Whole slide imaging and analysis for biomarker evaluation in digital pathology

X. Moles Lopez¹², O. Debeir², I. Salmon¹³ and C. Decaestecker¹²

¹ DIAPATH, Center for Microscopy and Molecular Imaging, Université Libre de Bruxelles, Rue Adrienne Bolland 8, B-6041 Gosselies, Belgium
² Laboratories of Image, Signal processing and Acoustics, Brussels School of Engineering, Université Libre de Bruxelles, 50, av. Franklin Roosevelt, B-1050 Brussels, Belgium
³ Department of Pathology, Erasme Hospital, Université Libre de Bruxelles, Route de Lennik 808, B-1070 Brussels, Belgium

Whole-slide imaging (WSI) is involved in a growing number of developments and applications in various fields spanning basic science, pathology and pharmaceutical research. Imaging an entire histological slide at high magnification can now be achieved in a few minutes. Therefore, this technology enables a progressive shift towards complete microscopic image digitization in pathology, i.e., digital pathology. In research, the acquisition speed featured by whole-slide scanners (WSSs) enables image acquisition for large-scale studies involving tissue-based biomarkers, such as protein expression revealed by immunohistochemistry. Image analysis tools can then be used to extract objective and quantitative features from these virtual slides to characterize these biomarkers. In the present work, we review the current challenges in WSI and describe our recent contributions to different aspects of image analysis involved in biomarker studies.

Keywords: digital pathology; tissue-based biomarker; immunohistochemistry; image analysis; image registration

1. Introduction

1.1 Current trends in pathology

Pathology is a medical specialty that addresses the four components of disease: the cause of the disease, the mechanisms of development, the morphological changes induced and the clinical manifestations. In particular, pathologists are responsible for the definitive diagnosis of disease based on the tissue samples resected from a patient. Their diagnosis relies on gross and microscopic examination of the tissue combined with the results of additional tests, such as immunohistochemistry (IHC), to evaluate protein expression in tissue samples. Proteins are the actors in the (normal or disrupted) physiological processes, and IHC is an efficient method of visualizing and locating protein expression in tissue samples (Fig. 1). For this reason, IHC is playing a growing role in surgical pathology for diagnosing various diseases and guiding personalized therapy.

Fig. 1 Examples of IHC-revealed protein expression showing different cellular locations at 10X magnification: a) nuclear staining: Ki67, b) cytoplasmic staining: EGFR, c) cytoplasmic staining: E-cadherin, d) membranous staining: HER2.

Over the last decade, pathology has entered into the digital era by benefiting from the developments of image acquisition, which have enabled the rapid conversion of glass slides into digital images for microscopic examination of histological or cytological samples (as detailed in section 3.1); and from worldwide image distribution via the Internet. This digital niche is rapidly evolving in the pathology domain, and its popularity should keep growing at the rate of the
technological improvements. As illustrated in Fig. 2, digital pathology fosters efficient communication between general and specialized pathologists. These new communication media facilitate second opinion requests and enable more precise diagnoses for difficult cases, among other advantages. Thus, patients benefit from faster and better diagnosis and treatment, which leads to reduced healthcare costs. In addition, samples from rare cancer cases can be scanned and archived in a virtual tumor bank, which would constitute a “gold mine” for teaching and research.

Currently, in addition to identifying and classifying diseases based on cell and tissue morphology, the pathologist is asked to specifically evaluate the molecular characteristics to guide the therapeutic choice for each patient. This personalized medicine requires the development of new diagnostic methods. In oncology, an increasing number of targeted therapies are made available, each in association with a companion test for selecting patients who will benefit from the specific therapy. As illustrated in section 1.2, such companion tests often require evaluating the expression of tissue-based biomarkers, such as those illustrated in Fig. 1. As a result of slide digitization, image analysis algorithms can be applied to objectively (i.e., quantitatively) characterize the expression of these biomarkers.

![Comparison between traditional and digital pathology workflows.](image)

**Fig. 2** Comparison between traditional and digital pathology workflows.

### 1.2 Companion tests and tissue-based biomarkers

A notable example of a biomarker used to guide cancer therapy is human epidermal growth factor receptor 2 (HER2), which is a prognostic and theranostic biomarker (i.e., it is able to predict disease evolution and therapy response, respectively) for patients suffering from breast cancer. Upregulation of the HER2 gene, which is accompanied by an over-expression of the corresponding protein, occurs in 10 to 30% of breast cancer cases and is associated with poor prognosis [1]. HER2 is the target of Trastuzumab (Herceptin®, Genentech), a monoclonal antibody that improves the prognosis of patients suffering from HER2-overexpressing breast cancers when combined with chemotherapy [1]. Therefore, HER2 expression must be evaluated for all new breast cancers to guide the therapeutic choice, which
requires robust and standardized testing methods. HER2 status can be evaluated by two techniques: upregulated gene expression can be evidenced by fluorescent in situ hybridization (FISH), whereas protein over-expression can be visualized by IHC (Fig. 1). Because IHC is widely used in pathology departments and inexpensive, this technique is used first to identify patients who could potentially benefit from the treatment. However, HER2 evaluation using IHC is performed by a pathologist by eye using a microscope and relies on a semi-quantitative assessment.

For each new breast cancer case, it is also recommended to evaluate the expression of estrogen receptor (ER) and progesterone receptor (PR), which are prognostic factors [2]. These proteins also predict the response to hormonal therapies such as those based on tamoxifen [2]. The expression of these hormonal receptors is also evaluated by IHC. In the literature, different scoring systems have been proposed, such as the HSCORE, which combines the percentage of stained area and discretized intensity levels [3].

Another example of the application of tissue-based biomarkers relates to neuroendocrine tumors of the digestive tract. Cell proliferation evaluation has recently been proposed as a prognostic biomarker [4]. This evaluation involves counting the mitotic figures and determining an index of proliferation using the anti-Ki67 antibody, which labels the nucleus of any cell undergoing cell division (Fig. 1a). The evaluation of Ki67 expression in three categories (<2%; >3% and <20%; >20%) is restricted to "hot-spot" regions [4]. The detection of these regions can be challenging, as detailed in section 4.2.

1.3 Challenges in tissue-based biomarker evaluation

In the majority of laboratories, pathologists perform tissue-based biomarker evaluations by eye using a microscope. Many studies have indicated intra- and inter-observer variations for the interpretation of IHC staining, such as for HER2, hormone receptors or Ki67 [5–7]. Digital pathology is playing an important role in improving reproducibility by promoting the submission of difficult cases to a specialist or a board of experts and by improving the training of younger pathologists. Given the strong clinical impact of these results, numerous efforts, such as automation and quality controls, have been made to optimize the IHC technique. Furthermore, image analysis algorithms are developed with the intention of objectively evaluating IHC biomarkers or tissue morphology to guide therapy. However, the implementation of image analysis in daily practice remains controversial [3].

Currently, an entire histological slide at high magnification can be imaged in a few minutes, which allows for image acquisition in large-scale studies involving cell- or tissue-based biomarkers for research applications. Image analysis tools can then be used to extract objective and quantitative features from these virtual slides (VSs) to characterize these biomarkers. Databases can then be established for a large sample series by merging the multiple quantitative features characterizing IHC stains and clinical and/or pharmaceutical data, depending on the context of the study. All these data are then submitted to multivariate data analysis and machine learning methods to extract information able to identify and validate new biomarkers (beneficial for diagnostic, prognostic or therapeutic procedures) and to understand pathological processes and therapy responses [8–12].

Despite the strong technological advances that have recently been made, some challenges remain regarding the image acquisition process and, more particularly, the image format, size and quality management. Fig. 3 illustrates the essential steps in the typical workflow required for VS analysis. After calibrated image acquisition (detailed in section 3), image quality assessment and sharpness evaluation are essential for ensuring the quality of subsequent analyses, making a slide review process mandatory. This time-consuming task requires the scanner operator to carefully assess the entire slide image at high magnification. Automated methods for whole slide imaging (WSI) quality assessment are therefore required to help the operator identify blurred regions in the images [13] (Fig. 3, step a).

In addition to determining specific regions of interest (ROIs), which is useful for limiting subsequent analyses and morphological characterization (Fig. 3, steps b and c), efficient image visualization methods and annotation tools are needed for biomarker research purposes. In particular, these tools play essential roles to enable experts to easily label a small set of cells in order to automate cell type (or phenotype) classification. Cell features can then be extracted and submitted to machine learning methods to infer the phenotype classes of unlabeled cells remaining on the images [14]. A similar approach can be used to recognize particular histological structures in tissue sample images. However, this latter problem appears more challenging because of the high variability of the morphological and architectural features characterizing pathological tissue samples [15].

Another challenge is the extraction of original features to better characterize the protein expression patterns indicated by IHC and, more specifically, their heterogeneous distribution across whole pathological tissue slides (Fig. 3, step d) [16]. Furthermore, extracting relevant information from the complex biological processes involved in pathologies, such as cancers, requires the simultaneous targeting and evaluation of multiple proteins in tissue slides. Such information can be extracted by imaging and registering adjacent tissue sections on which expression patterns of different proteins are evidenced (Fig. 3, steps e and f). This approach also enables the more accurate characterization of cellular heterogeneity in tissue samples.
Fig. 3  Typical WSI workflow for VS analysis. Steps (a) to (f) indicate current challenges that are addressed in the present paper.

Collectively, these challenging problems motivated us to develop efficient methods for the assessment of image sharpness and for original feature extraction and marker colocalization in the context of brightfield WSI. In the present paper, after briefly describing the typical techniques used to identify and evaluate tissue-based biomarkers (section 2), we review image-related keys and present tools to automate the tedious image quality control task (section 3). In section 4.1, current methods for IHC staining assessment are presented, along with their limitations. We then describe a method to extract more robust quantitative features that overcomes some of these limitations (section 4.2). However, analyzing each biomarker separately does not provide information about the colocalization of biomarkers being studied, which provides a more detailed characterization of disease processes and/or treatment responses. In section 5, we present available staining techniques for characterizing such colocalization and detail a method to perform a multiresolution registration of tissue slides stained for different biomarkers, which constitutes a key tool to address marker colocalization. Finally, in section 6, we summarize the general contributions of the developed methods.

### 2. Usual techniques to identify and to evaluate tissue-based biomarkers

As mentioned in the Introduction, biomarkers are used in pathology for diagnostic, prognostic and therapeutic purposes. These biomarkers can target cell or tissue morphology (e.g., structural changes in microarchitecture), as well as protein and gene expression, which can be identified in tissue samples using specific staining techniques, such as IHC and in situ hybridization (ISH) [17]. IHC is used to label proteins specifically using antigen-antibody reactions. Staining is achieved through the chromogenic reaction that occurs when an enzyme (conjugated to an antibody) reacts with its substrate (e.g., 3,3’-diaminobenzidine (DAB) to produce a dark-brown staining, see Fig. 1). In contrast, ISH is used to label a specific DNA or RNA sequence in tissue sections (i.e., in situ) using a complementary DNA or RNA probes. The probe is labeled with a fluorochrome (FISH) or a chromogen (chromogenic in situ hybridization (CISH)). This technique quantifies the number of gene copies in the cell nucleus, thereby enabling the evaluation of gene amplification or deletion.

Fluorophores or brightfield chromogens can be used to visualize gene or protein expression. Fluorescence-based labeling works better on frozen sections and has some advantages for colocalizing antigens (e.g., regarding multilabeling facility and individual labeling detection), but it also presents some disadvantages in practice, such as labeling fading, tissue autofluorescence and the cost of the image acquisition equipment. In contrast, the brightfield-based approach is easier to implement because it can be performed on formalin-fixed paraffin-embedded (FFPE) tissue samples, which are easier to store and slice and offer better tissue morphology preservation compared with frozen sections [17]. Consequently, brightfield IHC staining on histological slides from FFPE tissue sample is more appropriate for routine clinical practice and for high-throughput tissue sample processing performed in preclinical and clinical research. Brightfield IHC staining is universally adopted in pathology laboratories because it is both economical and convenient [17]. The adoption is such that the complete sample processing workflow can now be fully automated. Similarly, current CISH technology uses conventional enzymatic reactions that can be used on FFPE tissues and visualized under brightfield microscopy, thus enabling gene expression evaluation in a morphological (i.e., tissue-based) context.

The scoring of IHC tests can be time-consuming and difficult due to tumor heterogeneity, the absence of precise cutoff values and variable staining patterns. Despite these problems, this evaluation is used for most common types of cancer, and the drugs developed in this context have blockbuster status. Quantifying IHC staining patterns has thus become a crucial need in pathology practice. For this task, automated image analysis has multiple advantages, such as avoiding the effects of human subjectivity resulting from visual evaluation [18]. Additionally, in clinical research, more standardized and robust biomarker quantification methods are needed to avoid any potential bias due to tissue preparation and/or subjective evaluation. Staining image analysis also increases the range of quantitative information that can be extracted from complex IHC staining patterns. Glass slide digitization, detailed in the next section,
constitutes an essential technological step to address the need for the quantitative characterization of tissue-based biomarkers.

3. Digital shift

3.1 Whole slide imaging technology

Over the last decade, several technical solutions for WSI have been proposed [19]. These solutions can be categorized as either a motorized microscope (MM) or a whole slide scanner (WSS). The first category features all of the conventional equipment of a standard microscope (e.g., eyepieces, objective lens turret, position control, camera), except the stage is motorized and the controller of the stage can be programmed so that the stage follows a user-defined path. Using MMs, WSI is achieved by capturing and stitching the images of every field of view (FOV) of the slide. Compared with MMs, the devices of the second category (WSSs) often lack eyepieces and position control, and they usually have only one objective. This simpler configuration enables a simplified user interface and, in particular, much faster image acquisition for a complete slide compared with a MM. For high-throughput facilities, a WSS featuring a slide loader is able to scan batches of hundreds of histological slides. Although WSSs capture complete slides faster than MMs do, they are less versatile and do not allow a pathologist to freely move the stage for direct examination of the slide. However, once the scan is completed, the file containing the VS can be shared or viewed remotely using image servers [14]. VSs enable users to navigate within slides at magnifications up to 100X [20]. However, most WSSs feature a 40X magnifying lens with a resolution of 0.25 micrometers per pixel (mpp), which is satisfactory for pathologists. For a typical sample of 20 mm x 15 mm, this resolution leads to images with approximately 80,000 x 60,000 pixels [20]. Such a large image cannot be opened using a conventional, “naïve” viewer on a standard computer. Therefore, WSSs save images in specialized file formats, which are often proprietary and can be opened with a dedicated viewer provided by the manufacturers. To reduce the image file size, images are compressed using JPEG or JPEG2000 algorithms [20].

To be reproducible, the quantitative evaluation of IHC staining patterns using image analysis requires calibrating the acquisition parameters according to a reproducible and documented procedure. In [21], we detailed such a procedure for MM-based acquisition systems and observed that fulfilling these requirements is not easy, partly because few (if any) MM-based system manufacturers provide thorough instructions to perform the task correctly. Therefore, MM users have the responsibility of developing and conducting this calibration procedure. In contrast, WSSs usually provide built-in calibration routines that ensure reproducible illumination conditions. These calibration routines are automated processes that use either a calibration slide or a blank FOV on the slide. Depending on the scanner and the availability of a calibration slide, the calibration may be performed for each scanned slide, once per batch or on a regular basis (e.g., daily).

Unfortunately, the requirements for illumination parameters do not ensure a perfectly in-focus image throughout the whole tissue slide. Indeed, even when performed carefully by a skilled technician, tissue sections are not perfectly flat after being mounted onto glass slides. Therefore, both MM-based acquisition systems and WSSs provide autofocus methods. Autofocus methods with different strength and limitations are used, depending on the acquisition method. The tiling method captures the VS by acquiring high-magnification FOVs, or tiles. Some devices acquire overlapping tiles and use an image stitching method to create the VS. To increase speed, it is possible to skip empty tiles (i.e., those without tissue), which are identified using a low-resolution image of the entire slide. MM-based acquisition systems and WSSs based on tiling offer the most versatile autofocus method. Indeed, with these systems, it is possible to perform an autofocus on each FOV of the tissue slide. This feature is also referred to as "extended focus". Although this feature may be of interest for the case of highly non-planar samples (e.g., cytology), focusing consumes a considerable portion of the scanning time when acquiring a whole slide image. Therefore, in practice, users prefer a lower focusing frequency (e.g., focusing every 5 or more FOVs). In contrast, some WSSs capture the VS with a linear sensor acquiring non-overlapping bands, and fast-focusing procedures based on using a reduced number of focusing points distributed on the slide are employed. For example, the NanoZoomer 2.0 (Hamamatsu Photonics, Hamamatsu City, Japan) that we used in our work utilizes a focusing procedure that defines separate focal planes covering the tissue slide. The size and location of these focal planes and the location of the focusing points that helped determine them have a significant impact on the focusing quality. Therefore, a quality control step is required to ensure valid acquisition, which constitutes an essential step for any future use of the VSs, particularly for the quantitative analysis of biomarkers. In the next section, we present a tool that automates the sharpness assessment of VSs, which is a tedious task that dramatically reduces the throughput of scanning facilities when performed manually.

3.2 Automated sharpness assessment of virtual slides

WSSs can scan an entire tissue slide in under two minutes and enable high-throughput image acquisition (modern WSSs feature slide loaders with capacities up to 400 slides). Obtaining reproducible and in-focus images is of utmost importance for further image analysis. Recently, some WSS manufacturers have provided scanners with new
“continuous” focusing methods that enable the focusing of each FOV while maintaining scanning speed. In contrast, the previous generation of WSSs, such as the one in use in our laboratory, utilizes a pre-scanning step, during which the tissue sample is detected on the slide and focal planes along with focusing points are positioned (Fig. 4). The scanner operator must therefore verify that the tissue was detected accurately and that the focusing points positioned automatically were not misplaced on artifacts such as tissue folds, air bubbles under the coverslip, or dust. Although this method works well most of the time, we observed that approximately 25% of the slides from our routine scanning had to be rescanned.

Regardless of the WSS used, the operator must verify after scanning that the (automatic or manually assisted) focus settings were adequate to obtain a focused image everywhere on the slide. During the elaboration of our quality control step, we found that focusing problems were visible only at magnifications of 10X and above, forcing the operator to assess the sharpness of VSs at high magnifications. Due to the size and number of VS images to assess, this quality control step drastically reduces the scanning throughput. To accelerate this task and alleviate the workload of scanner operators, we developed an automated tool using a supervised classification method that is able to categorize image tiles as “sharp” or “blurred” [22]. Our method uses classification trees, which are trained to classify tiles according to the values of features characterizing the sharpness. These features are evaluated and then selected for both their ability to discriminate blurred from sharp tiles and their computation speed. We observed that two different classifiers were necessary to optimally distinguish blurred regions on H&E and IHC slides. Based on the image tile classification, our method generates a map of blurred tissue regions that indicates to the operator which regions need additional focus points to improve image quality (Fig. 4b and c). This method proved to be efficient and reduced the workload placed on operators during the sharpness assessment while augmenting the scanning throughput of the complete analysis workflow (Fig. 3) compared with the manual method [22].

Throughput could be further increased by interfacing our method with the scanner driver, as in [13], to automatically define new focusing points and scanning areas. Computation of the features used for the classification could also be accelerated using a parallel implementation. However, our first efforts to parallelize the feature computations on a graphics processing unit (GPU) were inconclusive. Indeed, the time required to transfer the large amount of data to the GPU outweighed any benefit of the parallelization. Finally, even when continuous focusing methods become widespread, a quality control step will remain necessary before publishing VSs.

4. Tissue-based biomarker analysis

4.1 Conventional biomarker expression analysis

Quantitative image analysis is often presented as a means of achieving more precise and reproducible biomarker evaluation measurements. The most common quantitative features used for staining characterization estimate the proportion of labeled cells (or labeling index (LI)) and the staining intensity, as described in [21]. To evaluate the staining intensity on an image, we use a measure that is inversely correlated to the conventional pixel intensity, i.e., a low staining intensity value corresponds to a weak (light) staining, whereas a high staining intensity value corresponds to a strong (dark) staining. Staining intensity is thus positively correlated to the conventional optical density measurement and is sometimes strictly equivalent to this value (depending on the image analysis software package used).

Conventional features often rely on segmentation of the labeled (i.e., positive) versus unlabeled (i.e., negative) tissue parts in the slide being observed. However, the spectral characteristics of IHC tissue staining, which are usually combined with counterstaining to visualize the morphological context (e.g., DAB chromogen and hematoxylin), cause difficulties with respect to quantification. Multispectral imaging can assist in this task [23], although the high cost of the associated acquisition technology often prohibits its use in more standard applications. Furthermore, this acquisition technology has not yet been implemented in WSSs. Image processing techniques, such as colorspace conversion, can be used to ease this important segmentation step. In [24], the authors proposed an efficient method to separate stains into different channels or components, a process often referred to as “color deconvolution.” This method is based on the
transformation of the original RGB image into another colorspace based on the colors of interest. In the context of IHC slide analysis the two colors of interest correspond to hematoxylin (“blue” counterstaining) and DAB (“brown” staining). A third component (corresponding to a direction that is orthogonal to both “blue” and “brown”) is also computed during color deconvolution and can be used as a quality control to evaluate the validity of the color deconvolution step. Indeed, an empty image in this component indicates that the stain separation was optimal, while a non-empty image implies that the “blue” and “brown” components should be adjusted.

Because accurate cell segmentation is difficult and highly dependent on cell density, LI computing based on automated cell counts does not provide the most accurate estimation. In contrast, the ratio of positive to reference area can provide a more accurate assessment. The term “reference area” applies to either the whole tissue area or the total cell nucleus area that are used as references for staining evaluation (application dependent) [21].

To evaluate staining intensity, commonly used measures are based on the total integrated intensity (TII), which is the sum of the staining intensity values of every pixel in positive areas. To compare staining intensities from tissue samples with different sizes, the TII can be normalized by either (i) the positive surface area or (ii) the (tissue/nucleus) reference surface area. The first normalization case gives the mean staining intensity for the positive area only, whereas the second quantifies the mean staining intensity of the entire reference area, such that negative pixels are considered to have zero staining intensity. The first feature is usually referred to as the mean intensity (MI), while the second feature is sometimes termed the “quick score” (QS) because it corresponds to the product of LI and MI [21].

Though the results obtained by image analysis are more precise and reproducible compared with manual scoring (assuming that the protocols ensure reproducible image acquisition parameters), conventional measures such as LI, MI and QS can be influenced by staining batch, which may introduce variations in staining intensity. In addition to intensity-based features (such as MI and QS), these variations impact the segmentation of positive versus negative pixels (and, consequently, LI). Segmentation is often performed using a binary thresholding procedure, for which the threshold was chosen by investigating “representative” FOVs. In a heterogeneous tumor, the threshold should be set cautiously and tested on multiple FOVs exhibiting various staining patterns because a poor FOV selection may bias segmentation. If all tissue samples involved in a study cannot undergo all of the staining steps (antigen retrieval and counterstaining) at the same time (i.e., in a single batch), a common set of tissue samples covering the distribution of the IHC staining expression patterns being analyzed should be included in the different batches for staining feature normalization. As detailed in [21], tissue microarray (TMA) technology is a useful tool for such inter-batch normalization.

