Visual Analytics in Histopathology Diagnostics: a Protocol-Based Approach

A.Corvò [1], M.A.Westenberg[1], M.A.van Driel[2] , and Jarke J.van Wijk[1]

[1]Eindhoven University of Technology
[2]Philips Digital Pathology Solutions, Best, The Netherlands

Abstract

Computer-Aided-Diagnosis (CAD) systems supporting the diagnostic process are widespread in radiology. Digital Pathology is still behind in the introduction of such solutions. Several studies investigated pathologists' behavior but only a few aimed to improve the diagnostic and report process with novel applications. In this work we designed and implemented a first protocol-based CAD viewer supported by visual analytics. The system targets the optimization of the diagnostic workflow in breast cancer diagnosis by means of three image analysis features that belong to the standard grading system (Nottingham Histologic Grade). A pathologist's routine was tracked during the examination of breast cancer tissue slides and diagnostic traces were analyzed from a qualitative perspective. Accordingly, a set of generic requirements was elicited to define the design and the implementation of the CAD-Viewer. A first qualitative evaluation conducted with five pathologists shows that the interface suffices the diagnostic workflow and diminishes the manual effort. We present promising evidence of the usefulness of our CAD-viewer and opportunities for its extension and integration in clinical practice. As a conclusion, the findings demonstrate that it is feasibile to optimize the Nottingham Grading workflow and, generally, the histological diagnosis by integrating computational pathology data with visual analytics techniques.

CCS Concepts
●I.3.8 Computer Graphics → Applications;

1. Introduction

The advent and the uptake of digital pathology technology represented a vast boost in the discipline of anatomical pathology, both in research and the clinic. In the research setting, the availability of large volumes of digitized tissue slides, also known as Whole-Slide-Images (WSIs), enabled to experiment with image analysis algorithms that promise to improve the accuracy and the reproducibility of the diagnosis [LST*16, GBC*09]. At the same time, many hospitals invested in the transition from a microscope-based environment to digital platforms [SVHvD13]. The diagnostic workflow, however, has not changed substantially and the current software products still tend to mimic the current routine at the microscope [Fin14]. Whereas radiology diagnosis is empowered across many image modalities by Computer-Aided-Diagnosis (CAD) systems [HSC06, SSMD10], digital pathology lacks methodology to optmize the diagnostic process. The laboratories that adopted digital pathology see the pathologists working on standard viewers that provide different tools for measurement, annotation-taking and quantification. Some automatic tools are available to facilitate counting and quantification of histologic primitives. In general, these applications do not significantly improve pathologists' diagnostic workflow. However, premises for its improvement are currently emphasized by a wide collection of image analysis techniques specifically addressing typical histologic primitives [VPvDV14]. Besides traditional segmentation algorithms, researchers have been investigating promising techniques based on machine learning and deep learning that are boosting the field of computational pathology [FB11]. For example, Janowczyk *et al.* [JM16] present several use cases based on task specific features. The foremost aspect now becomes to integrate the output of these techniques visually to equip pathologists with interactive and smart interfaces towards the optimization of the diagnostic process.

To achieve this goal, it is important to understand the nature of the histopathology diagnostic process. In fact, diagnosis is the result of a set of complex and subjective cognitive tasks [PAF09] that lead towards the completion of standard protocols such as the *College of American Pathologists* templates [LBC*09]. Pathologists complete the diagnostic report by describing and quantifying a list of required findings and the aggressiveness of the disease. Some of these steps are error prone (*e.g.*, glands grading and mitotic counting) [OEC*16, VvDJ*16, SKM*13] and time consuming (*e.g.*, measurements). The final information on the report reflects a semi-quantitative and personal exploration that might be affected by subjectivity [AWM*17].
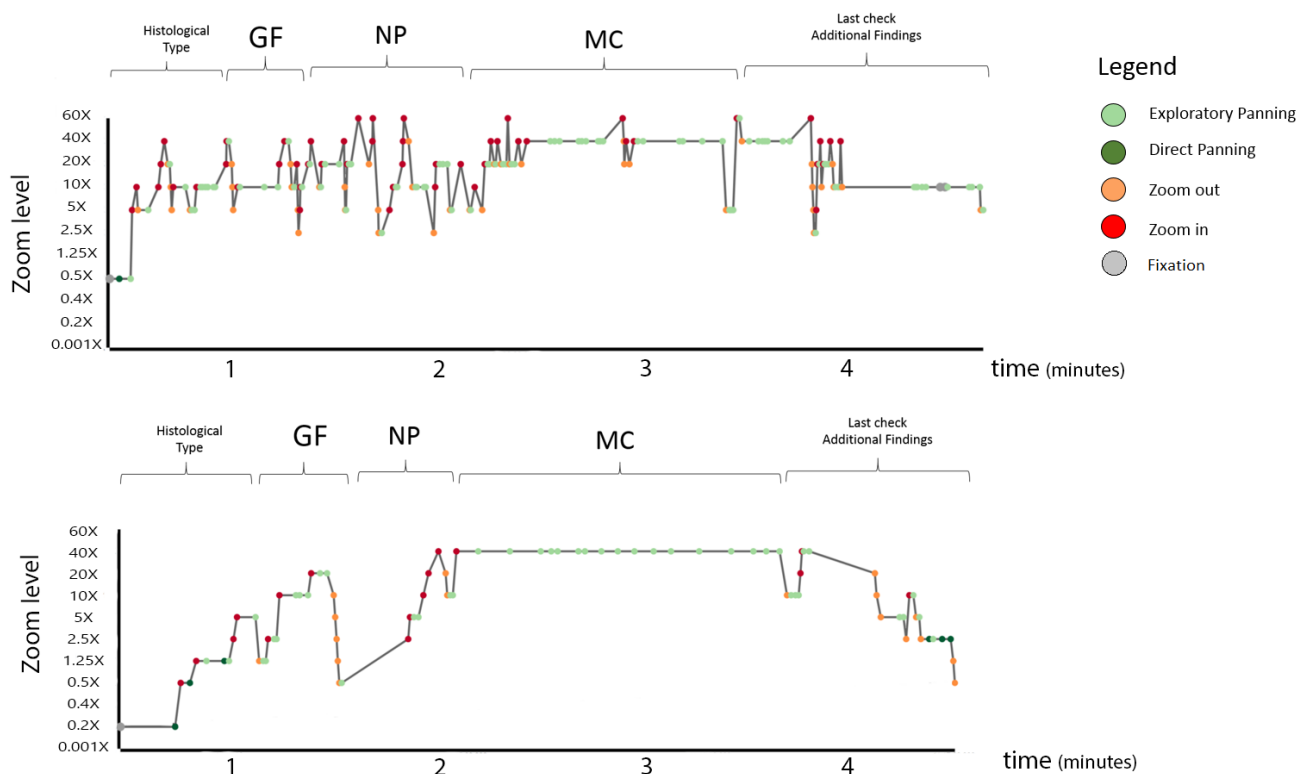
Figure 1: *Two examples of diagnostic traces. The pathologist reviewed two Invasive Ductal Carcinoma breast tissue slides according to the CAP protocol [LBC*09]. In the first part of the process, the histological type and the gland Formation (GF) are examined. Most of the traces cover the steps involved in Nuclear Pleomorphism (NP) assessment and Mitosis Counting (MC). In the first trace (top) the pathologist often uses zooming interaction to relocate himself on the tissue slide. After these first steps, he proceeds with examination at high magnifications. The first part of this case shows the pathologist using the zooming interaction more often than in the second case (bottom). In this process, the pathologists identified a region where to focus on tumor cells morphology and mitosis. As evidenced, he did not need to zoom out or move away much further than his chosen target.*

The work of Cervin *et al.* [CML16] is one of the few studies on the optimization of pathologists' workflow. Their solution diminishes the effort in note-taking and measurements towards the generation of a final report in a dynamic way. The first evaluation of the software underlines the potential of a digital solution to reduce the cognitive effort. The authors, however, do not discuss the integration of image analysis methods in the diagnostic process.

