UNIVERSITY OF MEDICINE AND PHARMACY CRAIOVA
PhD STUDIES SCHOOL

Ph.D. THESIS

-ABSTRACT-

THE ROLE OF COMPUTER ASSISTED DIAGNOSIS FOR THE EVALUATION OF CONFOCAL LASER ENDOMICROSCOPY IMAGES IN GASTROINTESTINAL PATHOLOGY

SCIENTIFIC COORDINATOR:
PROF. UNIV. DR. SĂFTOIU ADRIAN

PhD STUDENT:
ȘTEFĂNESCU (BADEA) DANIELA

CRAIOVA
2019

CONTENTS

I. Introduction……………………………………………………………………………………3
II. Stage of knowledge…………………………………………………………………………4
1. Inflammatory bowel disease……………………………………………………………4
2. Colo-rectal cancer………………………………………………………………………..5
3. Laser confocal endomicroscopy…………………………………………………………..5
4. Computer Assisted Diagnosis……………………………………………………………..6
III. Own contributions………………………………………………………………………..7
1. Study 1……………………………………………………………………………………7
2. Study 2……………………………………………………………………………………9
3. Study 3……………………………………………………………………………………10
IV. Discussions……………………………**….**……………………………………………11
V. Conclusions…………………………………………………………………………….12
VI. Bibliography……………………………………………………………………**..**…13

KEYWORDS: inflammatory bowel disease, colorectal cancer, laser confocal endomicroscopy, computer-aided diagnosis, neural network.

SUMMARY OF MAIN PARTIES:

I. Introduction

Inflammatory bowel diseases (IBD) represented by Crohn's disease (= CD) and ulcerative colitis (= UC) are autoimmune pathologies that affect the digestive tract, evolving with periods of remission interspersed with progression. These are non-healing diseases; several factors may be involved in the etiology such as genetics, intestinal flora, immune system, smoking, hygiene, infections or stress.

Colorectal cancer (CRC) occupies the 3rd place among the most common neoplasms worldwide, being the 4th leading cause of cancer death and is expected to double its incidence over the next 10 years. It develops slowly, in years, in most cases progressing from polyps developed in the colonic or rectal epithelium.

The etiology is multifactorial: genetics, colon microenvironment, lifestyle, family history, presence of colonic polyps, IBD.

The diagnosis of IBD and CRC are obtained using imaging techniques and are histopathologically confirmed.

Confocal laser endomicroscopy (CLE) is a novel imaging technique that contributes to the diagnosis of these pathologies. Thus, images are obtained at the cellular level of the digestive tract mucosa. The structure of the intestinal wall can be studied in vivo in real time. CLE can also be successfully applied in monitoring disease progression, identifying relapses or evaluating therapeutic efficacy.

When it comes to establishing a diagnosis, subjective factors and a limited examiner's experience may interfere. For these reasons artificial intelligence such as computer-assisted diagnosis (CAD) is more and more used for diseases diagnosis.

CAD is applied on medical imaging, this way it "learns" to recognize as normal or pathologic different images thus managing to provide a diagnosis.

The role of CAD applications is to help the examiner evaluate certain parameters from pathological images such as: type, severity, stage, progression or regression.

Another component of artificial intelligence is the artificial neural networks (ANNs) that are successfully used for medical classifications. A special type of ANN is the convolutional type used for image recognition and classification.

Studies describing the utility of CAD systems in diagnosing the pathologies of several organs such as breast, brain, lung or prostate are presented in literature.

Currently there is no study describing the applicability of artificial intelligence in diagnosing CRC or IBD on CLE images, which led me to choose as theme for doctorate the utility of computer assisted diagnosis in identifying digestive tract pathologies.

II. Stage of knowledge

1. Inflammatory bowel disease

IBD, represented especially by CD and UC, are chronic pathologies that affect the gastrointestinal tract, evolving with periods of remission or progression. These are associated with an increased risk of developing CRC (1). In CD the inflammation is transmural and can affect any region of the digestive tract from the mouth to the anus, causing ”skip lesions” (2). UC is affecting only the large intestine, inflammation being present in the mucosa (rarely the submucosa) (3).

