How to implement digital pathology in tissue-based diagnosis (surgical pathology)?

Klaus Kayser*,1, Stephan Borkenfeld2, Gian Kayser3

* - corresponding author: Institute of Pathology, Charite, Charite Platz 1, D-10118 Berlin, Germany, email: klaus.kayser@charite.de
1 - Institute of Pathology, Charite, Berlin
2 - International Academy of Telepathology, IAT, Heidelberg
3 - Institute of Pathology, Medical University, Freiburg

Abstract

Background: Digital pathology has proven to become a reliable tool to be applied in pathology education and experimental measurements. However, it is still waiting at the front door of routine surgical pathology or tissue-based diagnosis. Why?

Definitions: Tissue-based diagnoses includes all diagnostic tools and algorithms that serve or can serve for diagnostic purposes in general. Any medical diagnosis procedure maps a panel of “findings” to a set of discrete diseases that include recommendations for the most efficient treatment. In surgical pathology, the “findings” include data that are derived from a) image content information, b) clinical history, c) expertise of the pathologist, d) knowledge about the disease. These are transferred using a statistical decision algorithm (neural network, discriminate analysis, factor analysis, etc.) into a diagnosis.

Image content information (ICI): Several commonly not well trained pathologists assume that microscopic images or any morphology is representative for all kinds of diagnoses. They overestimate ICI which includes the information that can be derived from an image without any external knowledge. In fact, a disease can manifest with different morphologies and vice versa. Therefore, a correct diagnosis has to include all contributions of information, especially when computerized algorithms are used. Electronic communication is a useful tool to collect all available information for diagnostic purposes. Such a system should consider the different clinical impact and the level (details) of a stated diagnosis. It ranges from crude diagnosis such as cancer, infection, etc. to predictive diagnosis or even the evaluation of a disease risk (breast cancer associated genes) prior to its manifestation.

Preparations of image content information measurements: Any measurement is a communication to transfer image information to interpretation. This procedure has to be standardized by pre-analysis algorithms. They include optical issues (shading, gray
value distribution, image size, object size, etc.) and issues of interpretation (regions of interest, object features, structures, and texture. They can be used to control and improve the laboratory’s quality.

**Conclusions:** Digital pathology offers numerous tools that can be composed to an effective and appropriate diagnosis system. The tools should collect data from all contributing information sources, standardize the collection and interpretation algorithms and grade the obtained diseases in appropriate disease levels that direct the available treatments. The coordination and standardization of the essential factors seem to be a main constraint of digital pathology implementation into routine pathology.

**Key words:** Digital pathology; image content information; region of interest; image standardization; automated diagnostics

**Introduction:**

Tissue – based diagnosis which includes the macroscopic and microscopic diagnostic procedure remains the most reliable disease classification and treatment advise in man [1-5]. It is composed of several data sets comprising a) symptoms of the patient and their development (history), b) blood and serum analysis (clinical pathology), c) search for localized or generalized organ abnormalities (in vivo imaging, X-ray, radiology, ultrasound), d) structural changes within the lesion at the microscopic level (necrosis, haemorrhages, fibrosis, abnormal cells or abnormal cellular composition, foreign body deposits, external living or dead organisms), e) so – called functional data (receptors, release of macromolecules, falsified genetic memory (DNA and gene losses, inversion, enzyme releases).

The workflow of a realistic diagnostic procedure follows a time related direction and a space related sequence [6-9]. Time related direction means from the past to the presence, and space related sequence from the gross to the smallest volume, or, in other words, from the lowest to the highest magnification [6, 10, 11].

In this article we want to describe and to derive general aspects and details on how to diagnose and, if possible, on how to replace human interaction by computerized algorithms, or by information technology.

**Theory**

Definition: A disease (of a patient) is a time dependent function that disturbs the original feedback which ensures an internal labile steady state of a cell, a cellular society, an organ and finally of a human and his society [12-15].
The labile steady state requires and transforms free energy into forces that guarantee constant structures (during live time), adjust internal structures to variations of external influences (heat, movement, attacks, injuries), and regulate the amount of free energy by adjusting its import in relation to its consumption (need).

The system can be described by the thermodynamic term entropy applicable for closed systems and its counterpart entropy flow for open thermodynamic systems in general [14, 16-18].

The characterization of the parameters that induce the disturbances of the original feedback and to limit the number of their components to a small set of repair mechanisms is the final aim of any disease classification.

In medicine the applicable limited set of repair mechanisms results in \{yes – no\} decisions; for example surgery \{yes – no\}, fever reduction by aspirin \{yes – no\}, etc.

Several distinct medical disciplines come into action, often in a defined regular order: surgery (and / or) internal medicine (and / or) paediatric medicine (and / or) psychiatry (and / or) ophthalmology (and / or) dermatology, etc.

The diagnostic procedure starts with a crude estimation which of the medical disciplines might be involved and which not.

In our days all of them depend mainly upon visual information which is often transformed from other information sources such as acoustic or sensitive signals (sounds, temperature, behaviour, etc).