Another limitation, which is discussed further in the next section, is that conventional features fail to distinguish very different expression patterns. These limitations motivated us to develop a topological approach for staining characterization that is less sensitive to staining segmentation errors (and reasonable inter-batch staining variations). Our approach aims to characterize heterogeneous IHC expression distribution for nuclear biomarkers, as described below.

4.2 Novel features for biomarker expression analysis

Conventional features for characterizing IHC expression patterns have been successfully used in research to assess the prognostic value of several tissue-based biomarkers in patient cohorts [9–12]. To do so, biomarker expressions for groups of patients with different prognostic or theranostic outcomes are characterized with these features and their values are compared. However, these features cannot provide information regarding the staining distribution pattern, which can be, e.g., diffuse or organized into different dense regions. In recent studies on glioblastoma tumors, tumor heterogeneity has been investigated by measuring the LI evaluation on tissue areas showing high densities of stained cells, or "hot-spots" (HSs) [26, 27]. In these studies, HSs were identified by pathologists using visual (and therefore "manual") examination of low-magnification images.

In a first attempt to characterize the distribution of nuclear biomarkers, we tested graph-based methods (see [25] for details) without obtaining the expected success. Thus, we developed an alternative approach that uses an unsupervised learning method to cluster positive nuclei in HSs [28]. We tested our method on Ki67-stained high-grade glioma sections. Ki67 is a nuclear biomarker that is expressed in the nuclei of proliferating cells; therefore, Ki67 HSs are representative of highly proliferating regions. While developing our method, we observed that a strict definition for what should be considered a HS was lacking, which caused high interobserver variability during the manual annotation of Ki67 HSs by pathologists and prevented the use of a supervised classification approach. Therefore, we developed a hybrid clustering method that identifies good HS candidates (which may be highly variable in their number, shape, size and density) and helps improve inter-expert agreement [28].

Our method, referred to as Seedlink, improves DBSCAN, a well-known density-based clustering method that is able to detect isolated data points to exclude them from the clustering process [29]. DBSCAN has two parameters whose values must be adapted and should fit the data characteristics in any region of the data space. This latter condition was not observed in our real-world data, which can exhibit strong spatial heterogeneity. Our method automatically sets the DBSCAN parameter values to identify a (too) large number of small and dense clusters (that we labeled “seeds”) and then aggregates them using the single-linkage hierarchical clustering method. This additional hierarchical clustering step allowed us to address the spatial heterogeneity encountered in our data.
Our method can be applied to any nuclear marker (such as the tumor suppressor protein P53) to locate densely stained regions. Identifying these regions enables the better characterization of the staining because each cluster/HS can be considered an automatically defined ROI, from which morphological features (e.g., size, internal nuclear density) can be extracted (Fig. 3, steps b-c). Importantly, these data can extend the descriptive features used for statistical analysis with diagnostic, prognostic and/or theranostic purposes. Finally, the HSs identified with our method could be used in conjunction with an image registration method (section 5.2) to analyze and characterize other biomarkers inside or around the ROIs determined.

5. Biomarker colocalization

5.1 Colocalization of IHC protein expression in tissue slides

Multiple biomarkers can be evaluated on a single tissue sample by applying the biomarker analysis workflow (Fig. 3, steps e and f) to different sections (one per marker) of the same tissue block. Quantitative measures, such as LI, QS and HS-related features, can then be obtained for each marker and used to identify and characterize subgroups of patients with different prognoses or therapeutic responses [9–11]. However, such separate characterization of the expression of different biomarkers does not provide accurate information about the coexpression or the colocalization of protein expression, i.e., the simultaneous presence of certain proteins in cells or in particular tissue regions, such as histological structures or specific groups of cells. Such information should provide a more detailed characterization of disease processes and/or treatment responses because (i) proteins rarely act alone [30] and (ii) different cells usually interact in their microenvironment by means of secreted molecules and cell surface receptors or via other protein binding strategies [31, 32]. These actors could be identified using IHC (see, e.g., [11]), but additional tissue and/or image processing is required to colocalize them and extract more accurate information on their potential interactions. In the following text, we use “colocalization” as the general term and “coexpression” (if needed) to describe when different proteins are identified in the same cell.