What causes the appearance of IBD is not fully understood. An inappropriate response of the immune system (4) seems to be involved, which may or may not be triggered by the presence of microorganisms. Certain environmental factors that target genetically susceptible hosts to develop IBD are needed. (5) Some of the risk factors involved in pathogenesis are smoking in CD, (it has a protective role for UC), hygiene seems to be associated with an increased incidence of IBD, while a poor hygiene environment is a protective factor, infections and antibiotics, genetic susceptibility, stress. (6)

Clinically, patients with UC tend to have pain in the left iliac fossa, accompanied by fever, diarrhoea with mucus and sometimes blood. Those affected by CD have pain in the right iliac fossa, low appetite, diarrhoea and weight loss. The most common complication of CD is intestinal occlusion due to the thickening of the intestinal wall produced by inflammation. Another problem encountered in patients with CD is the development of malnutrition or nutritional deficiencies, caused by poor absorption of food. (7)

IBD diagnosis can be imaged guided by several techniques: computerized tomography, positron emission tomography, magnetic resonance imaging, transabdominal ultrasound, contrast enhanced ultrasound, endoscopic video capsule, but the gold standard is represented by digestive endoscopy with biopsies and histopathological examination that offers a certain diagnosis.

2.Colorectal cancer

CRC is classified as 3rd among the most common neoplasms worldwide, being the 4th leading cause of cancer death and is expected to increase by 60%, which represents 2.2 million new cases and 1.1 million deaths from cancer by 2030.

CRC develops slowly, in a few years, even decades. Most cancers develop from polyps localized in the colonic or rectal epithelium. Several factors such as genetics, epigenetics or environmental factors (as diet) are involved in the etiology. It has also been observed that intestinal flora and chronic inflammation may precede tumor development. (8)

Clinically, CRC has three phases, the asymptomatic invisible phase, followed by an asymptomatic visible phase, and then the symptomatic one, suggestive for advanced disease. Primary symptoms that are attributed to localized tumor include rectal bleeding, blood in the stool, changes in intestinal transit (diarrhea or recent constipation), abdominal pain, cramps, bloating. General symptoms, such as anorexia, significant unintended weight loss, fatigue, and anemia-related symptoms, which usually suggest an advanced stage, may also be present. The combination of rectal bleeding associated with altered transit or rectal bleeding in the absence of peri-anal symptoms are a common way of presenting the disease. (9)

The diagnosis is established by imaging techniques and histopathological confirmed.

3.Confocal laser endomicroscopy

CLE is an endoscopic technique developed to obtain images with magnification and very high resolution from the mucosa of the gastrointestinal tract. The images are obtained in vivo in real time, offering the possibility of microscopic evaluation of tissue structures at the cellular level. (10)

The operating principle is based on illuminating the tissue at a certain depth with a blue laser light (an argon-ion laser that generates a light with a wavelength of 488 nm). Due to CLE, tissue images can be obtained from both the mucosal surface and the depth of the intestinal wall, without affecting the tissues integrity. (11) The term "confocal" defines the alignment in the same focal plane of both the lighting system and the collection system. (12) The laser light is focused through a lens at a certain depth in the tissue. The reflected light is then refocused to the detection system by the same lens. Only the refocused reflected light is detected, any other light beams reflected or scattered at other geometrical angles or unfocused through the detection lens are excluded. This aspect increases the spatial resolution of CLE and allows the analysis of images at the cellular level, as well as the tissue architecture from the focal plane. The entire operation is performed during endoscopy. (13) The laser scans the area of interest, resulting a complex image. The images thus obtained are from a depth of 0-250 μm and have a field of view of 475 x 475 μm. Exogenous fluorescent substances can be administered topically and systemically to obtain confocal images. (14)

Currently, there are two types of CLE: one integrated with the endoscope, which contains a mini-scanner located in the distal portion of the video endoscope (eCLE - Pentax, Tokyo, Japan) and the other is a standalone probe that is introduced through the channel work of most endoscopes (pCLE - Cellvizio, Mauna Kea Technologies, Paris, France). (15)

CLE can be used to examine luminal structures, such as the esophagus, stomach, colon, but also ductal structures, such as biliary or pancreatic ducts. Also, CLE can be used to optimize endoscopic diagnosis, thereby reducing unnecessary resections, avoiding repeated biopsies and post-investigation follow-ups. Thus, the risks and costs associated with repeated endoscopic examinations are indirectly reduced. (16)

4.Computer-assisted diagnosis

In order to obtain accurate and correct diagnoses, lately several variants of computer-assisted diagnosis and detection applications (= CAD) have been developed. The purpose of these applications is to assist the physician in interpreting medical images, because the human eye-brain system has some limitations such as little experience, fatigue, distraction or satisfaction produced by the research work. All of these can contribute to the suboptimal use of available information. (17)

By processing the medical images, we try to modify them so that we can provide the reader as much detail as possible about the visualized structures. However, single image processing cannot reduce some limitations such as fatigue, distraction, or lack of experience. CAD systems go beyond the medical images processing, managing to provide specific information about the location of lesions or their appearance. (18)

The purpose of computer-aided detection systems is to mark during examination the regions of medical images that can provide information about any anomalies.

The purpose of computer-assisted diagnostic systems is to provide the clinician assistance in the evaluation of pathologies by marking parameters such as: type of pathology, severity, stage, progression or regression.

OWN CONTRIBUTIONS

1.Study 1

The purpose of the study implies the implementation of an automatic diagnostic algorithm meant to identify the CRC. Endomicroscopy images are processed using a CAD system which will provide information about the presence or absence of specific lesions, classifying them as normal or cancer. (19)

Materials and working method

For the study we used a batch of 1035 endomicroscopy images. All images provided from the database of the Center for Research in Gastroenterology and Hepatology Craiova, University of Medicine and Pharmacy Craiova, Romania between January 2010 and June 2012.

356 images with normal mucosa and 679 images with cancer were used for analysis. On average were used 44.5 ± 21.3 images per patient with normal mucosa, respectively 75.4 ± 59.4 images per patient with cancer aspect.

Technical data

Endomicroscopic examination of the colon was performed using a colonoscope model EC-3870 CIFK (Pentax, Tokyo, Japan). The resulting images were in gray scale having a 1024x1024 pixels resolution, being obtained at a rate of 0.8 images / second and were kept for further analysis (200-300 images per examination). Endomicroscopic images were obtained from different depths of the terminal ileum and four segments of the colon (caecum, transverse, descending and rectum). Simultaneously with the CLE examination, biopsies were taken from the same areas examined, in order to have a certain histopathological diagnosis.

Definition of the computer-assisted diagnosis parameters and the image analysis protocol

Endomicroscopy images were processed using a CAD module, a proprietary of the medical imaging system (NAVICAD) developed using Matlab programming software (Matlab, The MathWorks Inc. USA). Only images without artifacts such as intestinal peristalsis or the presence of various particles such as gas bubbles or faeces were used.

The CAD application comprises three modules:

1. the module for identifying the anatomical features realized with functions of the MathWorks image processing tools: the isocontour function based on Marching Squares and linear interpolation, and the polygeom function for calculating the perimeter of the obtained features;

2. the gray level analysis module based on the gray-level co-occurence matrix (GLCM) of the greycomatrix function;

3. the fractal analysis module.

To interpret the results obtained using the CAD system, we developed a feedforward ANN with two layers. Its role is to diagnose images as normal or cancer. ANN is based on seven parameters calculated for an image (fractal dimension, lacunarity, contrast, correlation, energy, homogeneity, number of characteristics). In the hidden layer for the recognition of the characteristics we used 100 neurons, as a balance between precision and speed. We used two neurons for the output layer, which correspond to the normal diagnosis versus cancer.

The average result for 1000 trainings, tests and validations is 87.6%, which represents the accuracy of the neural network. 100% represents no misclassification, and 0% indicates maximum misclassification.

The entire batch of sample images (both normal and pathological) were randomly divided into three sets: 725 for training, 155 for validating training efficiency (training stops when no improvement is observed), and 155 images to test diagnostic accuracy of the ANN after being trained and validated. The results of the simulation of the neural network are expressed in terms such as cross-entropy and decision accuracy error.

During the training phase, the cross entropy was 0.53 and the decision accuracy error for testing was 15.48%.

2.Study 2

The purpose of the study implies the implementation of an automatic learning algorithm that uses a convolutional neural network (CNN) that can differentiate the normal mucosa of the colon from the inflammatory one, on endomicroscopy images.