**Image Content Information**

The information that can be derived solely from an image without any knowledge of external additional data that might contribute to the diagnosis is called image content information (ICI).

ICI itself can be separated into three different contributes, namely objects, structures, and textures [1, 19-22]. Objects have to be defined by the observer, for example nuclei of cells by the pathologist. Their features derive from spatial magnitudes. These include basically area (volume), circumference, gray value distribution within the area, and their moments. Derived features are surface/area fraction, entropy, and relations between different color spaces (for example red, green, blue (rgb) [23-25]. The identification and the measurement of objects require a segmentation algorithm that separates the image into objects and the background. Dependent upon the image
quality, accurate segmentation might be difficult or even impossible [8, 26-35], especially at low magnification microscopic images [10, 24, 34, 36].

A Structure is the spatial arrangement between objects. Its prerequisite is the definition of a neighbourhood condition [14, 25, 37-39]. The most frequently applied algorithm has been proposed by Voronoi, its tessellation by Delaunay [22, 40, 41]. Knowledge of graph theory is essential [1, 9, 18, 22, 41, 42] as it can serve for selecting distinct or continuous features when creating a neighbourhood and its derivatives such as symmetry operations, image neighbourhood or similarity measurements [1, 9, 18, 41].

Structures derive from the definition of objects. Their contribution to the diagnostic algorithm is, however, independent from that of objects as they define a different diagnostic “parameter set” [6, 8, 9, 11, 18, 43].

The analysis of textures does not require any external information. It is a solely pixel based image analysis [6, 9, 13, 17, 44]. Any pixel is by definition an “object”. Thus, the classic algorithm approach to separate an image into a non–overlapping object and background space is not applicable [13].

**Proposed set of image primitives**

![Image primitives](image.png)

*Figure 1: Proposed set of image primitives. Two classes contribute: 1) Points and derived lines, closed curves, i.e. one dimensional and areas (two dimensional).*
Figure 2: Example of a diagnosis algorithm based upon image primitives:  
A: original image, invasive breast carcinoma, Her2-new, grade 3; DAB, *40;  B: Derived segmentation of DAB stain; C: segmented image primitives; D: computed entropy according to staining intensity.

Syntactic structure analysis can be applied to measure parameters of both structures and textures. Commonly, the coordinates of objects define the centers of gravity of structure vertices (nodes). The length of edges (distances) as well as their associates features (gray value distribution, entropies, etc.) can then easily measured [13].

Textures can be analyzed in a different manner too, the classic syntactic structure analysis as proposed by Fu et al. This algorithm searches for so called image primitives which are then composed to a complete “sentence” [13]. A set of simple image primitives is shown in <Figure 1>. Its practical evaluation and composition is shown in <Figure 2>.

The three described image analysis algorithms complement one another.

<Figure 3> exemplarily demonstrates how such an algorithm can be implemented.
Examples of algorithms to evaluate image content analysis

A: Object analysis (pleural effusion)
B: Tissue microarray stereoogy
C: Structure analysis
D: Texture analysis
E: Image primitives

**Figure 3: Examples of algorithms that can be applied for image content information measurements**

**Implementation**

Herein, we describe the implementation of microscopic image analysis only. Before doing this, we want to draw attention, that such an implementation should only be performed as an - obviously mandatory – tool of a general algorithm that takes into account underlying external clinical information as described above. For example, it is a waste of resources if one tries to identify a lung metastasis from a primary lung cancer by extremely expensive laboratory and IT efforts, when a simple look at the patient’s history tells us that this patient got operated two years ago by a stomach cancer.

The aim of image analysis should be clearly defined prior to implementation. There are additional difficulties when such an implementation focuses on general application with reproducible standards, images that are acquired in different laboratories, and the highest even possible diagnosis and disease classification (therapy advice) or when the proposed algorithm should be applied for an individual pathology team that delivers images with a continuous image quality and evaluates its own specifically biased set of diagnoses [6, 13, 25].
Image standardization is obligatory if an automated image analysis will be undertaken. It can be combined with an interactive measurement, i.e. the image has to be adjusted in colour, magnification, shading before measuring [45, 46].

Image standardization can be performed either in an objective or in a subjective manner. Objective measures include illumination, focus, gray value distribution, and adjusted magnification (area of objects / total image area) [45, 46]. Subjective standards include the adjustment of an image to the pathologist’s performance, which include color saturation, shift of white balance to the color sensitivity of the pathologist, sharpness of the object boundaries [1, 9, 47, 48].

Interactive diagnosis means that the pathologist is evaluating the image similar to his judgement of a glass slide under a conventional microscope. The viewer should allow navigation, changes of magnification, snap shots, contemporary viewing of several images (for example HE and IHC images), and acoustic commands [1, 9, 47, 48].

There exist different levels of commands when a diagnosis should be evaluated. These include commands to switch from one case to the next one, commands that are related to sampling (selecting ROIs), commands that are related to image analysis (image content information), commands that are related to external information (previous diagnoses, history, additional diagnostic procedures such as X-ray, CT, serum data, etc.), commands that transfer the obtained diagnosis to the HIS, or to the LIS that ask for additional laboratory examinations / data, and those that close the diagnostic procedure. The general scheme of a digitized pathology work flow is shown in <figure 4>.