Most approaches enabling antigen colocalization in tissue sections rely on fluorescence labeling (e.g., [33–35]). In contrast, few methods are proposed for colocalizing antigens using brightfield IHC [36, 37]. Although brightfield IHC is undoubtedly easier than fluorescence labeling for identifying one antigen, multichromogenic methods for multiple antigen labeling on the same slide (in which each biomarker is labeled with a different color in the visible spectrum) are limited (cf. e.g., [23, 38–40]). In addition to antibody cross-reaction (which is also a problem for fluorescence labeling), critical problems related to interpretation exist due to color merging or masking (e.g., if the targeted proteins are expressed in the same cellular compartment).

An alternative to multichromogenic methods is to identify protein expression by simply repeating standard IHC on adjacent, or serial, slides. The conventional slide thickness obtained with microtomy is 3 to 5 μm, which is smaller than the mean size of human cells. Therefore, histological structures (such as glands or epithelium) are often well conserved across a few slides. However, data characterizing coexpression at the cell level cannot be reasonably achieved by analyzing protein expression in serial slides.

When processing serial slides, colocalization analysis requires more challenging image analysis procedures, such as accurate VS registration, for viewing and extracting protein colocalization information. Until now, staining colocalization was often manually and qualitatively evaluated by imaging the same histological region in different slides and then comparing the different images side-by-side [12, 37, 41]. As presented in the following section, image registration is an efficient tool for viewing and extracting protein colocalization information from adjacent IHC slides [36, 42, 43].

Recently, an interesting approach was proposed to perform successive IHC stains on the same tissue slide [44]. This “Sequential Immunoperoxidase Labeling and Erasing” (SIMPLE) method requires slide digitization after each staining to allow virtual staining superposition via image registration (as in the case of adjacent slides). After digitization, staining is erased (or washed) through an antibody elution technique, and a new cycle of staining/digitization/erasing can be performed. The recurrent digitization and elution steps prevent the complete automation of this sequential method. However, this approach circumvents the limitations encountered when using multichromogenic techniques. The essential remaining limitation relates to the elution step, which should preserve tissue and antigen epitopes for the next staining step. By correcting certain parameters and including a new elution methodology, we recently improved the original SIMPLE method and successfully utilized it to identify antigens expressed in the same cellular compartment (unpublished data).

5.2 Multiresolution registration of serial slides

As mentioned in section 3, VSs acquired with a 40X magnification have sizes of approximately 80,000 by 60,000 pixels. Thus, to analyze biomarker colocalization as accurately as possible, strategies that enable image registration at high magnification are required. To register VSs, we extended the work of [42] in a two-step approach to reach the required accuracy [51]. In the first step, low-resolution registration was applied on the 1X equivalent magnification...
images, which are 4,000 by 3,000 pixels (Fig. 5, step 1). Because these images are still considered large compared with other medical images, we used the multiresolution, or pyramidal, registration framework available in the elastix software [45]. The result of the registration at each level of the pyramid served as the starting point of the registration occurring at the next level. In the second step, the result of the low-resolution registration step was used to initialize high-resolution registrations independently performed on 20X equivalent FOVs (Fig. 5, steps 2 and 3). This latter stage was also performed on a downsampled pyramid to accelerate registration and avoid local minima.

Fig. 5  Sequence of operations for the two-step registration procedure (see main text for details).

To optimally determine the various parameters involved in our two-step registration strategy, we performed a benchmark study comparing various transformation models (rigid and affine), similarity functions (mutual information and normalized cross-correlation), input images (luminance and deconvoluted hematoxylin channel) and optimization parameters (such as the number of iterations and the number of pixels sampled to compute the similarity function). This benchmarking demonstrated that our registration strategy can be successfully applied on serial as well as “SIMPLE” VSs. Furthermore, it revealed that the combination of simple transformation models, such as a low-resolution affine transform followed by a high-resolution rigid one, is able to correct for highly non-linear deformations, such as folding, tearing or shearing, that might appear during tissue processing. This study also showed that high-resolution registrations can be efficiently optimized using a simpler similarity function, namely the normalized cross-correlation (NCC), if computed on the hematoxylin channel of the deconvoluted image.

Serial slide staining and VS registration can therefore be combined to obtain an accurate measure of biomarker colocalization at the scale of a few cells (few dozens of μm) or even at the scale of the cell (less than 10 μm) if registration is applied on SIMPLE VSs. Colocalization measurements on registered VSs can then be computed by adapting the colocalization indices used to characterize the degree of overlap between two channels in fluorescence microscopy images [46]. In the case of registered VSs, similar indices can be computed on small tiles. The tile sizes must be determined in relation to the registration accuracy. Fig. 6 illustrates a method for computing colocalization maps and indices based on the staining features (LI or QS) computed on tiles.

6. Discussion and Conclusions

In this work, we focused on the applications of digital pathology to clinical research. In this context, we proposed tools and methods to improve three key steps of the image analysis workflow, as illustrated in Fig. 3. As mentioned in section 3, WSS limitations imply that thorough quality control is required prior to any analysis. This quality control step constitutes a bottleneck and limits the throughput of the complete analysis workflow. Furthermore, this task is often considered tedious and non-rewarding by the scanner operator (Fig. 3, step a). To increase the throughput and alleviate the operator workload, we implemented an automated sharpness assessment tool to produce a low-resolution map of the tissue and show sharp and blurred regions. The scanner operator uses this map to adjust the focusing planes and points of the pre-scan step (Fig. 4). Eventually, our tool can be interfaced with WSSs to further reduce manual interventions and increase throughput [13]. In addition, regions detected as blurred can be definitively excluded from further analyses (i.e., defined as regions of “non-interest”: Fig. 3, step b), reducing processing time and improving analysis results.

Calibrated lighting conditions and sharp images are required for the valid quantification of biomarker expression. However, the conventional features used to quantify biomarker expression, which are presented in section 4.1 (LI, QS and MI), are highly sensitive to variations in the IHC-staining process and thus require inter-batch normalization to avoid any bias [21]. In contrast, features that are related to the topology (i.e., spatial distribution) of the biomarker expression avoid this sensitivity. We proposed a method to extract features that are able to describe the heterogeneity of the IHC-expression pattern (Fig. 3, step d). Our method can be applied to any nuclear biomarker to extract densely stained regions, also referred to as HSs. Number, size, density and shape descriptors of the HSs are new morphological features that can be easily extracted from these newly defined ROIs (Fig. 3, steps b-d). We tested our method on IHC staining for Ki67, a proliferation biomarker that is widely used in pathology. The prognostic and diagnostic values of these Ki67 features remain to be validated on clinical series, for which the outcome of each patient is known.
The concept of considering quantitative staining features to be biomarkers [47] can be extended to features extracted from the analysis of the colocalization of multiple biomarkers [48–50]. As described in section 5.1, multichromogenic methods, which are dedicated to evaluate biomarker colocalization in tissue samples, require advanced IHC techniques (in fluorescence or brightfield microscopy). Due to the serious limitations of the existing methods, we proposed an alternative approach that registers VSs of adjacent tissue sections labeled for different biomarkers by standard IHC (Fig. 3, step e). A systematic benchmark proved that our approach is efficient and offers a median registration accuracy of approximately 15 μm. This precision enables the evaluation of biomarker colocalization at the scale of a few cells using well-controlled and widely adopted IHC techniques. Applying our method on VSs obtained using the SIMPLE technique greatly improves the registration precision, making this combination of methods suitable for applications requiring cell-scale colocalization. Indeed, the SIMPLE technique combines a “nearly” standard monochromogenic IHC-staining procedure, a digitization step and finally an elution step that can be applied a few times to the same tissue section. Registration of the VSs acquired after each IHC-labeling step is still required. However, because the same tissue slide was imaged with different stains, our method effectively provides, in a very large majority of cases, a registration accuracy of less than 5 μm (which is approximately the size of a human cell nucleus). After registration of high-resolution FOVs, a detailed map of biomarker expression colocalization can be generated by combining measures that are computed in small tiles to characterize the expression of each biomarker (Fig. 6). Quantifying the colocalization of protein expression at specific tissue locations or compartments should enable better data mining and improved accuracy in the characterization of cellular heterogeneity in tissue samples.

Fig. 6 Example of colocalization application where the overlap coefficient is computed locally in every tile and a global overlap coefficient is computed (limited to user-specified ROIs). The checkerboard image provides a visual means to assess a valid registration and consists of the combination of tiles originating successively from one image to the other.

The tools highlighted in the present review constitute a first step toward a better integration of the processes composing the complete IHC image analysis workflow that is involved in clinical research. Indeed, the current challenges of quantitative IHC image analysis are two-fold. First, the biological questions raised at the tissue level are increasingly complex and, consequently, require more complex image processing algorithms. Second, these more complex algorithms must process an increasing amount of data for validation purposes. An integrated workflow should facilitate automatic ROI definition, staining segmentation, image registration and customizable feature extraction, all with the possibility of manually reviewing and adjusting the intermediate results to provide validated starting points for subsequent analyses. This latter aspect, which is critical to obtain the support of the medical community, is often absent in commercial solutions.

Acknowledgements The support provided by the Fonds Yvonne Boël (Brussels, Belgium) and the Télévie program of the “Fonds National de la Recherche Scientifique” (FNRS, Brussels, Belgium) is gratefully acknowledged. The CMMI is supported by the European Regional Development Fund and the Walloon Region. C.D. is a Senior Research Associate with the FNRS.

References