CNN is a special type of ANN used for image recognition and classification. This name is used due to the convolution layers from its architecture. CNN focuses on classifying certain features from the images used.

Material and method

The study was performed on a batch of 6205 endomicroscopy images, all coloured with shades of grey, from 54 patients with CD. The images were obtained from five areas of the digestive tract: rectum, descending colon, transverse colon, ascending colon and terminal ileum. The final diagnosis was established by histopathological evaluation. The examination was performed at the Herlev University Hospital in Copenhagen, Denmark.

Technical data

In order to demonstrate the usefulness of artificial intelligence in the diagnosis of IBD, a CNN was developed to classify endomicroscopic images as normal or with inflammation present in the colon’s mucosa. The CNN architecture contains four convolution layers, two max-pooling layers, six batch normalization layers and four dropout layers.

Experiments

i. Dataset

The dataset consists of 6205 endomicroscopy images, all coloured with shades of grey, with the initial resolution of 1024 x 1024 pixels. These images are divided into two categories: 2533 images with normal colon mucosa and 3672 images with CD.

ii. Learning and assessment

Training

All 6205 images were divided in two categories: 5081 images were used for training and 1124 for validation. The initial resolution of 1024 x 1024 pixels was reduced to 64 x 64 pixels. For the training set we used 2892 images with inflammation and 2189 images with normal mucosa. For the validation set we used 780 images with inflammation and 344 images with normal mucosa.

Results

The model described by CNN runs an epoch in about 600 seconds. The algorithm has been trained for 200 epochs, obtaining a final accuracy of 92.79% and an adaptation error of 7.21%. The adaptation error represents the difference between the output of the network and the desired result. The CNN model begins to converge after about 75 epochs for the training set and the convergence process is stable. For the validation set, the algorithm begins to converge after 180 epochs. The convergence process fluctuates slightly, but the overall tendency of loss is clearly decreasing.

The ideal situation is obtained when the training curves follow the validation ones, which means that the algorithm has a very good generalization capacity.

3.Study 3

The purpose of the study implies the implementation of a CNN meant to identify the presence of CRC at microscopic level, on images obtained using the CLE technique.

Material and method

The study was performed on a batch of 1204 endomicroscopy images obtained from patients diagnosed with colon/rectal adenocarcinoma. All images are from the database of the Center for Research in Gastroenterology and Hepatology Craiova, University of Medicine and Pharmacy Craiova, Romania between January 2010 and June 2012.

The images were divided in two categories: normal 500 and pathological (with cancer) 704. Endomicroscopic examination of the colon was performed using a colonoscope model EC-3870 CIFK (Pentax, Tokyo, Japan).

Technical data

The images obtained with CLE had an initial resolution of 1024x1024 pixels. They were divided in two sets: 984 images were used for training and 220 images for validation. The image resolution was reduced to 64x64 pixels. CNN is trained using Adam's optimization algorithm on an NVIDIA Quadro K4200 graphics card for 50 epochs (running in approximately 60 seconds per epoch). The architecture of the experiments is Intel (R) Xeon (R) CPU E5-1620 v3 @ 3.50GHz, 32 GB RAM.

Results

Following the CNN application on the images from the training set we obtained an average accuracy of 90.7%. From the images in the test set we obtained an accuracy of 75.6%. This CNN model achieves a final accuracy of 87.73%, with an adaptation error of 12.27%.

IV. Discussions

The most commonly used diagnostic method in gastrointestinal pathology - the gold standard - is digestive endoscopy with biopsy sampling. During the examination, approximately 70% of rectal cancers and 30% of colon cancers are identified. The accuracy of the examination increases proportionally to the doctor's experience. An advantage of this examination is the ability to locate the tumour. A disadvantage of endoscopy is the inability to obtain a concrete diagnosis, because the macroscopic aspect is not always conclusive. Therefore, the final diagnosis is sustained by the lesion’s histopathological evaluation. Among the disadvantages of the histopathological examination we mention the long time until a result is obtained, that can be extended up to 3 weeks, as well as the fact that it depends on the examiner's experience. (20)

CLE can determine the histopathological type of the identified lesions in vivo during the procedure, without the need for biopsy sampling, thus shortening the time required for classical histopathological methods. Once the lesion structure is identified, the endoscopist will determine whether any resection is necessary. This increases the investigation’s efficacy and reduces the costs. (13)

However, the endomicroscopic interpretation is dependent on the histology and histopathology knowledge of the examining gastroenterologist. Errors could be caused by subjectivity or fatigue, thus causing quite large variations between different examiners. Another disadvantage is the large number of images obtained that must be selected and interpreted, limiting the advantage of working in real time.

By developing an automatic computer-based diagnostic algorithm, the dependency of a well-trained specialist for optical diagnosis obtained with CLE is diminished.

Various studies have established the utility of CAD in the medical field, targeting several organs, such as the brain (21) in order to identify various neurological diseases with positive results, prostate (22) with optimistic results for identifying cancerous lesions or thyroid gland (23). Most studies focused on the utility of CAD systems in the medical imaging have supported their practical importance in identifying lesions observed on medical images, helping the physician to determine the location or the diagnosis of suspicious lesions.

Automatic diagnosis obtained from medical images is difficult to apply because of the large number of objects shapes and permutations, the image position or brightness. The association between high performance calculation and machine learning represents the future for medical images interpretation when a big number of images have to be analysed in order to obtain an accurate and efficient diagnosis. Deep learning allows the selection and extraction of specific characteristics and the construction of new ones in order to obtain a diagnosis. It also offers models on which measurements can be realised to facilitate the doctor’s work. (24)

CNN are used with machine learning applications based on image recognition due to their ability to localise features. They can be applied in the medical field for image recording, anatomical structure detection, image segmentation, microscopic image analysis or computer-assisted diagnosis.

V. Conclusions

In the first study, CAD based on fractal analysis and other parameters from endomicroscopy images can be successfully applied to characterize the structures observed in the normal mucosa and to differentiate the areas with malignancy.

In the second study we obtained results that support the utility of CAD using CNN for the diagnosis of IBD on endomicroscopy images. It can differentiate the normal mucosa of the colon from the inflammatory one, the purpose being to establish mucosal healing.

In the third study, the results obtained support the applicability of CAD for differentiating CRC images from those with normal mucosa.

All the results obtained from the three studies support the usefulness of artificial intelligence in the medical field regarding the identification of CRC and IBD on endomicroscopy images.

VI. Bibliography

1. *Inflammatory Bowel Disease.* Wehkamp, Jan, și alții. 5, February 2016, Deutsches Arzteblatt International, Vol. 113, pg. 72-82.

2. *Chronic intestinal inflammation: inflammatory bowel disease and colitis-associated colon cancer.* Deborah C. Rubin, Anisa Shaker, and Marc S. Levin. 107, 6 May 2012, Frontiers in Immunology, Vol. 3.

3. *Advanced Endoscopic Imaging Techniques for the Study of Colonic Mucosa in Patients with Inflammatory Bowel Disease.* Ştefănescu, D, et al. 1, 2016, Romanian Journal of Internal Medicine, Vol. 54, pp. 11-23.

4. *Intestinal homeostasis and its breakdown in inflammatory bowel disease.* Powrie, Kevin J. Maloy and Fiona. 16 June 2011, Nature, Vol. 474, pg. 298-306.

5. *Recent advances in inflammatory bowel disease: mucosal immune cells in intestinal inflammation .* M Zaeem Cader, Arthur Kaser. 2013, Gut, Vol. 62, pg. 1653-1664.

6. *A clinical review of recent findings in the epidemiology of inflammatory bowel disease.* Ponder, Alexis și Long, Millie D. July 2013, Clinical Epidemiology, pg. 237-247.

7. *Clinical Presentation and Disease Course of Inflammatory Bowel Disease Differs by Race in a Large Tertiary Care Hospital.* M. Anthony Sofia, David T. Rubin, Ningqi Hou, and Joel Pekow. 9, September 2014, Vol. 59, pg. 2228-2235.

8. *Dietary induction of colonic tumors in a mouse model of sporadic colon cancer.* Yang K, Kurihara N, Fan K, Newmark H, Rigas B, Bancroft L, Corner G, Livote E, Lesser M, Edelmann W, Velcich A, Lipkin M, Augenlicht L. 19, October 2008, Cancer Research, Vol. 68, pg. 7803-7810.