Acoustic or even emotional commands are superior to visual commands for case transfer and diagnostics submission (image content information description) in contrast to image analysis which is the best controlled by visual commands [1, 9, 47, 48]. Most commercially available virtual slide systems are equipped with visual commands only.
Figure 4: Workflow of digital pathology in routine diagnostic surgical pathology. In principle, digital pathology adds to the work burden of the laboratory, and does not change the glass slide preparation. In reimbursement, it assists the pathologist’s work load.

Present stage

Digital pathology came into the focus of routine diagnosis at the beginning of this century [4, 19, 49-53]. Nearly all larger institutes of pathology are equipped with a virtual slide scanner at present. These scanners are in use for teaching and education as well as for glass slide storage purposes, i.e. to replace the storage of glass slides by electronic image data banks. Routine diagnostics are performed in only a few institutes of pathology. These pathologists prefer to work with both, conventional glass slides and virtual slides contemporary. M. Nap reports that he uses virtual microscopy in approximately 60% of his diagnostic cases. He does not select the organ or clinically proposed diagnosis or any other reason to apply virtual microscopy [54].

Interactive virtual microscopy seems to be in the focus of vendors, and not any automated system. The technological improvement focuses on initial image analysis systems to be applied in immunohistochemistry and on digitalization of immunofluorescent dyes. These laser equipped digital slide scanners significantly expand the application of fluorescent signalling. The unavoidable bleaching of the glass slides does not play any role because the primary images are stored electronically. In
addition, the virtual slides can be analysed in a three dimensional manner, if the glass slide is scanned at different focus levels [55-61].

Recently developed algorithms provide the pathologist with an accurate counting of fluorescent signals [60, 62, 63].

Ongoing efforts to include virtual microscopy into routine tissue – based diagnosis include a) performance of glass slide preparation in connection with feeding of virtual scanner magazines, b) implementation of acoustic information transfer as described above, and c) standardization of virtual sides [13, 20-22, 54]. The smoothing of the laboratory scanner workflow has been monitored and discussed on several pathology conferences. Indeed, the filling of magazines, handling of scanners, emptying of slide magazines and replacement of scanned glass slides include additional man powered work flow steps.

Several vendors hesitate to open their software and image formats to the public. They are afraid of competition and loss of head positions, and consider the potential advantages such as improved attraction to pathologists, improvement of the whole field of digital pathology of lower level.

These constraints are not overcome by the obviously noticed advantages of virtual microscopy. Why? The constraints apply to the workflow in the laboratory; the advantages to the pathologist’s diagnostic work. Therefore, it is difficult to convince any administration to introduce virtual microscopy. It consumes non negligible financial resources, and reimburses only quality and some efficiency of the pathologist’s diagnosis.

Discussion:

Digital pathology is one of the medical fields that are open for embedding features of our modern communication world [13, 20-22, 54]. Despite all efforts to replace the aged glass slide technology (tissue fixation, dehydration, paraffin embedding, cutting, rehydration, staining) digital pathology focuses on digitalization of glass slides (virtual images) and some image analysis, especially of IHC images [1, 25, 50, 64-67]. The advantages to handle and diagnose on virtual slides in comparison to conventional methods are obvious and not to contradict [1, 25, 50, 64-67]. Indeed, digital pathology seemed to be a promising scientific, medical and financial issue in its childhood. Viewing several different images contemporary on a screen, automated image measurements, simple image storage and retrieval, embedding into HIS and LIS as well
as easy expert consultation seemed to promise a fast and profit oriented implementation into routine tissue based diagnosis [68-72].

However, these expectations did not come true. The scanners turned out to be expensive. The hospital administration could not be convinced to invest money in the laboratories in order to improve the pathologist’s work, especially as the development of molecular biology and molecular genetics eat a lot of investment and, at the same time, reimburse the laboratory, because they open new promising markets such as predictive diagnosis or individualized therapy [24, 73-75].

In addition, the common business model of clinical pathology, which covers the main investment by consumables does not work either: Consumables do not exist in digital pathology with the only exception of storage and maintenance of digital images and/or cases. However, pathologists who are willing to leave their treasure of gold to an external person or institution are very rare to find if at all.

This situation explains the miracle: Pathology institutes cannot invest in digital pathology in contrast to radiology institutes which save a big amount when avoiding paper prints.

Shall we who are in favour of digital pathology, give in? Not at all! There is no other way than to implement digital pathology into the pathologist’s daily routine workflow. Industry has invested too much money, and wants to get its reimbursement. In addition, the concrete advantages will push the pathologists to communicate and take use of electronically available issues that are only accessible in the electronic world. It is only a question of break through!

When it will happen?

Very soon. Because new image acquisition machines are waiting before the door, easy to handle, cheap, and open for numerous small application programs (APPs). Prototypes of these systems are already under discussion on several conferences [13, 21, 22, 76]. They will at least destroy the financial argument: digital pathology is expensive!
References:


