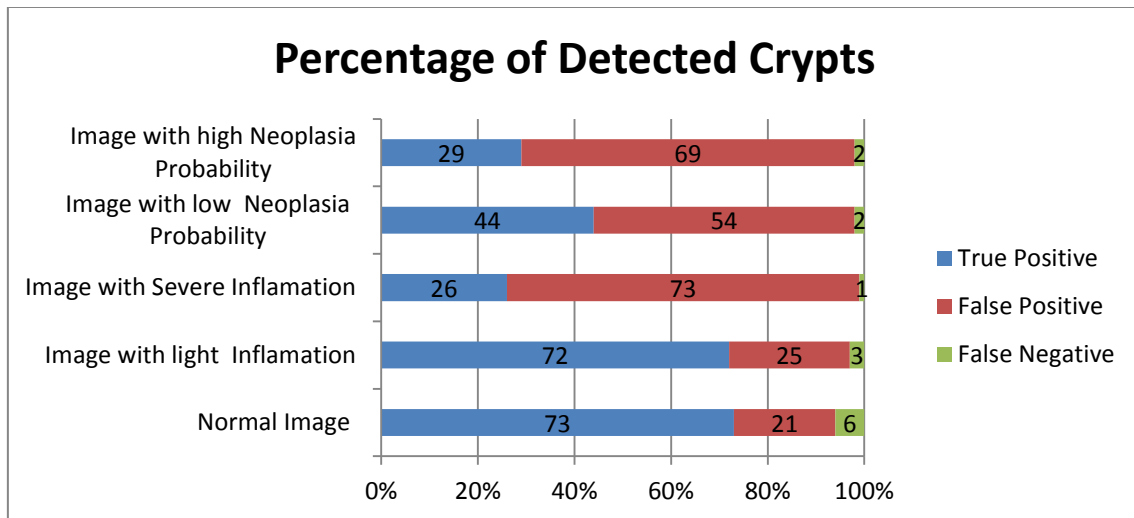
**Local Maximum, Active Contour and Ellipse fitting**

As we can verify in the symmetry image results, the greatest energy corresponds to the crypts center. So we use the local maximum to find the crypts lumen. To detect the crypts shape we utilize the active contour without edges [23], code available on-line in [24]. The lumen location allows obtaining the initial contour.

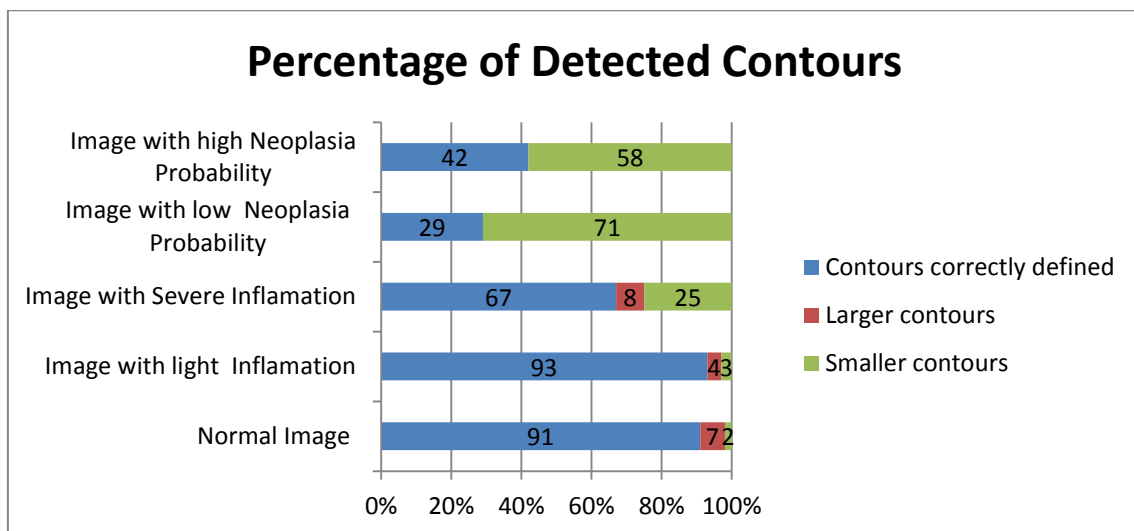
Generally the contour obtained with this method is smaller than the original crypts thus to detect the crypts shape we use the ellipse fitting together with the binary morphological operator dilation. Figure 31 shows some results obtained using the techniques described previously.



**Figure 31:** Results acquired by Active Contour applied in symmetry image. The contours of the crypts are marked in green color. The ellipse fitting was applied on the symmetry image. The ellipses of the crypts are marked in green color.



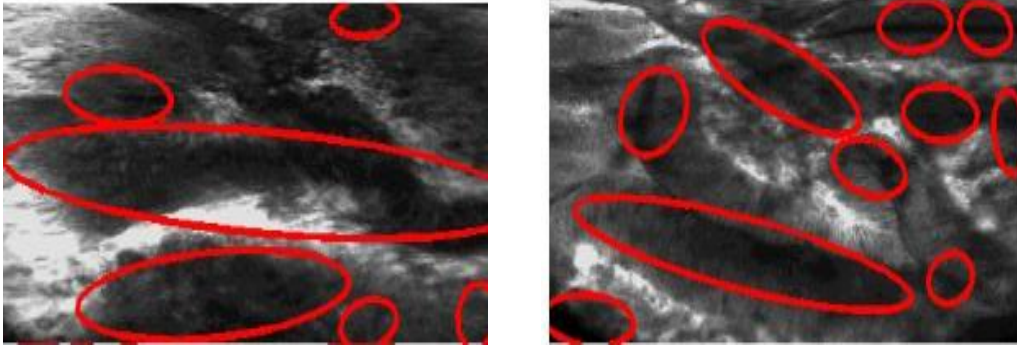
**Graphic 8:** Graph representation of the percentage of detected crypts acquired by local maximum, Active Contour and ellipse fitting applied in symmetry images.



**Graphic 9:** Graph representation of the percentage of detected contours acquired by local maxima, Active Contour and ellipse fitting applied in symmetry image.

The Graphic 8 and 9 we observed that in the majority of normal images (73%) and the image2 with light inflammation (72%) the crypts are well detected and it's contour correctly defined (above 90%).

Regarding the images of unhealthy tissue, the results are not satisfactory because the crypts are mixed-up with dark shares of lamina propria. There are many false positives because there are some structures in the laminas propria identical to the crypts. Regarding the contours the results is also bad because the majority of detected contour is smaller than the original.