Other studies observed pathologists' behavior for different purposes and provide valuable input for the implementation and the design of CAD systems. The work of Molin *et al.* [MFMTL15] looked at slide navigation patterns among pathologists with long experience on digital platforms. In their analysis, the authors identified different exploratory strategies covering the most common behaviors in the diagnostic routine. Unfortunately, this study does not define the specific tasks performed by the users (*e.g.*, nuclear pleomorphism assessment or mitotic counting). Hence, it is difficult to determine whether the pathologists would have benefited from automation or image analysis support. Recently, Langer *et al.* [LBG*15] presented a CAD system for digital pathology for

pancreatic cancer in mice. This represents a completely automated technique, which does not require human input. In a similar way, Tashk *et al.* [THDA14] developed a novel CAD system for mitosis detection. The evaluation of this last system showed the efficiency and feasibility of these techniques with respect to pathologists' needs. However, these studies do not provide any information concerning the way the computed information should be presented to the pathologist. The need for more intuitive and visual tools is also stated by Langer *et al.*, who affirm that "a visually interpretable model based only on quantifiable clinical measures will lead to computer-aided diagnostics systems which are more likely to be incorporated in a pathologist's workflow".

In this context, we apply principles from Visual Analytics (VA) as a mean to equip the pathologists with an interpretable CAD application. Medical imaging is the field that so far saw the profusion of new VA techniques in support of radiologists and physicians [NRS*14, RBV17]. Lundström and Persson [LP11] characterized VA in the field of radiology. They identified efficiency as the most challenging aspect in radiology routine and

the need for exploratory techniques to examine medical images. Likewise, similar opinions have been expressed with the advent of digital pathology. The workload has increased in the last years and will maintain this trend [Wor14]. In this scenario, efficiency is on the priority list in laboratories. In addition, whole-slide-images are large medical images that require different interaction strategies from common radiology images. Therefore, further methods have still to be disclosed to apply visual analytics to digital pathology. Thus far, little literature has been presented on the topic. A valuable viewpoint is given by Fine [Fin14], who suggests an interesting design and ideas towards an optimized workflow wherein triaged regions of interest (ROIs) are presented to the pathologist in an automated way. Moreover, the author proposes the introduction of pre-existing templates able to capture the selected ROIs and to quantify relevant data without manual entry. These ideas, however, did not lead to a real solution for clinical practice.

Our goal is to fill the gap between histopathology diagnostic process and image analysis integration. As a conclusion of this study, our contribution is an actual implementation of a CAD system for histopathology diagnosis.

2. Clinical Background

In this paper, we target the histopathology examination of Invasive Ductal Carcinoma (IDC) of breast cancer specimens. IDC is the most common type of breast cancer with about 80% of all diagnosed cases [AHZ*13]. Despite the high incidence of breast cancer, discordance among pathologists on final diagnosis still presents large variability [ELC*15]. The standard diagnostic workflow of a pathologist consists of two main steps: review of the tissue slides and reporting. Tissue slides are typically processed by different stainings to intensify the visibility of specific biological phenomena. For instance, a mixture of Hematoxylin and Eosin (H&E) is used to highlight cellular structure and it is often considered the gold standard [Cha14]. Pathologists examine H&E breast tissue slides according to the Nottingham Grading (NG) system [RESL*08]. This grading system is part of the CAP protocol and is currently used in clinical practice. NG classification consists of three histologic primitives, which must be examined by pathologists: Gland formation (GF), nuclear pleomorphism (NP) and mitotic counting (MC). Each of these primitives is scored from 1 to 3 and each score is added to give a total score in a range between 3-9. Despite the clear importance of the NG for patients' diagnosis and prognosis [RRFB*10], this kind of examination lacks reproducibility as well as inter- and intra-observer concordance [EE02].

3. Task analysis

In this scenario, our final goal was to design and develop a visual analytics CAD system by using computational pathology data. Along this study, we collaborated with a specialist in breast cancer diagnosis. First, we started collecting requirements from the literature. Next, we implemented our tracking component to observe the behavior of the pathologist during breast cancer examination. The tracked interactions are used to generate a diagnostic trace that is the sequence of interactive steps conducted by the pathologist on the viewer. This concept is already presented in our previous work

PathoVA [CvDW17]. We let the pathologist familiarize with the viewer and we recorded a series of diagnostic traces on three WSIs. Next, we identified the downsides in the process. Subsequently, we analyzed the behavior for each diagnostic task and we studied how image analysis data can aid the pathologist to diagnose.

3.1. Tracking component

We implemented a tracking component in the spirit of Molin *et al.* [MFMTL15]. This component captures the main interactions on the viewer. For our specific task analysis we adapted the specificity of the tracked interactions that Molin *et al.* named *navlets*. Different from their work, we differentiate only two types of *panning actions*: *exploratory panning* and *direct panning*. The first identifies interactions in the micrometers range, while the second captures the ones in the millimeters range. Beside these actions, we tracked *zooming* events and the *fixation* navlet when the pathologist keeps the view static for more than two seconds. In addition, we also record the annotation actions and measurements. This helped to characterize the workflow of the pathologist. The results are visualized as diagnostic traces (Fig. 1) wherein the interactions are displayed with different dots labeling the corresponding *navlet*.

After tracking our user on three WSIs, we discussed the diagnostic process on a dedicated view. We visualized the traces on a customized line chart (Fig. 1). Furthermore, we extrapolated the stages and the respective pitfalls of these workflows. Our study is based on the analytical thinking process used in design studies and extensively discussed by Brehmer *et al.* [BM13]. The method used is based on the typology terms *why, how* and *what*. We used this abstraction to define the design and user's interactions of our CAD application. The typological description of the targeted tasks is shown in Fig. 2 and described in detail in the next section. We summarize the tasks involving the NG primitives in Fig. 2. In more detail, we can describe these processes as follows.

Gland formation. Generally, pathologists examine the organization of glands at low magnifications (2x - 10x). Usually, *zooming* is only required to look at the edges and at the lumina of such structures to assess the entity of tumor growth. Thus, the quantification of this factor is performed at the early stage (Fig. 1). However, the discrimination between normal and abnormal glands might require more interactions at higher magnification levels and on larger regions.

Nuclear morphometry. In an IDC review, the pathologist identifies a large cancerous region and zooms at higher magnification levels (20x - 40x). At this magnification, nuclei can be judged for their morphology and appearance. Usually, a few glances and observations are enough to define a final score. At this stage, a quick strategy is to compare the size of tumor cells with the one of red blood cells or normal epithelium cells. Nevertheless, this task produces a semi-quantitative evaluation. The final score can be hampered by illusions of size and brightness of the cells and their surrounding context [AWM*17]. Computational pathology can aid diagnosis to become more quantitative and to reduce the interobserver variation.

Mitotic activity. The last task concerns assessing tumor proliferation. The protocol requires the pathologist to *identify* the mitotic cells in a certain field of view and *count* them at each 40x
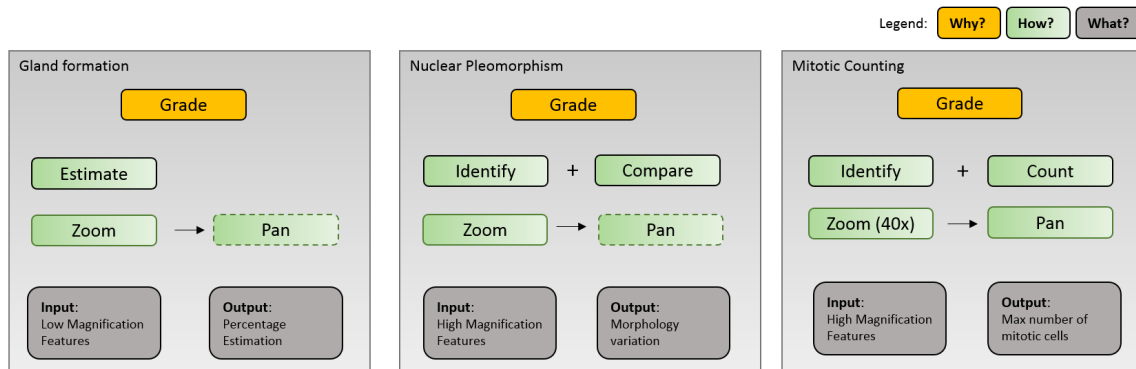
Figure 2: *A typology description of the three workflows for the NG primitives according to the why, how and what spans. Nuclear Pleomorphism and Mitotic Counting require the pathologist to zoom at the highest magnification (up to 40x) levels and perform several panning actions. The panning interaction is not always necessary for GF and NP. One or two views are enough in standard cases. The black border box represents a cognitive actions while the green border box represents the required interaction to feed the respective cognitive processes. Dashed borders indicate an optional interaction (not always required).*

magnification (approx. 10 non overlapping regions are required). In this stage, *panning* is the main interaction. Besides the high number of interactions, this task is time consuming and error prone [VvDJ*16] as visible in Fig. 1. This task took almost 40 per cent of the diagnostic time and it requires many sequential exploratory interactions.

3.2. Image analysis data

We gathered data from image analysis techniques for the detection of tubules, nuclei and mitosis counts. Regarding nuclei detection, nuclei objects are associated with several attributes such as area, short axis, perimeter and median Hematoxylin value. This information can be used to characterize nuclear pleomorphism. For tubule detection, the method described in Chen *et al.* [CQY*17] has been adopted. Despite the difficulty in discriminating normal tubules from abnormal tubules, in this work we use the results to illustrate our concepts. For instance, we could identify tubules filled by a high number of tumor cells, which leads to cancer invasion, and use it as a proof of concept in our design. This suffices for our purpose of providing useful features for the GF task.

To obtain mitosis counts, we used the detection technique described in Veta *et al.* [VvDJ*16]. The output of the model consists of the x and y coordinates of the identified mitotic cells and an associated *F-score*. This measure indicates the classification accuracy of the algorithm. It considers *recall* and *precision* of the test. *Precision* is the ratio of correctly predicted positive observations to the total predicted positive observations. In our context, this measure answers the question: "among all the detected mitotic cells, how many are actually true positives?" The *recall* is the ratio of correctly predicted positive observations to all observations in the actual class. It quantifies, therefore, the number of mitotic cells that have been detected among all the mitotic cells on the tissue slide. We consider the optimal threshold of 0.67 given by the *F2-score* of the mitotic

counts algorithm. This measure places a higher weight on recall. We integrate this information in our application to indicate to the user the detection criteria of the algorithm.

4. Requirements

We generated a list of requirements based on the conducted task analysis and on the computational pathology features at our disposal. These requirements represented the pillars for the design and implementation of our CAD-Viewer. In a broader scope, the following prerequisites intend to be a set of general concepts for CAD-implementation on digital pathology interfaces. The requirements can be distinguished in two main areas: the visual characteristics of the *histologic primitives* and the way the pathologists look at them. We refer to the image examination steps conducted on the viewer as the *diagnostic work* of the pathologist.

4.1. Histologic primitives (P): visual aspects

Contrary to radiology images, WSIs are high resolution images structured by different magnification levels. This characteristic adds a further challenge to give visual analytics support. We list the principal aspects to be considered when visualizing computed objects on digital tissue slides.

P1. Size. Features can be distinguished as low magnification features or high magnification features. The first ones such as tumor regions or luminar structures are visible and recognizable at low magnification. High magnification features can be identified in micro-features like tumor cells, lymphocytes and stromal cells that comprise the micro-environment of the tissue. These differences in size need to be converged in different visual representations.

P2. Function. Different from radiology, WSIs do not present an
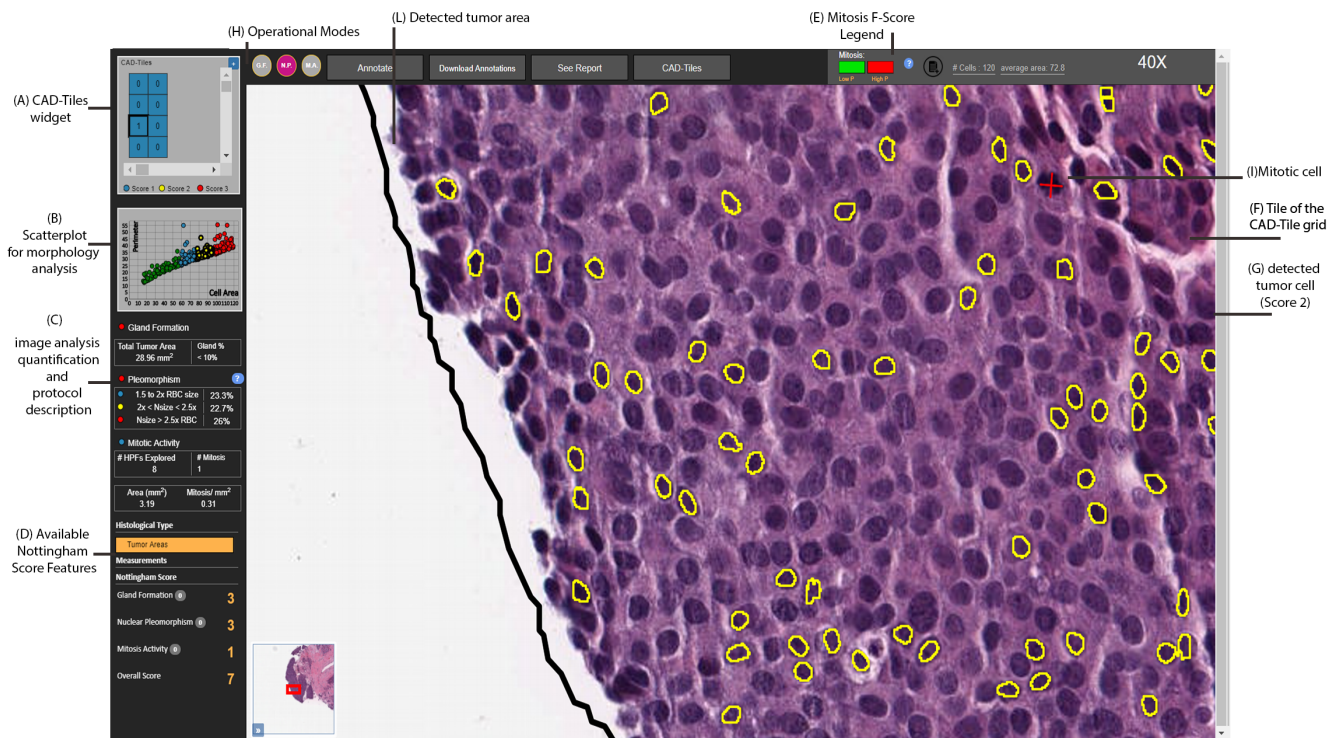
Figure 3: *Overview of our tool. On the left side, an interactive report enables the pathologist to ask for CAD support. On the right, a standard viewer is shown enhanced with the CAD-Tiles widget. Depending on the selected mode, the pathologist can get insight into nuclear morphology or mitotic activity at 40x magnification FOV. Zooming and panning are automated and triggered respectively by single-click actions and pressing arrow keys. In this case, the CAD-Tiles is generated by clicking on the black detected tumor area (L). Score 1 tumor cells (G) are displaying by hovering the cell cluster in the scatterplot (B).*

anatomical reference. However, pathologists refer to low magnification features to build a *map* in their mind and to define boundaries of the tissue architecture. Therefore, these features are functional to the exploration of high magnification elements.

P3. Color. As Pena *et al.* [PAF09] observe, the identification of a specific finding is simplified by the recognition of a set of colors and textures that corresponds to a specific biological event. Staining colors and textures are therefore important visible features to be preserved. Consistency of colors is also an ingredient to be considered when we plan to apply overlays of image analysis features onto tissue slides.

4.2. Diagnostic work (DW)

By observing the diagnostic work of the pathologist during grading we extracted the following requirements.

DW1. Whole picture. The whole slide view is used by the pathologist to formulate the first hypothesis and gradually characterize the case. Altering the aspect of this view or hiding it in the early stage can be perceived as a barrier to a correct diagnosis.

DW2. Interaction. During the examination, pathologists typically need to zoom in at 20x and 40x magnification to analyze cell details. In case of the NG, the protocol requires pathologists to

look at cell morphometry and count mitotic cells. In these cases, it is necessary to zoom in and manipulate the image at the highest magnification levels. Often, the pathologist pans over the WSIs to collect several fields of view and compare similar features.

DW3. Report-centric. The reviewing strategy is optimized towards the completion of the final report. Therefore, we consider the report as a central aspect to be integrated in a CAD viewer.

DW4. Automation. Quantification, comparison and memorization need to be facilitated. Also, the navigation efforts that the pathologist still makes can be mitigated by digital automation. Pathologist's main role is to bring in his domain knowledge [Fin14] to make a final decision, and other aspects should be taken care automatically.

As a result of our analysis, we decided to steer our implementation towards a protocol-based application. To optimize the process, we focus on the NG primitives and to the reporting needs. We only consider the case of H&E examination on Invasive Breast Cancer grading. We take into account a synoptic report (structured checklists document) comprising histological type, measurements and the NG. With these requirements in mind, we designed and implemented a CAD viewer in collaboration with digital pathology experts and an expert pathologist. In the next sections we explain the main components of the interface, the communication between

1.5X
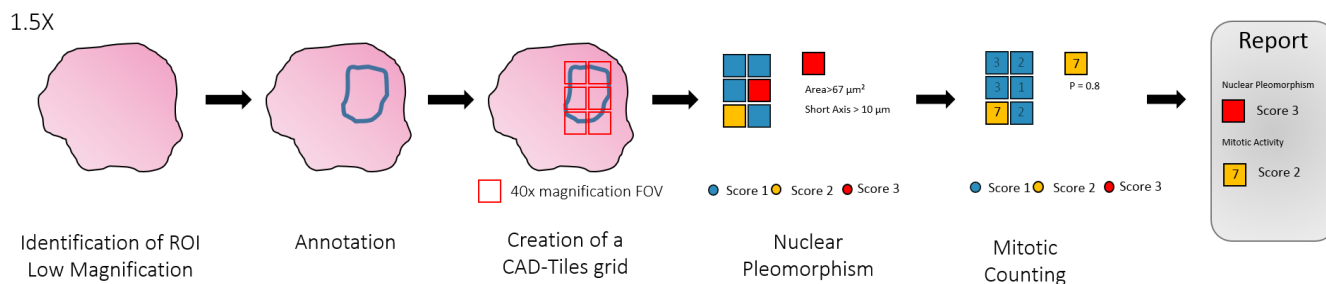
Figure 4: *The CAD-Tiles widget. The pathologist identifies one region on the tissue slide. A grid is automatically built within the extent of the annotation. Each cell of this grid is a 40x field of view (FOV). In the nuclear pleomorphism mode, each cell is assigned to a score on basis of the morphology characteristics of the tumor cells. In mitosis counting mode, the CAD system gives an estimation of the number of mitotic counts for each FOV. The FOVs are automatically attached to the report.*

the report template and the viewer and the supported interactions. Moreover, we introduce the *CAD-Tiles* tool for analyzing the NG micro-environment at low magnification levels by using image ana-lyisis data described in the following section.

5. CAD-Viewer: design and implementation

The interface of our application consists of two main elements: a standard viewer and an interactive sidebar for protocol-based vi-sual analytics. The sidebar on the left (Fig. 3) consists of an area dedicated to CAD support for the grading. The bottom part (Fig. 3 (D)) shows the updated *scores* for the current diagnosis. Each field is a clickable entity that activates the image analysis layer of the de-tected histologic primitive (Fig. 3 (C)). The report and the viewer area are fully connected. The top part of the sidebar (Fig. 3 (A)) yields tools for the NG. This widget forms the core of our appli-cation. We named it *CAD-Tiles*, a tool for analysis of High Power Fields (*e.g* 40x) at low magnification levels.
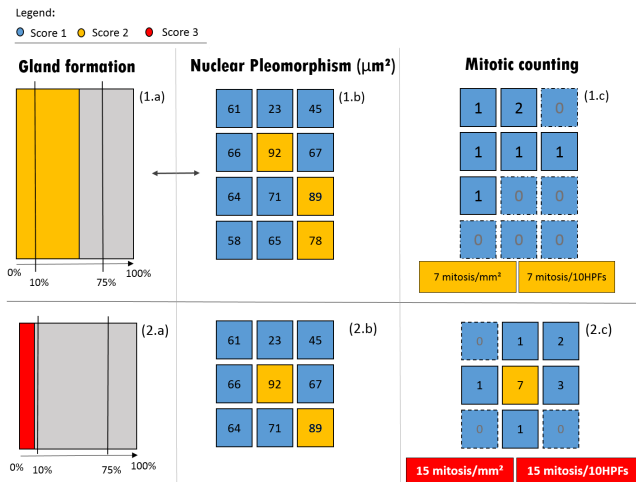
5.1. CAD-Tiles

CAD-Tiles was designed and implemented to serve four main pur-poses: view navigation, analysis, visualization of results and report-ing. The main advantage of this tool is to provide further visual in-sights of the image analysis primitives without cluttering the entire picture (**DW1**). Moreover, *CAD-Tiles* aims to reduce the panning and zooming required for the assessment of features at high mag-nification (**DW2**). Also, it gives the pathologist an automated and interactive way (**DW4**) to look at quantification of the detected fea-tures in strong connection with the report (**DW3**).

Generation. The generation of the *CAD-Tiles* grid and its external widget works as follows. First the pathologist identifies a region to examine as illustrated in Fig. 4. Next, he draws a polygonal annotation on the tissue slide. At this point, the CAD system au-tomatically produces a grid divided by tiles of 40x field of views. We restrained the tile size at the equivalent of a 40x magnifica-tion that represents the High Power field that pathologists use to count mitotic cells during standard breast cancer diagnosis.

Alternatively, the pathologist can click on a tumor area, upon which a polygon surrounding the identified area is automatically visualized. The tumor area polygon In this case the *CAD-Tiles* are created over the pre-defined region (Fig. 3 (F)). When the grid is composed on the viewer, an external view is updated on the left sidebar as shown in Fig. 3 (A). This serves as a map for the pathologist once he zooms to the high magnification level. The tool and its functionalities are designed to avoid cluttering of features on the tissue slide and disruption of the diagnostic workflow. Therefore, the grid serves as a tool to display the char-acteristics of high magnification features at middle range zoom level (**P1**). On the viewer, the *CAD-Tiles* grid is computed as follows. We extract the extent of the annotation *A*. We calculate the width *w* and the height *h* of a 40x magnification FOV. Each one of these views is a tile within our widget. Then, we divide the extent of *A* in *n* tiles of width *w* and height *h*. In the final visualization, each tile is outdistanced from its adjacent tiles by some micrometers. This distance is added to avoid cells to be considered in two different tiles.

Navigation. One of the main goals in the design process was to provide strong integration between image analysis data and the viewer area along with protocol observance. Like standard view-ers we support zooming, panning and measurements tools. Cus-tomized navigation interactions are given through the *CAD-Tiles* grid that supplies image analysis aid. Once a *CAD-Tiles* grid is generated from a manual or computed (*e.g.*, tumor area) anno-tation, the pathologist can interact with each tile. A click on a tile will automatically zoom at 40x magnification. From that moment, by using a left or right arrow key the view is automati-cally switched to the corresponding tile of the grid. This automa-tion has been specifically conceived to decrease the effort during the examination of vast regions for mitosis counting purposes or morphology assessment. The user can also use the *CAD-Tiles* widget on the left side (Fig. 3 (A)) to navigate and jump to other regions. A tile of the external *CAD-Tiles* widget is interactively connected to the grid visualization on the WSI. A click on one of the squares leads to a 40x zooming on the viewer.

Visualization. The visualization of the *CAD-Tiles* widget and the

Figure 5: *The CAD-Tiles widget with three views for two different examples. Each view encodes information for a NG primitive. Gland formation is displayed in 1.a and 2.a. The percentage of gland formation within the selected (tumor) area is shown in yellow in 1.a, corresponding to Score 2 and in red if (Score 3). In the NP view (1.b and 2.b), the average area of the tumor cells in each tile is displayed. The respective score is color-encoded. In the mitosis counting (1.c and 2.c) each tile shows the mitosis counts in the region. Dashed tiles do not contribute to the total. The final score is displayed at the bottom and shows the protocol score.*

In this way, we distinguish the two different functions of low magnification features and high magnification features **(P2)**.

Grade computation. The external *CAD-Tiles* widget supports three views, each one corresponding to the NG primitives: gland formation, nuclear pleomorphism and mitosis activity (Fig. 5). The high-level information displayed in each tile is computed and visualized with colors as described above and numbers characterizing the three histologic primitives.

The gland formation view (Fig. 5, 1.a - 2.a) encodes the extent of the annotation or the selected tumor region. The box is divided in three areas wherein the markers for the CAP thresholds of 10% and a 75% gland formation are indicated. In Fig. 5 an example shows a region wherein a *Score 2* is given on a computed gland formation in between the two thresholds.

By selecting the NP mode, each 40x FOV of the *CAD-Tiles* grid is graded with a *Score from 1 to 3* according to tumor cells morphology. Here we use a classification based on the average area of the tumor cells in that FOV. We used the thresholds provided by the CAP protocol that describes the score to be assigned to tumor cells morphology with respect to a normal blood cell aspect. We considered the diameter of a normal red blood cell roughly 8 micrometers. Hence, we calculated grading thresholds respectively for mean area and the perimeter of tumor cells. We also added the information regarding the values for the shortest axis of the tumor cells because this was demonstrated to have prognostic value for breast cancer outcome in IDC [PKC98]. Given these factors, we assign *Score 1* to tumor cells 1.5 to 2 times
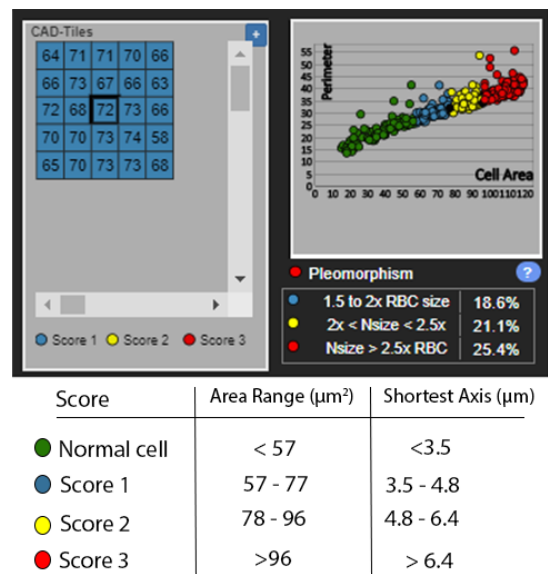
generated grid on the WSI is designed to support the three histologic primitives of NG in three different views (Fig. 5). Once the pathologist draws an annotation, the features contained in the drawn area are displayed accordingly to the selected mode (Fig. 3 (H)). The pathologist can select the operation mode of *CAD-Tiles* by clicking on the dedicated buttons (Fig. 3) or by selecting the respective layer from the left sidebar. The CAD engine computes a score for each tile on the basis of the primitives characteristics. High level *CAD-Tiles* information is encoded with four different colors. Green represents benign condition, blue corresponds to a *Score 1* in the NG, yellow and red indicate respectively *Score 2* and *Score 3*. This encoding is consistent along the three different views generated in the *CAD-Tiles* widget. Accordingly, the colors of tile borders follow this scale both on the viewer and the complementary widget. We similarly visualize the borders of detected nuclei within the annotation region as in Fig. 3 where *Score 2* tumor cells are displayed. The stroke of the tumor cells encodes the computed score. We do not use any overlay to preserve the staining color of the cell **(P3)**. Detected mitotic cells are indicated by a cross, which color indicates the F-score associated to the detection. The legend is displayed in the top right corner of the viewer bar (Fig. 3). The *CAD-Tiles* annotation is intended to be used at middle range magnification (e.g. 5X - 10X). Therefore, we decided to display inside the grid only high magnification features (tumor cells and mitotic cells). The low magnification features, such as tubules and tissue boundaries, can be requested on demand by the pathologist clicking on the side menu (Fig. 3 (D)).

Figure 6: *Focus on the CAD-Tiles widget. Here, the NP mode is enabled. In the scatterplot (right) each dot represents a nucleus. The color identifies the corresponding score. However, all the FOVs (left) are graded as score 1 according to the threshold. Here the pathologist can inspect the CAD quantification and understand the reason for its classification. In case of a different score between area and shortest axis, the priority is given to the first.*

larger than normal red blood cells, *Score 2* for those in between 2 and 2.5 times and *Score 3* if larger than 2.5 times. We indicate this distinction according to the CAP definition. Regarding the shortest axis we set thresholds of 3.5, 4.8 and 6.4 micrometers respectively for *Score 1 to 3*. The exact thresholds are given in Fig. 6. NP quantification is also supported by a scatterplot (Fig. 6) wherein we show the distribution of the detected cells on two axes. The basic view plots cell area on the x-axis and cell perimeter on the y-axis. Each nucleus is represented with a dot of which the color encodes the grade. By hovering one of the dots, all the tumor cells matching that grade are highlighted on the slide FOV (Fig. 3 (G)). By means of this view the pathologist gets insight into how the CAD determined the grade. This serves as an element designed for trust [JCvO15] where the pathologist can assess the quality of the detection. At last, in the mitotic counting mode, our CAD determines the number of mitotic events in each FOV. The detected number is converted to a score as it would contribute to the final score by definition in the standard guidelines [LBC*09]. Detected mitosis are visualized only at tile magnification level in order to avoid cluttering (**DW1**) during other diagnostic steps The total count is displayed at the bottom of the grid as shown in Fig. 3.

Reporting. We support prompt integration of any *CAD-tile* to the report. When the *CAD-Tiles* grid is generated and grades computed for each tile, our application automatically includes relevant regions to the NG section of an interactive report. The regions with the highest scores for the three histologic primitives are available on the CAP template. A click on a thumbnail directly shifts the viewer area to the respective FOV. Alternatively, interesting tiles for gland formation, nuclear plemorphism and mitotic regions can be added manually by means of a dedicated button. Every thumbnail can be edited and removed from the report. We also support a specific interaction to update the mitotic counting. The pathologist can annotate the cells in mitosis directly by typing a number on the keyboard. This number is displayed on the viewer area and automatically updated on the mitotic score. The application automatically includes conducted measurements, which can also be edited and labeled manually. The other fields of the template are provided as a standard form template with text boxes and checklists.

5.2. Implementation

Our tool was implemented as a server-client framework. The visual interface has been implemented in OpenLayers [Haz17] and D3.js [BVJ11]. The back-end comprises a GeoServer [geo17] to handle the digital slide images and the results from image analysis. The discussed WSIs were downloaded from The Cancer Digital Slide Archive (CDSA) [CVS*13].

6. Evaluation

We conducted a qualitative evaluation of our tool with five pathologists. We firstly introduced them to the concept of visual analytics and its purpose. Next, we presented the application in the use case of invasive breast carcinoma and the functionalities. Two pathologists actively interacted with the tool and examine a given WSI

Theme	Question	Agreement (0-4)
V.A.	I see the need of more visual analytics in digital diagnostic	4 (0.0)
FV.1	The necessary features for breast cancer grading are displayed in an intuitive way	3.3 (0.4)
FV.2	High magnification features information are clearly visible at middle-range magnification	3 (0.0)
FV.3	The grading process would not be disturbed by this feature visualization	3.2 (0.4)
WO.1	The Nottingham Grading and the reporting can be increasingly speeded up	2.8 (1.5)
WO.2	The CAD-Tiles tool seems easy to use	3.3 (0.4)
WO.3	Nuclear Grading and Mitosis counting are fully supported	3.3 (0.4)
T.1	I could trust this system for nuclear grading	3.5 (0.5)
T.2	I could trust this system for mitotic counting	3.8 (0.5)
T.3	I would always use this system	2.4 (1.4)

Table 1: *We interviewed five pathologists regarding the feature visualization (FV), the workflow optimization (WO) and the trust (T) of our tool. The average agreements are displayed for each question. We assigned score 4 to "strong agreement", 3 to "agreement", 0 to "disagreement" and -1 to "strong disagreement". The mean and the standard deviation (in the brackets) are shown for each theme question.*
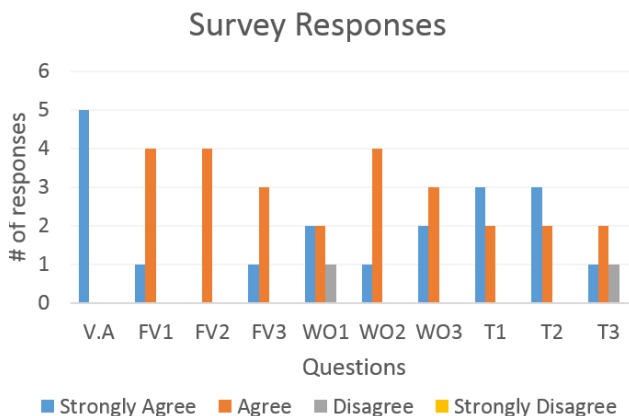
Figure 7: *View of survey results. The five pathologists agreed on the majority of our questions concluding that the visual analytics funtionalities of our CAD system suffice the NG diagnostic work. Refer to Table 1 as a legend to questions abbreviations.*

by using the *CAD-Tiles* tool. We then interviewed the five pathologists to collect their opinions on the application. Our survey covers the following themes: features visualization, workflow optimization and trust. As shown in Table 1 and Fig. 7, all the pathologists

agreed with the importance of visual analytics within the field of digital pathology. The pathologists appreciated the visualization of the three diagnostic primitives and they considered that the *CAD-Tiles* widget would not disturb the grading process. From the workflow perspective, the general opinion was that the tool would increase the speed in the diagnostic routine. Some concern was expressed regarding the number of grids required to perform the final diagnosis. In this case, the pathologists expressed the belief that automated grids can be generated on the tumor regions to further reduce the manual effort. In answer to this, we also support the creation of a *CAD-Tiles* grid generated upon the entire tissue area if considered necessary. Hence, we asked whether the pathologists would trust the proposed system in the scenario of highly accurate detection (for both specificity and sensitivity). The agreement was high and only one pathologist claimed that he would restrain the use of this tool only for complex cases. All pathologists admitted that the gland formation represents the most debatable feature within the *CAD-Tiles* widget. In fact, a simple indication of the percentage of gland formation in a region is enough according to the protocol. The difficulty in manual routine as in machine-detection is to discriminate between abnormal and normal gland formation. This distinction may require a more detailed visualization than the one presented in this work. A more extensive approach on this aspect will be addressed in future work with the availability of more accurate gland detection features. We think that visual concepts on this kind of feature can be extended to other low magnification features and to other histologic diagnosis. This evaluation represents a solid input for a further quantitative evaluation. We believe that this will be feasible when stable image analysis results and more trustful features will be integrated in the tool with a larger set of tissue slides.

7. Conclusions

In this paper we have introduced a CAD application for histopathology diagnostics in breast cancer. We presented its implementation and development throughout the standard design process of a visual analytics system. From the literature we extracted one of the strategies typically adopted to analyze pathologists' behavior. We observed that the first CAD tools implemented for digital pathology do not meet the needs of pathologists' diagnostic process with visual support of CAD findings. In order to address this gap, we observed some diagnostic traces to elicit general requirements to design visual analytics support in histopathology. From our knowledge, we conducted the first analysis on pathologists' behavior in relation to diagnostic tasks. Despite the limited experiment conducted with only one pathologist, our findings induce us to think that this can be inspiring for future analysis and in the development of new applications. For this particular study, the task analysis has been converted in a subdivision of the requirements in two areas: visual aspects of the histologic primitives and diagnostic principles. Accordingly, the design choices followed the requirements and the design conveyed in a protocol-centric application.

In support of the standard workflow, we created the *CAD-Tiles* widget to interact with high magnification primitives such as nuclei and mitosis at low magnification levels. Similarly, this widget can be used for the quantification and the presentation of data for other use cases beside breast cancer diagnosis. This can be the case, in the future, for the integration of new emerging important prognostic indicators like lymphocytes density [ADP*16] or insight on tumor-stroma ratio [DTSS15] once they will become established factors in the diagnosis.

In this context, we limited our CAD system to the scenario of the Nottingham Grading System. We illustrated the impact that visual analytics can have in tedious and time-consuming tasks such as mitotic counting. In addition to that, the adopted protocol-based approach can be extended to other histologic cases whether they are standardized across laboratories as the Invasive Breast Cancer protocol. In a first evaluation we received a positive feedback from five pathologists. We think that the visual analytics capabilities of our application will benefit of a further and more extensive evaluation in a broader setting.

In conclusion, this study represents the first integration of image analysis features into a histologic protocol-based diagnostic workflow. We aimed to pave the way to CAD applications that can be directly used by pathologists in a clinical setting. We consider that the described study can simplify the creation of such systems in digital pathology and accelerate their introduction in clinical routine. Our work was restricted to a limited number of users. It was, therefore, not possible to collect a significant range of opinions from the pathologists regarding the use of our tool in comparison to a fully automated workflow. As evidence of this work, we can affirm that an attentive translation of protocol requirements into visual analytics capabilities will find positive consensus among pathologists in the first adoption of computational pathology in diagnostics and towards fully automated steps of their routine tasks.

References

[ADP*16] ALI H. R., DARIUSH A., PROVENZANO E., BARDWELL H., ABRAHAM J. E., IDDAWELA M., VALLIER A.-L., HILLER L., DUNN J. A., BOWDEN S. J., HICKISH T., MCADAM K., HOUSTON S., IRWIN M. J., PHAROAH P. D. P., BRENTON J. D., WALTON N. A., EARL H. M., CALDAS C.: Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Research 18*, 1 (dec 2016), 21. 9

[AHZ*13] ARPS D. P., HEALY P., ZHAO L., KLEER C. G., PANG J. C.: Invasive ductal carcinoma with lobular features: a comparison study to invasive ductal and invasive lobular carcinomas of the breast. *Breast cancer research and treatment 138*, 3 (apr 2013), 719–26. 3

[AWM*17] AEFFNER F., WILSON K., MARTIN N. T., BLACK J. C., HENDRIKS C. L. L., BOLON B., RUDMANN D. G., GIANANI R., KOEGLER S. R., KRUEGER J., YOUNG G. D.: The Gold Standard Paradox in Digital Image Analysis: Manual Versus Automated Scoring as Ground Truth. *Archives of Pathology & Laboratory Medicine 141*, 9 (sep 2017), 1267–1275. 1, 3

[BM13] BREHMER M., MUNZNER T.: A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (dec 2013), 2376–2385. 3

[BVJ11] BOSTOCK M., VADIM O., JEFFREY H.: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics 17* (Dec 2011), 2301 – 2309. 8

[Cha14] CHAN J. K. C.: The Wonderful Colors of the Hematoxylin-Eosin Stain in Diagnostic Surgical Pathology. *International Journal of Surgical Pathology 22*, 1 (feb 2014), 12–32. 3

[CML16] CERVIN I., MOLIN J., LUNDSTRÖM C.: Improving the creation and reporting of structured findings during digital pathology review. *Journal of pathology informatics 7* (2016), 32. 2

[CQY*17] CHEN H., QI X., YU L., DOU Q., QIN J., HENG P.-A.: DCAN: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis 36* (feb 2017), 135–146. 4

[CvDW17] CORVO A., VAN DRIEL M. A., WESTENBERG M. A.: PathoVA: A visual analytics tool for pathology diagnosis and reporting. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)* (oct 2017), IEEE, pp. 77–83. 3

[CVS*13] CLARK K., VENDT B., SMITH K., FREYMANN J., KIRBY J., KOPPEL P., MOORE S., PHILLIPS S., MAFFITT D., PRINGLE M., TARBOX L., PRIOR F.: The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal of digital imaging 26*, 6 (dec 2013), 1045–57. 8

[DTSS15] DOWNEY C. L., THYGESEN H. H., SHARMA N., SHAABAN A. M.: Prognostic significance of tumour stroma ratio in inflammatory breast cancer. *SpringerPlus 4* (2015), 68. 9

[EE02] ELSTON C. W., ELLIS I. O.: Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. C. W. Elston & I. O. Ellis. Histopathology 1991; 19; 403-410. *Histopathology 41*, 3A (sep 2002), 151–2, discussion 152–3. 3

[ELC*15] ELMORE J. G., LONGTON G. M., CARNEY P. A., GELLER B. M., ONEGA T., TOSTESON A. N. A., NELSON H. D., PEPE M. S., ALLISON K. H., SCHNITT S. J., O 'MALLEY F. P., WEAVER D. L.: Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens Michael's Hospital and the. *JAMA* (2015). 3

[FB11] FUCHS T. J., BUHMANN J. M.: Computational pathology: Challenges and promises for tissue analysis. *Computerized Medical Imaging and Graphics 35*, 7-8 (oct 2011), 515–530. 1

[Fin14] FINE J. L.: 21(st) century workflow: A proposal. *Journal of pathology informatics 5*, 1 (2014), 44. 1, 3, 5

[GBC*09] GURCAN M. N., BOUCHERON L. E., CAN A., MADABHUSHI A., RAJPOOT N. M., YENER B.: Histopathological image analysis: A review. *Biomedical Engineering, IEEE Reviews in 2* (2009), 147–171. 1

[geo17] Geoserver. http://geoserver.org/, 2017. "[Online; accessed February-2017]". URL: http://geoserver.org/. 8

[Haz17] HAZZARD: Openlayers 4.11. http://openlayers.org/, 2017. "[Online; accessed February-2017]". URL: http://openlayers.org/. 8

[HSC06] HADJIISKI L., SAHINER B., CHAN H.-P.: Advances in computer-aided diagnosis for breast cancer. *Current opinion in obstetrics & gynecology 18*, 1 (feb 2006), 64–70. 1

[JCvO15] JORRITSMA W., CNOSSEN F., VAN OOIJEN P.: Improving the radiologist-CAD interaction: designing for appropriate trust. *Clinical Radiology 70*, 2 (feb 2015), 115–122. 8

[JM16] JANOWCZYK A., MADABHUSHI A.: Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics 7*, 1 (2016), 29. 1

[LBC*09] LESTER S. C., BOSE S., CHEN Y.-Y., CONNOLLY J. L., DE BACA M. E., FITZGIBBONS P. L., HAYES D. F., KLEER C., O'MALLEY F. P., PAGE D. L., SMITH B. L., TAN L. K., WEAVER D. L., WINER E., MEMBERS OF THE CANCER COMMITTEE, COLLEGE OF AMERICAN PATHOLOGISTS: Protocol for the examination of specimens from patients with invasive carcinoma of the breast. *Archives of pathology & laboratory medicine 133*, 10 (oct 2009), 1515–38. 1, 2, 8

[LBG*15] LANGER L., BINENBAUM Y., GUGEL L., AMIT M., GIL Z., DEKEL S.: Computer-aided diagnostics in digital pathology: automated evaluation of early-phase pancreatic cancer in mice. *International Journal of Computer Assisted Radiology and Surgery 10*, 7 (jul 2015), 1043–1054. 2

[LP11] LUNDSTRÖM C., PERSSON A.: Characterizing visual analytics in diagnostic imaging. *International Workshop on Visual Analytics* (2011). 2

[LST*16] LITJENS G., SÁNCHEZ C. I., TIMOFEEVA N., HERMSEN M., NAGTEGAAL I., KOVACS I., HULSBERGEN - VAN DE KAA C., BULT P., VAN GINNEKEN B., VAN DER LAAK J.: Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports 6*, 1 (sep 2016), 26286. 1

[MFMTL15] MOLIN J., FJELD M., MELLO-THOMS C., LUNDSTRÖM C.: Slide navigation patterns among pathologists with long experience of digital review. *Histopathology 67*, 2 (aug 2015), 185–92. 2, 3

[NRS*14] NUNES M., ROWLAND B., SCHLACHTER M., KEN S., MATKOVIC K., LAPRIE A., BUHLER K.: An integrated visual analysis system for fusing MR spectroscopy and multi-modal radiology imaging. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)* (oct 2014), IEEE, pp. 53–62. 2

[OEC*16] OZKAN T. A., ERUYAR A. T., CEBECI O. O., MEMIK O., OZCAN L., KUSKONMAZ I.: Interobserver variability in Gleason histological grading of prostate cancer. *Scandinavian Journal of Urology 50*, 6 (nov 2016), 420–424. 1

[PAF09] PENA G. P., ANDRADE-FILHO J. D. S.: How does a pathologist make a diagnosis? *Archives of pathology & laboratory medicine 133*, 1 (jan 2009), 124–32. 1, 5

[PKC98] P. KRONQVIST T. K., COLLAN Y.: Morphometric grading of invasive ductal breast cancer. I. Thresholds for nuclear grade. *British journal of cancer 78*, 6 (sep 1998), 800–5. 7

[RBV17] RAIDOU R., BREEUWER M., VILANOVA A.: Visual analytics for digital radiotherapy: Towards a comprehensible pipeline. *Computer Graphics Forum (Proceedings of Eurographics) 36* (Apr. 2017). 2

[RESL*08] RAKHA E. A., EL-SAYED M. E., LEE A. H. S., ELSTON C. W., GRAINGE M. J., HODI Z., BLAMEY R. W., ELLIS I. O.: Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology 26*, 19 (jul 2008), 3153–8. 3

[RRFB*10] RAKHA E. A., REIS-FILHO J. S., BAEHNER F., DABBS D. J., DECKER T., EUSEBI V., FOX S. B., ICHIHARA S., JACQUEMIER J., LAKHANI S. R., PALACIOS J., RICHARDSON A. L., SCHNITT S. J., SCHMITT F. C., TAN P.-H., TSE G. M., BADVE S., ELLIS I. O.: Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research 12*, 4 (aug 2010), 207. 3

[SKM*13] SABO E., KLORIN G., MONTGOMERY E., DRUMEA K. C., BEN-IZHAK O., LACHTER J., VIETH M.: Subjective grading of Barrett's neoplasia by pathologists: correlation with objective histomorphometric variables. *Der Pathologe 34*, 2 (mar 2013), 133–7. 1

[SSMD10] SUGIMOTO K., SHIRAISHI J., MORIYASU F., DOI K.: Computer-aided diagnosis for contrast-enhanced ultrasound in the liver. *World journal of radiology 2*, 6 (jun 2010), 215–23. 1

[SVHvd13] STATHONIKOS N., VETA M., HUISMAN A., VAN DIEST P. J.: Going fully digital: Perspective of a Dutch academic pathology lab. *Journal of pathology informatics 4* (jan 2013), 15. 1

[THDA14] TASHK A., HELFROUSH M. S., DANYALI H., AKBARZADEH M.: A Novel CAD System for Mitosis detection Using Histopathology Slide Images. *Journal of medical signals and sensors 4*, 2 (apr 2014), 139–49. 1

[VPvDV14] VETA M., PLUIM J. P. W., VAN DIEST P. J., VIERGEVER M. A.: Breast Cancer Histopathology Image Analysis: A Review. *IEEE Transactions on Biomedical Engineering 61*, 5 (may 2014), 1400–1411. 1

[VvDJ*16] VETA M., VAN DIEST P. J., JIWA M., AL-JANABI S., PLUIM J. P. W.: Mitosis Counting in Breast Cancer: Object-Level Interobserver Agreement and Comparison to an Automatic Method. *PLOS ONE 11*, 8 (aug 2016), e0161286. 1, 4

[Wor14] WORLD HEALTH ORGANIZATION: World cancer report 2014. *IARC Publications* (2014). 3