DOCTORAL THESIS

- ABSTRACT -

COMPUTATIONAL METHODS AND MODELS IN PATHOLOGY

PhD Supervisor:
Profesor Universitar Dr. Iancu Emil PLEŞEA

PhD Student:
Mircea Sebastian ŞERBĂNESCU

2015
TABLE OF CONTENTS

Table of contents
Introduction
State of the art
Chapter I Image preprocessing techniques
Chapter II Image noise reduction techniques
Chapter III Images segmentation techniques
Chapter IV Feature extracting techniques
Chapter V Detection and classification techniques
Personal contribution
Chapter VI Computational methods and models in gross pathology. Model: aortic morphometry
Chapter VII Computational methods and models in cytopathology. Model: pleural smear pathology
Chapter VIII Computational methods and models in histopathology. Model: myocardial morphometry
Chapter IX Computational methods and models in histopathology. Model: stromal hepatic morphometry
Chapter X Computational methods and models in histopathology. Model: liver cancer - primitive or metastasis?
Chapter XI Computational methods and models in histopathology. Model: a new parameter in prostatic stromal morphometry
Chapter XII Normalization of digital histological stainined images for automatic diagnosis
Conclusions
References
Publication list

KEY WORDS

digital pathology, computational algorithm, image analysis, image quantification, image normalization, automated diagnosis
Abstract

Digital pathology, as a distinct entity, is among the newer fields in medicine. It began to be mentioned in peer reviewed publications (journals) only since 2000, although the beginnings date back to 1960s. In a book on digital pathology (Digital pathology - Sucaet Yves, Wim Waelput - SpringerBriefs in Computer Science), it shows that in 2014 there were 188 relevant articles regarding the subject, of which 164 (87%) were published after 2009.

Although the increased number of articles is impressive, it is only the tip of an iceberg. This is an entity that is comprised of several concepts, each of them representing a new research area in itself. When studying medicine only a few subjects contain the word medicine. This applies to digital pathology too: researchers are interested in defining and researching its components, rather than the entity itself.

At first glance digital pathology represents converting an optical image of classical histological slide (glass) in a digital image that can be loaded on a computer.

Digital pathology is an information medium based on images, developed using a computer, which allows the management of information generated by a virtual slide. Digital pathology is only possible through the use of digital microscopy, which involves converting a histological slide (made of glass), into a digital slide that can be displayed, manipulated and analyzed by a computer.

The emergence of digital slides opens multiple directions for usage and automated analysis, for research, such as:
1. Collection, storage and management of virtual slides.
2. Telepathology.
3. The assistance and the automation of the diagnosis allowing its objectification and its transformation from qualitative into quantitative.

Acquisition, storage and management of virtual slides represents a first step in promoting scanners and a new diagnostic method. Most of the major medical equipment manufacturers designed for pathology offer slides scanners (Olympus, Leica, Hamamatsu, 3D Histech), in addition to management software.

Telepathology is practicing remotely pathologically.
Telepathology uses telecommunications technology to facilitate the transmission for pathological images from different locations for diagnostic, educational, or research purposes. Again, major manufacturers of medical equipment have a quasi-monopoly of this area. Newly emerged software companies offer dedicated (and often more agile) solutions, such as Pathomation.

By far the largest area of digital pathology is assisting and automating the diagnosis. Regarding the two others directions: the formalization of the concepts was carried out, at present attempting their best possible implementation, while concerning the assistance and the automation of the diagnosis, everything is ongoing. The broad base of development direction is given by several factors including, without limitation, staining techniques which are difficult to standardize, and the diversity of the aspects of the analyzed images. Last but not least, I discuss the staff involved: doctors without technical knowledge in informatics and computer scientists without medical knowledge. This oftentimes results in difficult collaboration.

In this thesis I intended to trace the possible directions regarding assisting and automation of digital pathology. I present the implementation, the results and the conclusions of several projects. This is based both on software and hardware diagnostic methods. I propose both quantitative diagnostic and technical improvement in the pathology that I have developed within 4 years of doctoral training.

In the first part I reviewed the main computer digital techniques that simulate human diagnosis. Thus, making an analysis of human mode of processing the images, I’ve split the state of the art in 5 chapters, as follows: Chapter I Image preprocessing techniques, Chapter II Image noise reduction techniques, Chapter III Images segmentation techniques [Serbanescu 2012], Chapter IV Feature extracting techniques, and Chapter V Detection [Serbanescu 2013] [Serbanescu 2013b] and classification [Serbanescu, 2011], [Belciug et. All, 2013], [Belciug et. All, 2013b] techniques.

The second part, personal contribution, is divided into 7 chapters, representing independent studies where we have applied several computational methods and models to quantify various morphometric parameters or to improve histopathology image quality. These studies are completely different in background, material and methods, so they will be presented separately, as opposed to classical thesis, every
The chapter will have designated subchapters for introduction, materials and methods, results, discussion and for conclusions. In Chapter VI we quantified aortic diameters in various locations for patients with cardiac or non-cardiac death [Mirea et. All 2014], [Ancuta et. All 2013]. The resulted data was used to train a neural network capable of predicting the cause (cardiac / non-cardiac) of death according to the evaluated parameters [Serbanescu et. All 2014]. In Chapter VII we quantified cellular populations of pleural fluid smears from TB pleurisy, creating numerical profiles for the main smear classes. In Chapter VIII we quantified the degree of cardiac fibrosis in patients scheduled for cardiac transplantation and generated an average profile for the patient in need of heart transplant. In Chapter IX, using the trichromic Masson staining, we quantified the degree of fibrosis in C hepatitis patients and confirmed the Metavir score quality. In Chapter X we quantified the growth pattern of liver cancer, the data were used to train a neural network capable of indicating whether the neoplastic process is primitive or metastatic [Gheonea et. All 2013]. In Chapter XI we quantify the growth pattern in prostate cancer using our own algorithm for computing the fractal dimension [Serbanescu et. All, 2015c], [Serbanescu et. All, 2015d], [Serbanescu, 2015] also making assessments of the stromal and vascular association in different growth patterns [Stoiculescu et. All, 2012], [Plesea et. All, 2013], [Plesea et. All, 2015a], [Plesea et. All, 2015b], [Plesea et. All, 2015c]. In Chapter XII we have developed a method of normalization of histological staining for obtaining standard images, suitable automatic diagnosis [Serbanescu et. All, 2015a], [Serbanescu et. All, 2015b].

In the context of evidence-based medicine and a growing needs of diagnostic accuracy, presents work composed of seven independent studies, showing that computational algorithms can be used on one hand to numerically assist and quantifying the diagnosis and, on the other hand, to improve digital image quality needed for a accurate diagnosis.

In the first study using macroscopic images, calibrated with a marked ruler, we quantify the exact size of the aortic artery diameter at 4 levels (ring, cross, descending thoracic and descending abdominal) thus establishing two profiles of variation in diameter, one for patients with cardiovascular death and another for those with non-cardiovascular death. Next the data was used to train a neural network able to predict the cause of death. The results of correct classification
(accuracy) - around 80% - and the possibility to evaluate the same parameters, in-vivo, using ultrasonography, offers great practical potential for the method to predict vascular / non-vascular in live patients.

In the second study, we quantify populations of isolated cells in smears of pleural fluid from patients with TB pleurisy. Ignoring other possibly present cells in the smears, cells were accurately quantified and divided into 4 classes: lymphocytes, erythrocytes, polymorphonuclear and mesothelial cells. The results were used for tracing smear profiles with numerical quantification of the cell types.

In the third study, using classical histological slides, stained with hematoxylin and eosin we quantified the scope of intramuscular fibrosis and of the average cardiac muscle fiber diameter in patients with disorders that required cardiac transplant. We could thus draw a histological profile requiring a heart transplant and noticed that the process of fibrosis affects different components of the heart walls in different ways.

In the fourth study, using classical histological slides, stained with special trichromic Masson staining we quantified the amount of collagen fibers in patients with hepatic fibrosis of viral cause. We observed strong correlations between our quantitative assessment and assessment through Metavir score, confirming the quality of this score. Clinical patient's profile data shows minimal differences with gender and large differences with age in terms of the degree of fibrosis. Morphological, the presence of bridging fibrosis and increased degree of fibrosis is accompanied by an increase in total portal area, but there is a minimal increase in the amount of portal fibrosis, proving that at the portal level there is a balance between the ratio of the amount of collagen fibers and the portal area, regardless of the degree of fibrosis.

In the fifth study, using classical histological slides, stained with hematoxylin and eosin and immunohistochemicaly with CD34/CD31 staining for vessel growth, we quantified metstatic and primary liver cancer growth patterns. By computing the fractal dimension of the nuclei placement and of the blood vessels in both normal liver tissue and neoplastic liver tissue we trained a neural network capable of indicating the type of malignant neoplasia, primary or metastatic disease. The excellent results of the network have made us appreciate that the proposed model
Abstract

can be successfully used in telemedicine, medical training, or time efficiency for liver cancer diagnosis.

In the sixth study, we proposed a new algorithm for quantifying the fractal dimension. The algorithm is based on the assessment of the ratio of current volume and maximum amplitude and has the advantage of direct application to grayscale images, thus overcoming the binarization process which is a necessary step for the classical box-counting algorithm. The algorithm was tested on a public database of textures as well as on a set of prostate cancer images, labeled with Gleason patterns. In both sets of data the algorithm had better discriminatory performance than the standard algorithm.

In the seventh study, we proposed an algorithm to normalize the color of digital images with histological staining. The proposed normalization, needed for any quantitative assessment on image colour composition, is based on a solid protein solution, with constant concentration, added next to each histological product and that completes all the steps that the biological product goes through in order to obtain the final digital image. Digital pictures are corrected using this protein marker to obtain a new image with standardised colours. The method was tested both numerically and perceptual and proved superior results to a digital normalization.

Finally after we demonstrated the advantages of using various computational methods and models to assist pathologic diagnosis in macroscopy, cytopathology, histopathology and improving the histological technique for viewing / measuring purposes we conclude that this is just the beginning of a new era and in time more algorithms will be involved in pathological diagnosis and, taking in consideration anatomo-clinical correlations, in clinical prevention.

OWN REFERENCES


19. **Serbanescu Mircea Sebastian.** A new parameter with fractal behavior that can be applied directly to grayscale images. APHS2015 - Actualitati si perspective in hard si soft, Editia a VIII-a, Tibiscu, Timisoara, 2015.