**Figure 32:** Example of ellipse fitting to detect crypts in image with severe inflammation (left) and image with high probability of neoplasia (right).

The two images shown above are similar but the left image correspond to an image of tissue with severe inflammation and do not have crypts. Nevertheless six crypts were detected. The right image corresponds to an image with high neoplasia probability and it has five crypts, but ten crypts were detected. Due these disorders and incongruence the results to detect crypts in these classes is not satisfactory.

These results meet the results obtained in symmetry energy. This shows that the local maxima together with ellipse fitting present good results, but it is necessary to improve the symmetry energy. We can affirm that the naïve classification does not work with endoscopic image and that it is necessary to take into consideration much more information and features.

## Chapter 6

### Conclusion and Future Work

This chapter evaluates and draws some conclusions regarding the objectives achieved through this project.

#### 6.1: Conclusion

The main goals of this project are to develop an application that allows storing and annotating the image acquired by confocal endomicroscopy and to use the collected data to develop basic algorithms for crypts detection and segmentation. To achieve these objectives it was necessary to execute the following task:

- Study of the basic principle of confocal microscopy to fully understand the endomicroscopic system; observation of clinical procedure; understanding of diagnosis criterion.
- Establishment of a protocol for the systematic images acquisition by the doctors.
- Development of an application in *Java* for labeling the images and build a database.
- Research in pattern recognition techniques: detection and segmentation of the crypts.

The developed application is easy to use and practical so the doctors have little difficulties in making familiar with it.

In relation to crypts segmentation and detection the results show that the technique that uses symmetry [19] presents good results but not enough for all classes. Therefore in the future it is needed to take into consideration other feature to complement this technique.

### 6.2: Future Work

The results of this study show that the naïve classification does not allow characterizing and distinguishing the different images types due the great variability intra-classes and small differences between classes. To do an accurate classification it will be necessary a large database with many examples of each class to train the automatic classifier. To acquire large quantities of data the application developed to build the database need changed, in order to support multiple local clients installed in different medical user's computers.

In relation to pattern recognition, specifically in crypts segmentation and detection, the obtained results is not satisfactory especially in image of tissue with some pathology.

As mentioned previously the naïve classification does not enable to discriminate the endomicroscopic image, so to develop a computational tool for automatic interpretation of endomicroscopic image it will be necessary: (i) to acquire image with high quality. The main artifact of CEM images is caused by the endoscope motion. In the future it can develop a tool that enables the doctor to keep the endoscope fixed. (ii) In classification to distinguish the normal and pathological images it is essential to complement the technique of the symmetry energy with other characteristic and information, namely the information of the spatial arrangement.

# Appendix A

## Protocol

To do the images insertion in the database it is necessary to obtain the images with optimal quality and export them out of the clinical *software*.

### 1. Images Acquisitions:

In system calibration: (i) define the laser power such that the image highlights the indicators of each image types; (ii) Choose the slower scanner that has better resolution.



Figure 33: Image Acquisition screens.

After the exam conclusion the images is saved in the file to export.

### 2. The data exportation:

To do the data exportation it is necessary to navigate to the initial screen and select the review mode. The exported file is saved as ZIP file. Select the button *Review Procedure Results* that allows to review and export archived processes. After selecting this button the new window that allows doing choose of process is opened. Next select



## APPENDIX A PROTOCOL

the searching criterion (can be name, surname, ID, among others) click on the button “*Next*”.

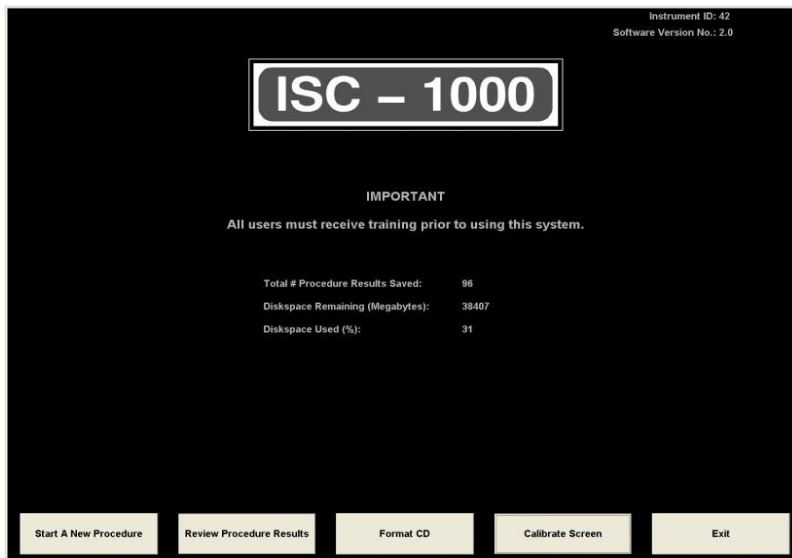


Figure 34: Start-up screens.

Subject Details

Subject ID: \_\_\_\_\_

Surname: \_\_\_\_\_

First Name: \_\_\_\_\_

Middle/Initial: \_\_\_\_\_

Title: \_\_\_\_\_

Sex:  Male  Female  Unspecified

D.O.B.

Month: \_\_\_\_\_ ('01' or '1' for January)

Day: \_\_\_\_\_ ('01 or '1' for 1st)

Year: \_\_\_\_\_ ('2000' for 2000)

Telephone Numbers:

Home: \_\_\_\_\_

Business: \_\_\_\_\_

Mobile: \_\_\_\_\_

Address:

Street: \_\_\_\_\_

City: \_\_\_\_\_

State: \_\_\_\_\_

ZIP/Postcode: \_\_\_\_\_

Country: \_\_\_\_\_

Clear Search Fields      Next      Exit

Figure 35: Procedure results search criteria screens.

In the Review procedure results screens, click in the button *Export Results*. The default exported process name is a set of numbers. It is necessary to change the file name for *word* and then choose the place where it wants to save the exported process.

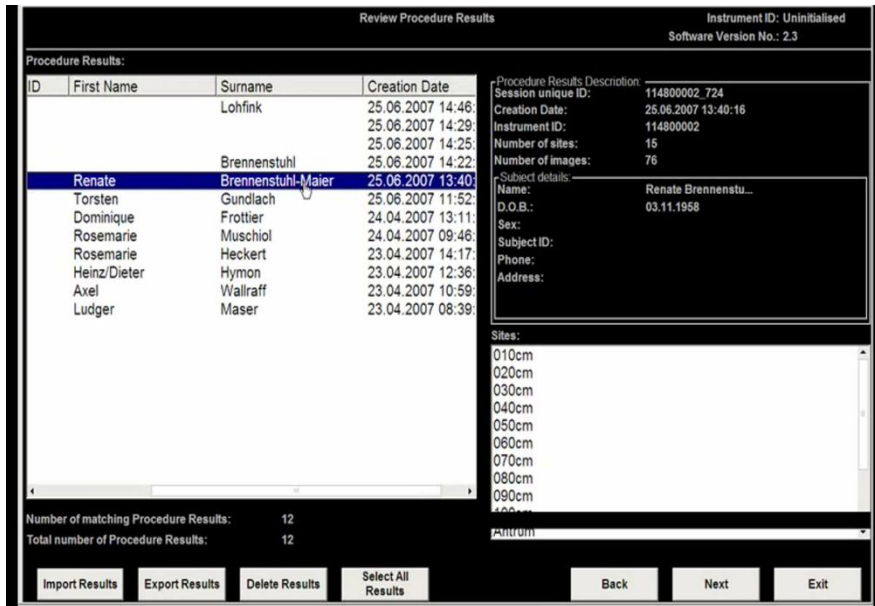


Figure 36: Review procedure results screens.

### 3. Annotation and insertion in the data base

To insert the exported data use the developed application. This application must always be in the same location of its library (file lib). This application open with a double click. The *Login* screen is equal to all users, it allows the user to access the data or make a new registration if it is an administrator.

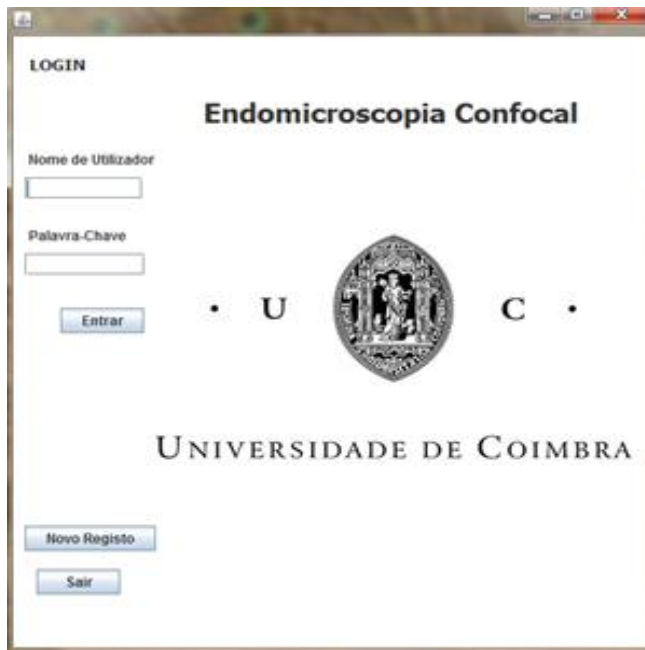


Figure 37: Schematic representation of the first application screen – *Login* screens

## APPENDIX A PROTOCOL

If the user is new it is necessary that the administrators make a new registration (button new registration in the figure 5). The new registration screen has several fields filled by default. If necessary change these fields. The administrator can fill the fields and confirm the registration in the end. In the category field, the administrator can choose if the new user is a normal user or an administrator.



**Figure 38:** Schematic representation of the first application screen – *Novo Registo* screens.

After the login the *Processos* windows is opened and we can visualize all the processes discriminated by: name, exam number, the responsible doctor and the current diagnosis.

There are several buttons in this window with different functions. The *Pesquisar* button allows to search the process by name. The *Novo Processo* and *Novo Exame* buttons serve to introduce a new process and new exam respectively. The button *Ver Exames* allows to visualize all the exams of the selected process. The button *Abrir* shows all information of a process and we can modify this information. The button *Sync* allows to make a synchronization between two databases: local and remote.

## APPENDIX A PROTOCOL

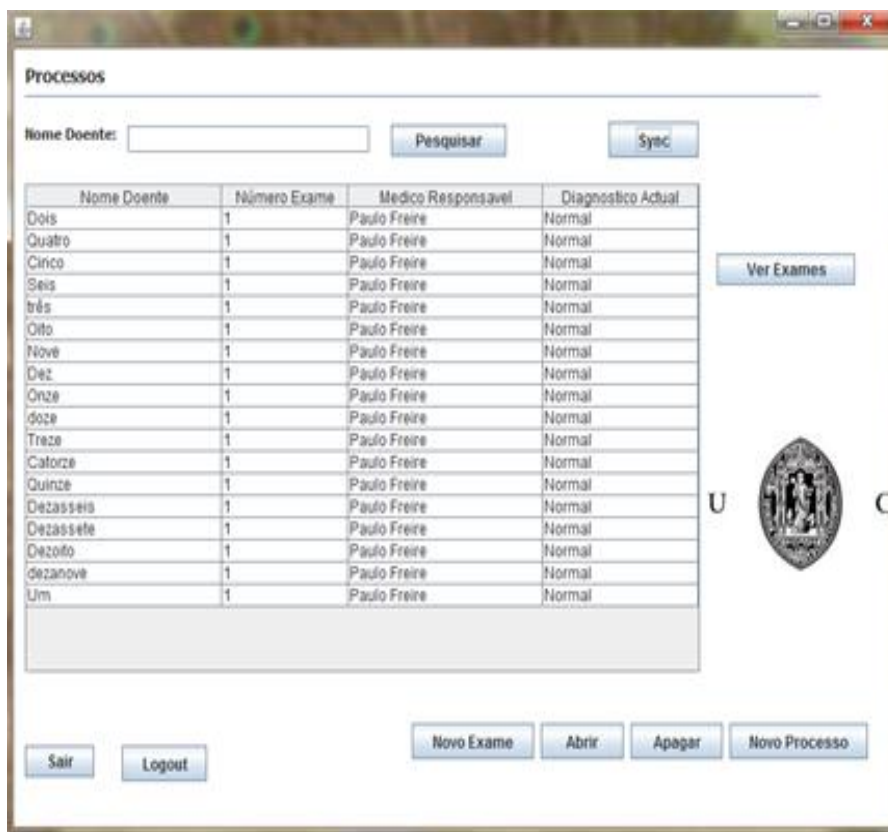


Figure 39: Schematic representation of the third application screen – *Processos* screens.

After the doctor selects the button *Novo Processo*, the window *Novo Processo* appears. In this screen the doctor should fill all the fields that characterize the process.

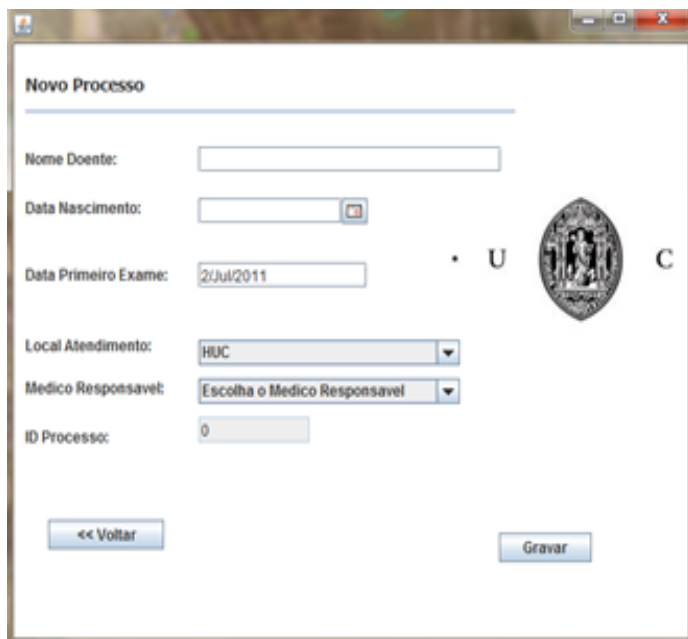
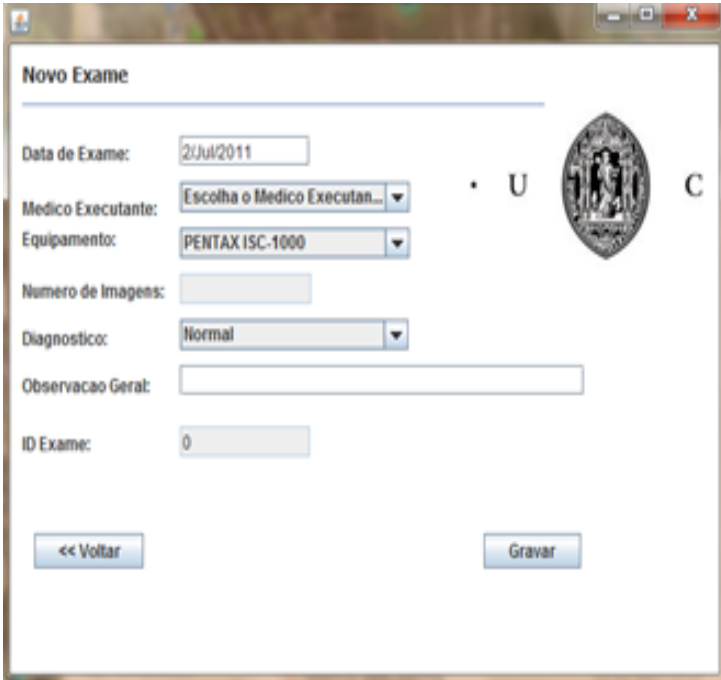


Figure 40: Schematic representation of the fourth application screen – *Novo Processo* screens.

## APPENDIX A PROTOCOL

After inserting a new process the *Novo Exame* window opens automatically. The doctor should fill the all exam fields. The field *Medico Executante* corresponds to the doctor that realized the exam and it can be different of the doctor that is following the patients.



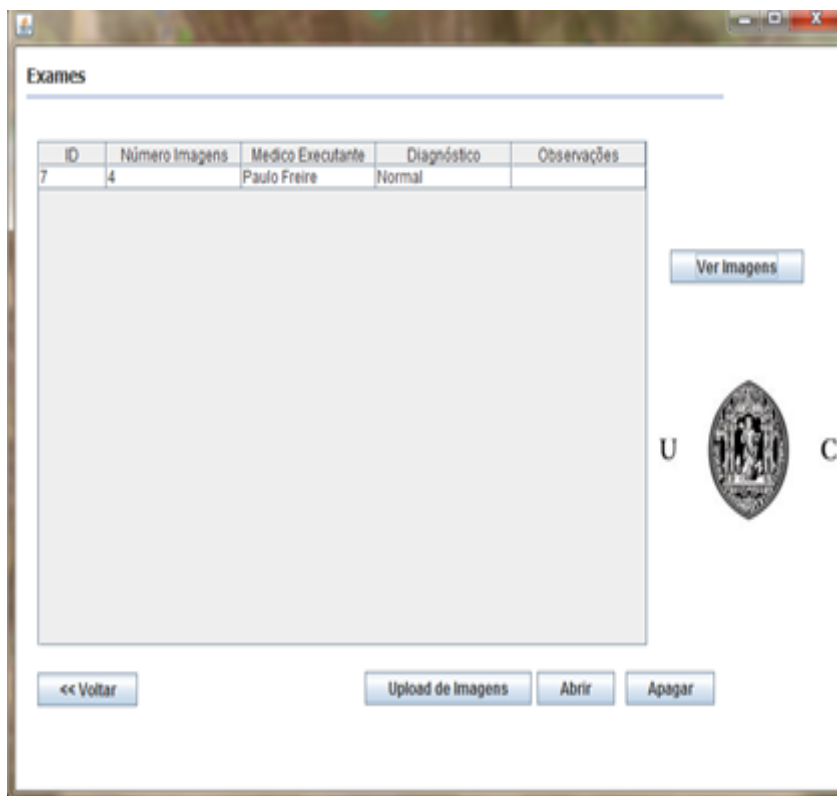
The image shows a screenshot of a software window titled "Novo Exame". The window contains the following fields and controls:

- Data de Exame:** A text input field containing "2/Jul/2011".
- Medico Executante:** A dropdown menu with the selected option "Escolha o Medico Executan...".
- Equipamento:** A dropdown menu with the selected option "PENTAX ISC-1000".
- Numero de Imagens:** An empty text input field.
- Diagnostico:** A dropdown menu with the selected option "Normal".
- Observacao Gerat:** An empty text input field.
- ID Exame:** A text input field containing "0".

At the bottom of the window, there are two buttons: "<< Voltar" on the left and "Gravar" on the right. A circular logo is positioned to the right of the "Medico Executante" and "Equipamento" fields.

**Figure 41:** Schematic representation of the fifth application screen – *Exame* screens.

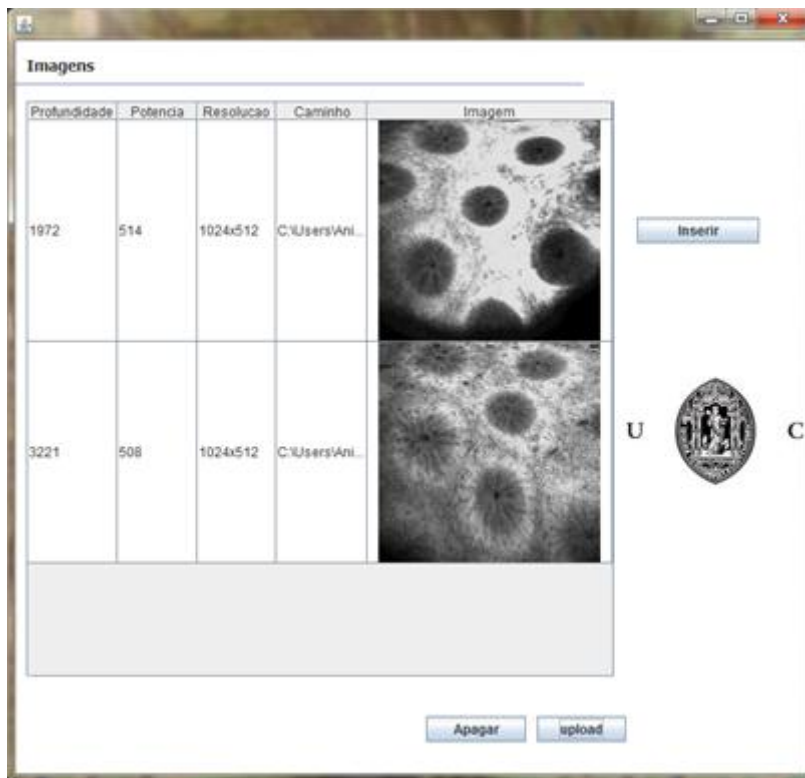
When we use the button *Ver Exames* in the *Processo* screen it opens the *Exames* windows where all exams identified by the images number, the executant doctor, the diagnosis and the observation are listed.



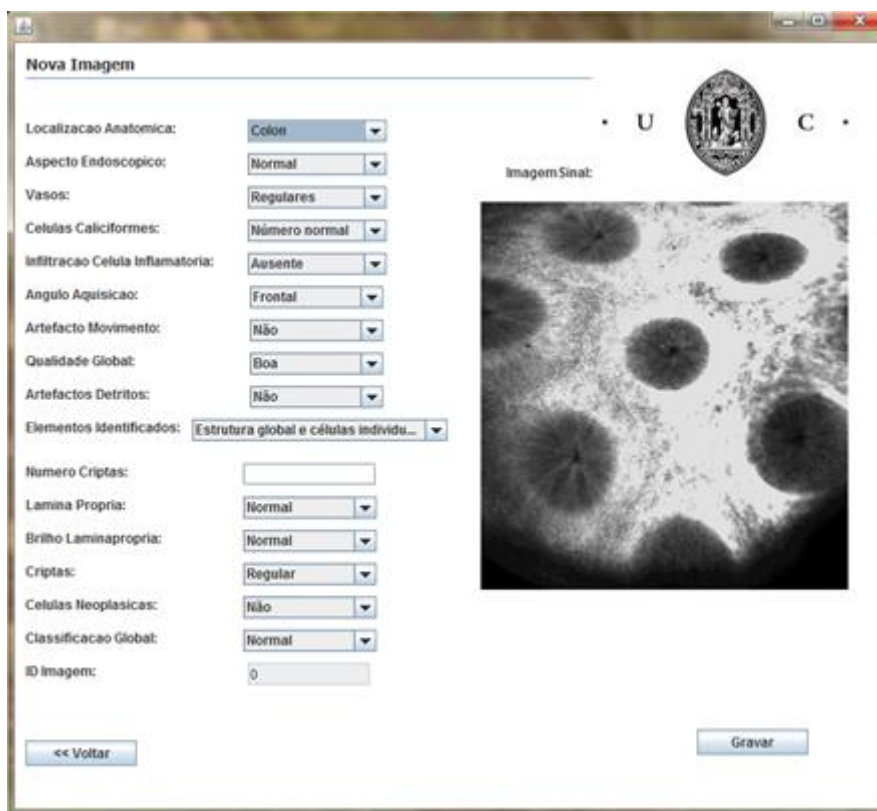
**Figure 42:** Schematic representation of the sixth application screen – *Exames* screens.

The screen *Imagens* arises after clicking in the button *Gravar* in the *Novo Exame* window or when we select *Upload de Imagens* on the previous screen. To insert the image in the database it is necessary to upload the file. For this we use the *Upload* button. The selected images are introduced in the database one by one. When we click in the button *Inserir* the window *Nova Imagem* appears with all the attributes necessary to characterize an endomicroscopic image and distinguish the different types of image.

## APPENDIX A PROTOCOL



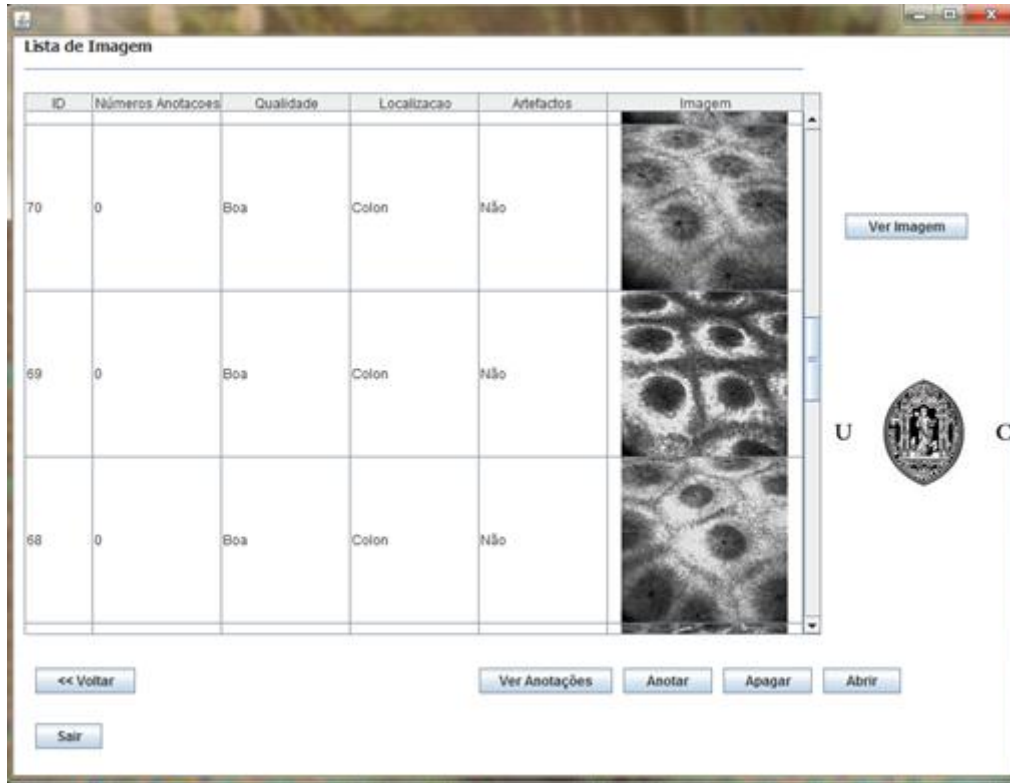
**Figure 43:** Schematic representation of the seventh application screen – *Imagens* screens.



**Figure 44:** Schematic representation of the eighth application screen – *Nova Imagem* screens.

## APPENDIX A PROTOCOL

When clicking in button *graver*, one of the two following events can occur: 1) come back to the previous screen if there are images to insert; 2) the window *Lista de Imagem* appears where we insert the last image.



**Figure 45:** Schematic representation of the ninth application screen – *Lista de Imagem* screens.

This window has several buttons that have the following functions: The *Abrir* button shows all the attributes introduced by the doctor about the selected image and it also allows modifying the image information. The button *Ver Imagem* is like a button *Abrir*, but this show all the information about the image and it includes the technical data and the annotations. The button *Anotar* allows the doctor to select an area of the image that is relevant and then annotate. The button *Ver Anotações* enables to see all annotations that the selected image has. The button *Sair* closes the application.

After selecting the button *Ver Imagem* the tenth screen of this application, *Imagem* windows, is opened. In this window all parameters that characterize the image including the technical features are illustrated.



## APPENDIX A PROTOCOL

**Imagem**

Localizacao Anatomica: Colon

Aspecto Endoscopico: Normal

Vasos: Regulares

Celulas Caliciformes: Número normal

Infiltracao Celula Inflammatory: Ausente

Angulo Aquisicao: Frontal

Qualidade Global: Boa

Artefacto Movimento: Não

Artefactos Detritos: Não

Elementos Identificados: Estrutura global e células individuais

Numero Criptas: 6

Lamina Propria: Normal

Brilho Laminapropria: Intenso

Forma Criptas: Regular

Celulas Neoplasicas: Não

Classificacao Global: Normal

Numero Sequencia: 17

ID Imagem: 76

Imagem Sinal:

Estrutura	Caracteristicas
Cripta	Regulares
Cripta	Regulares
Cripta	Regulares
Cripta	Regulares

Caracteristicas Técnicas:

Potência: 513

Resolução: 1024x512

Profundidade: 2088

<< Voltar

**Figure 46:** Schematic representation of the tenth application screen – *Imagem* screens.

Next to use the button *anotar* the window *Nova Anotação* is opened. The doctor can select an area of the image that is relevant and then annotate. There are predefined fields to make note of the crypts, vases, lamina propria, neoplastic cells, etc.

**Nova Anotação**

Criptas: [dropdown]

Vasos: [dropdown]

Lâmina própria: [dropdown]

Células neoplásicas  Estrutura Indiferenciada

Observacao: [text area]

TLC: X; Y

BRC: X; Y

ID Anotacao: 0

<< Voltar Gravar

**Figure 47:** Schematic representation of the eleventh application screen – *Anotação* screens.

## Appendix B

### Code to create the Tables of data base

#### Table: MEDICO

```
create table MEDICO (  
  NOME_MEDICO      VARCHAR(100)    not null,  
  ESPECIALIDADE    VARCHAR(100)    null,  
  LOCAL            VARCHAR(100)    null,  
  NOME_UTILIZADOR  VARCHAR(100)    not null,  
  PALAVRA_CHAVE    varchar(15)     not null,  
  EQUIPAMENTO      varchar(100)    not null,  
  CATEGORIA        varchar(100)    not null,  
  constraint PK_UTILIZADOR primary key (NOME_UTILIZADOR)  
);
```

#### Table: PROCESSO

```
create table PROCESSO (  
  NOME_DOENTE      VARCHAR(100) not null,  
  DATA_NASCIMENTO VARCHAR(20)  null,  
  DATA_PRIMEIRO_EXAME VARCHAR(20) not null,  
  DATA_ULTIMO_EXAME VARCHAR(20) not null,  
  NUMERO_EXAME     INT          null,  
  LOCAL_ATANDIMENTO VARCHAR(100) null,  
  MEDICO_RESPONSAVEL VARCHAR(100) not null,  
  DIAGNOSTICO_ACTUAL VARCHAR(100) not null,  
  ID_PROCESSO      INT          not null,  
  constraint PK_PROCESSO primary key (ID_PROCESSO)  
);
```

#### Table: EXAME

```
create table EXAME (  
  DATA_EXAME      VARCHAR(20)    not null,  
  MEDICO_EXECUTANTE VARCHAR(100)  not null,  
  EQUIPAMENTO      VARCHAR(100)  not null,  
  NUMERO_IMAGEM    INT          not null,  
  DIAGNOSTICO      VARCHAR(100)  not null,  
  OBSERVACAOGERAL TEXT          null,  
  ID_EXAME         INT          not null,  
  ID_PROCESSO      INT          not null,  
  constraint PK_EXAME primary key (ID_EXAME)  
);
```

## APPENDIX B CODE TO CREATE THE TABLES OF DATA BASE

alter table EXAME

```
add constraint FK_EXAME_POSSUI_PROCESSO foreign key (ID_PROCESSO)
references PROCESSO (ID_PROCESSO)
on delete CASCADE on update restrict;
```

### Table: IMAGEM

```
create table IMAGEM (
NUMERO_ANOTACAO INT null,
LOCALIZACAO_ANATOMICA VARCHAR(100) null,
NUMERO_SEQUENCIA INT null,
ANGULO_AQUISICAO VARCHAR(100) not null,
Aspecto_endoscopico VARCHAR(100) not null,
QUALIDADE_GLOBAL VARCHAR(100) not null,
ARTIFACTO_MOVIMENTO VARCHAR(3) not null,
ARTIFACTOS_DETritos VARCHAR(3) not null,
Elementos_identificados VARCHAR(100) not null,
NUMERO_CRIPTAS INT not null,
LAMINA_PROPRIA VARCHAR(100) not null,
BRILHO_LAMINAPROPRIA VARCHAR(100) not null,
CRIPTAS VARCHAR(100) not null,
Vasos VARCHAR(100) not null,
Celulas_caliciformes VARCHAR(100) not null,
Infiltracao_Celula_Inflamatoria VARCHAR(100) not null,
CELULAS_NEOPLASICAS VARCHAR(10) not null,
Classificacao_global VARCHAR(100) not null,
IMAGEMSINAL TEXT not null,
PROFUNDIDADE INT not null,
POTENCIA INT not null,
RESOLUCAO VARCHAR(10) not null,
ID_IMAGEM INT not null,
ID_EXAME INT not null,
constraint PK_IMAGEM primary key (ID_IMAGEM)
);
```

alter table IMAGEM

```
add constraint FK_IMAGEM_CONTEM_EXAME foreign key (ID_EXAME)
references EXAME (ID_EXAME)
on delete CASCADE on update restrict;
```

### Table: ANOTACOES

```
create table ANOTACOES (
TLCX INT not null,
TLCY INT not null,
BRCX INT not null,
BRCY INT not null,
LAMINA_PROPRIA VARCHAR(100) null,
CRIPTAS VARCHAR(100) null,
Vasos VARCHAR(100) null,
Celulas_caliciformes VARCHAR(100) null,
CELULAS_NEOPLASICAS VARCHAR(100) null,
ESTRUTURA_INDEFERENCIADA VARCHAR(100) null,
OBSERVACAO TEXT not null,
```

## APPENDIX B CODE TO CREATE THE TABLES OF DATA BASE

```
ID_ANNOTACOES    INT    not null,  
ID_IMAGEM        INT      not null,  
constraint PK_ANNOTACOES primary key (ID_ANNOTACOES)  
);
```

```
alter table ANOTACOES  
add constraint FK_ANNOTACOE_DESCREVE_IMAGEM foreign key (ID_IMAGEM)  
references IMAGEM (ID_IMAGEM)  
on delete CASCADE on update restrict;
```

### **Criação de Sequencia para gerar ID automáticos**

```
Create sequence ANOTACOES_ID_SEQ;
```

```
alter table ANOTACOES alter column ID_ANNOTACOES set default  
nextval('ANOTACOES_ID_SEQ');
```

```
Create sequence EXAME_ID_SEQ;
```

```
alter table EXAME alter column ID_EXAME set default nextval('EXAME_ID_SEQ');
```

```
Create sequence IMAGEM_ID_SEQ;
```

```
alter table IMAGEM alter column ID_IMAGEM set default nextval('IMAGEM_ID_SEQ');
```

```
Create sequence PROCESSO_ID_SEQ;
```

```
alter table PROCESSO alter column ID_PROCESSO set default  
nextval('PROCESSO_ID_SEQ');
```

## Appendix C

### Code to process the image in Database

```

% Parâmetros da base de dados
dbhost = 'localhost';
dbuser = 'postgres';
dbpassword = 'postgres';
dbName = 'endomicroscopia';
jdbcString = sprintf('jdbc:postgresql://%s/%s', host, dbName);
jdbcDriver = 'org.postgresql.Driver';

% alguns parametros usado no processamento de imagem
Imagem = 'C:\Users\Aniana\Desktop\BaseDadosEx\Imagens\';
ImageCaminho = 'C:\Users\Aniana\Desktop\SegundaFase\ImageProcesing\';
TipoImagem = {'Normal' 'Inflamação' 'Lesão circunscrita'};
Clasificacao = {'Normal', 'Ligeira Inflamação ', 'Inflamação severa', ...
                'Probabilidade ligeira de Neoplasia', 'Probabilidade
                elevada de Neoplasia'};

%% Ligar a base de dados

javaaddpath('postgresql-8.4-701.jdbc4.jar');
dbConn = database(dbName, user , password, jdbcDriver, jdbcString);

%% Fazer Pesquisa na base de dados

for ii=1:length(Clasificacao)

    % Verificar se a base de dados foi acedida
    if isconnection(dbConn)
        result = get(fetch(exec(dbConn, ['SELECT imagensinal,
            id_imagem FROM imagem WHERE classificacao_global = ' ' ' ' '
            Clasificacao{ii} ' 'order by 2 asc'] )), 'Data'); % Query
    end
    close(dbConn);

    % Aceder a cada imagem
    for j=1:length(result)

        a = result{j,1};
        t = findstr(a, '%');
        a(t(length(t)-1)) = '/';
        a(t(length(t))) = '/';
        c = a(t(length(t)-1)+1:length(a));

        filename = fullfile(Imagem,c); % caminho da imagem
        im1 = imread(filename); % Ler a imagem

        %% Fazer análise da Textura

        % Obter a imagem de Gradiente
        [px,py] = gradient(double(im1),.1);

```

## APPENDIX C CODE TO PROCESS THE IMAGE IN DATABASE

```
im_gradients_Magnitude = uint8((px.^2+py.^2).^0.5);
im_gradients_Angle = atan2(py,px);

%I = im1; I = im_gradients_Magnitude; I =
im_gradients_Angle; % Fez-se a analise da simetria para as
três imagem.
Fractal = [Fractal ...]
Intensidade = [Intensidade ...]
Entropia = [Entropia ...]
GLCM = graycomatrix(I);
GLCM_Propriedade = graycoprops(GLCM,{'contrast','Energy'});
GLCM_Energy = [GLCM_E ...];
GLCM_Contrast = [GLCM_Con ...];
Momento = [Momento ...];
Asimetria = [Asimetria ...];

[pixelCount grayLevels] = imhist(I);
MediaHist = [MediaHist ...];
AsimetriaHist = [AsimetriaHist...];

%% Calcular a energia de simetria
[phaseSymC2] = phasesym(...);%Função da Simetria

%% Calcular os maximos locais
tamanho = size(phaseSymC2);
[y x] = meshgrid(1:1:tamanho(1),1:1:tamanho(2));
lMaxInd = localMaximum(...); % Função disponível
http://www.mathworks.com/matlabcentral/fileexchange/14498-
local-maxima-minima.

%% Active Contuor
m = zeros(size(I,1),size(I,2));
for i=1:length(xx)
    m(xx(i):xx(i)+3,yy(i):yy(i)+3) = 1; % Contuor inicial
end

figure(1);
subplot(1,2,1); title('Contornos');

seg = region_seg(Bw2, m, 530,im1); % Função Active Contuor
SE4 = strel('disk',23,8);
BW4 = imdilate(seg,SE4); % Operadores morfologico

%% Ellipse Fitting
s = regionprops(bwlabel(BW4), 'Area','Orientation',
'MajorAxisLength','MinorAxisLength','Centroid');

subplot(1,2,2); imshow(im1); title('Ellipse Fitting');
hold on

phi = linspace(15,2*pi,35);
cosphi = cos(phi);
sinphi = sin(phi);
for k = 1:length(s)

    xbar = s(k).Centroid(1);
    ybar = s(k).Centroid(2);

    a = (s(k).MajorAxisLength/2)*1; % Eixo maior
```

## APPENDIX C CODE TO PROCESS THE IMAGE IN DATABASE

```
b = (s(k).MinorAxisLength/2)*1; % Eixo menor

theta = pi*s(k).Orientation/180; % orientação da
ellipse

R = [ cos(theta)   sin(theta)
      -sin(theta)  cos(theta)];
xy = [a*cosphi; b*sinphi]; % Equação da ellipse
xy = R*xy;
x = (xy(1,:) + xbar);
y = (xy(2,:) + ybar);

plot(x,y,'r','LineWidth',2); % desenhar a ellipse

end

t2 = findstr(c,'/');
c(t2(length(t2))) = '_';
filename2 = fullfile(ImageCaminho, c); % local onde
vai      ser guardada a imagem processada

print('-f1','-djpeg', filename2); % guardar a imagem
processada

end

end
```

## Bibliography

- [1] Kiesslich, R., Galle, R. P., Neurath, F. M., 2008. Atlas of Endomicroscopy, Book.
- [2] <http://emedicine.medscape.com/article/179037-overview>. [18/05/2011].
- [3] <Http://www.alert-online.com/pt/news/health-portal/doenca-inflamatoria-intestinal-atinge-mais-de-14-mil-portugueses>. [18/05/2011].
- [4] <Http://www.endomicroscopy.org/> [08/02/2011].
- [5] Buchner, A.M., Gomez, V., Gill, Heckman, M.G., Shahid, M.W., Achem, S., Gill, K.R., Laith, J., Kahaleh, M. Simon, K., Picco, M., Riegert-Johnson, D., Raimondo, M., Sciemeca, D., Wolfsen, H., Woodward,T., Woodward, T. Wallace, M.B., 2009. The learning curve for in vivo probe based confocal laser endomicroscopy for prediction of colorectal neoplasia. *Clinical Endoscopy.*, 556-560.
- [6] Buchner, A.M., Shahid, M.W., Heckman, M.G., Krishna, M., Ghabril, M., Hasan, M.,Crook, J.E., Gomez, V., Raimondo, M., Woodward, T., Wolfsen, H., Wallace, M.B., 2010. Comparison of probe based confocal laser endomicroscopy with virtual chromoendoscopy for classification of colon polyps. *Gastroenterology* 138 (7), 834–842.
- [7] André, B., Vercauteren, T., Perchant, A., Wallace, M.B., Buchner, A.M., Ayache, N.,2009. Endomicroscopic image retrieval and classification using invariant visual features. In: *Proc. ISBI'09*, pp. 346–349.
- [8] André, B., Vercauteren, T., Perchant, A., Wallace, M.B., Buchner, A.M., Ayache, N., 2009. Introducing space and time in local feature-based endomicroscopic image retrieval. In: *Proceedings of the MICCAI 2009 Workshop – Medical Contentbased Retrieval for Clinical Decision (MCBR-CDS'09)*.



## BIBLIOGRAPHY

- [9] André, B., Vercauteren, T., Buchner, A.M., Wallace, M.B., Ayache, N., 2010. A smart atlas for endomicroscopy using automated video retrieval. *Jornal: Medical Image Analysis*.
- [10] Claxton, S. N., Fellers, J. T., Davidson, W. M. Laser scanning confocal microscopy, Department of Optical Microscopy and Digital Imaging, National High Magnetic Field Laboratory, The Florida State University.
- [11] Prasad, V., Semwogerere, D., Weeks, R. E., 2007. Confocal microscopy of colloids. Department of Physics, Emory University.
- [12] Hoffman, A., Goetz, M., Vieth, M., Galle, R. P., Neurath, F. M., Kiesslich, R., 2006. Confocal laser endomicroscopy: technical status and current indications. *Endoscopy* 2006; 38 (12): 1275-1283.
- [13] George, N., 2003. Confocal Microscope Systems - A Comparison of Technologies. Bioscience thecnology .Olympus America Inc.
- [14] Gheorghe, C., Lacob, R., Becheanu, G, Dumbrava, M., 2008. Confocal Endomicroscopy for in vivo Microscopic Analysis of Upper Gastrointestinal Tract Premalignant and Malignant Lesions. *Clinical Imaging*. Vol.17 No 1, 95-100.
- [15] Pentax. Confocal Imaging System, Pentax ISC-1000. Clinical Software Version 2.3, Reporter Software Version 2.3.
- [16] Kiesslich, R., Goetz, M., 2008. Confocal Endomicroscopy: In Vivo Diagnosis of Neoplastic Lesions of the Gastrointestinal Tract. *Anticancer Research* 28: 353-360.
- [17] Srinivasan, N. G., Shobha, G., 2008. Statistical Texture Analysis. *Engineering and technology*. Vol. 36. ISSN 2070-3740.
- [18] Srivastava, S., Rodríguez, J. J., Rouse, R. R., Brewer, A. M., Gmitro, F. A., 2005. Analysis of confocal microendoscope images for automatic detection of ovarian cancer.

## BIBLIOGRAPHY

[19] P. Kovési, “Symmetry and asymmetry from local phase”, Tenth Australian Joint Convergence on Artificial Intelligence, 1997.

[20] Antunes, M., 2009. Stereo from symmetry. The signal processing approach.

[21] Matlab code for calculating phase symmetry. Available from < <http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/PhaseCongruency/phasesym.m> >. [21/06/2011].

[22] Derpanis G. K., 2005. Mean Shift Clustering.

[23] Chan, F. T., Vese, A. L., 2001. Active Contours without edges. IEEE transactions on image processing, vol. 10, no. 2,

[24] Matlab code for active contour. Available from < <http://www.mathworks.com/matlabcentral/fileexchange/19567-active-contour-segmentation> >. [29/07/2011].

[25] Moussata, D., Goetz, M., Gloeckner, A., Kerner, M., Campbell, B., Hoffman, A., Biesterfeld, S., Flourie, B., Saurin, J-C., Galle, R. P., Neurath, F. M., Watson, M. J. A., Kiesslich, R., 2010. Confocal laser endomicroscopy is a new imaging modality for recognition of intramucosal bacteria in inflammatory bowel disease in vivo.

[26] Bi, Z., Manjunath, S. B. Analysis of confocal microscope images from retina detachment experiments using texture bases features. Department of Electrical and Computer Engineering and Center for Bio-Image Informatics, University of California.