9. *Symptomatic presentation of early colorectal cancer.* Smith, D, și alții. 2006, Annals of The Royal College of Surgeons of England, Vol. 88, pg. 185-190.
10. *Needle-based confocal laser endomicroscopy in pancreatic cystic tumors assessment.* Stefanescu, Daniela, și alții. 4, 2015, Romanian Journal of Morphology and Embryology, Vol. 56, pg. 1263-1268.
11. *Confocal Laser Endomicroscopy: Applications in Clinical and Translational Science—A Comprehensive Review.* Goetz, Martin. 2012, International Scholarly Research Notices, Vol. 2012, p. 13.
12. *Confocal laser endomicroscopy for gastrointestinal diseases.* Kiesslich, R., Goetz, M și Neurath, M F. 3, 2008, Gastrointestinal Endoscopy Clinics of North America, Vol. 18, pg. 451-466.
13. *Confocal laser endomicroscopy,.* Chauhan, Shailendra S., și alții. 6, December 2014, Gastrointestinal Endoscopy, Vol. 80, pg. 928-938.
14. *Confocal laser endomicroscopy: technical status and current idications.* Hoffman A, Goetz M, Vieth M, Galle PR, Neurath MF, Kiesslich R. 2006, Endoscopy, Vol. 38, pg. 1275-1283.
15. *Comparison between two types of confocal laser endomicroscopy in gastrointestinal tract.* Li CQ, Zuo XL, Guo J, Zhang JY, Liu JW, Li YQ. 5, May 2015, Journal of DIgestive Diseases, Vol. 16, pg. 279-285.
16. *Clinical impact of confocal laser endomicroscopy in the management of gastrointestinal lesions with an uncertain diagnosis.* Carlos Robles-Medranda, Maria Vargas, Jesenia Ospina, Miguel Puga-Tejada, Manuel Valero, Miguel Soria, Gladys Bravo, Carlos Robles-Jara, and Hannah Pitanga Lukashok. 8, 26 August 2017, World Journal of Gastroenterology, Vol. 9, pg. 389-395.
17. *Lesion conspicuity, structured noise, and film reader error.* Kundel, L și Revesz, G. 1962, Ameican Journal of Roentgenology, Vol. 126, pg. 1233-1238.
18. *Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation.* Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, Kim L, Summers RM. 5, May 2016, IEEE Transactions on Medical Imaging, Vol. 35, pg. 1170-1181.
19. *Computer Aided Diagnosis for Confocal Laser Endomicroscopy in Advanced Colorectal Adenocarcinoma.* Daniela Ştefănescu, Costin Streba,Elena Tatiana

Cârţână,Adrian Săftoiu,Gabriel Gruionu,Lucian Gheorghe Gruionu. 5, 4 May 2016, Plos One, Vol. 11.
20. *The diagnostics of colorectal cancer.* Magdalena Świderska, Barbara Choromańska, Ewelina Dąbrowska, Emilia Konarzewska-Duchnowska, Katarzyna Choromańska, Grzegorz Szczurko, Piotr Myśliwiec, Jacek Dadan, Jerzy Robert Ładny, Krzysztof Zwierz. 1, 2014, Vol. 18, pg. 1-6.
21. *Computer-Aided Diagnosis Systems for Brain Diseases in Magnetic Resonance Images.* Hidetaka Arimura, Taiki Magome, Yasuo Yamashita and Daisuke Yamamoto. 3, 2009, Algorithms, Vol. 2, pg. 925-952.
22. *Computer-aided detection of prostate cancer in MRI.* Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. 5, May 2014, IEEE Transactions on Medical Imaging, Vol. 33, pg. 1083-1092.
23. *The Role of Computer-aided Detection and Diagnosis System in the Differential Diagnosis of Thyroid Lesions in Ultrasonography.* Chang, Tien-Chun. 4, 2015, Journal of Medical Ultrasound, Vol. 23, pg. 177-184.
24. Muhammad Imran Razzak, Saeeda Naz, Ahmad Zaib. Deep Learning for Medical Image Processing: Overview, Challenges and Future. *Lecture Notes in Computational Vision and Biomechanics .* 2017, Vol. 26, pg. 323-350.