OpenHI: Open platform for histopathological image annotation

Author

Address

Abstract: Consolidating semantically rich digital histopathological image by annotating scanned glass slides known as whole-slide images requires a software capable of handling this type of biomedical data and a support for procedures which align with existing pathological routine. Demand for large-scale annotated histopathological datasets are on the raise because they are needed for developments of artificial intelligence techniques to promote automatic diagnosis, mass screening, or phenotype-genotype association study. This paper presents an open platform for efficient collaborative histopathological image annotation with standardized semantic enrichment at a pixel-level precision named OpenHI—Open Histopathological Image. The framework's responsive processing algorithm can perform large-scale histopathological image annotation and serve as biomedical data infrastructure for digital pathology. It is highly configurable and could be extended to annotate histopathological image of various oncological types. The framework is open-source and fully documented. It is publicly available at https://gitlab.com/BioAI/OpenHI.

Keywords: OpenHI; digital pathology; whole-slide image; WSI; image annotation; virtual slide; virtual magnification; histopathology; cancer diagnosis; cancer grading; genotype-phenotype association

1 Introduction

Crucial part of cancer diagnosis is an examination of histopathological slides performed by pathologists. Tissue samples obtained via biopsies are prepared—cut, stained, and fixed on glass slides, resulting in diagnostic slides containing patterns that would indicate if particular patient has sings of a disease. It could be stated that histopathological images present rich information about the cancer microenvironment (True and Jordan, 2008). Since the beginning of pathology discipline, interpretations of tissue slides have been manually carried out by human pathologists through grading standards and guidelines established by communities of pathologists, mainly, with a bright field light microscope (Joseph P. Houghton *et al.*, 2014). The knowledge in pathology was handed over to the

newer generations of pathologists and clinicians through established grading standards, case reports, etc. which often describe structures of tissues by natural language or technical pathological terms with limited visual examples, WHO/ISUP grading standard would be one example (Delahunt *et al.*, 2013). Two type of data from this branch of knowledge could be exploited including visual phenotypic data from large-scale histopathological images and expert knowledge which supports the decision-making process in cancer diagnosis. In this study, we investigate procedures for enriching tissue slides with semantic knowledge, our proposing open platform should help researchers to overcome technical and biomedical obstacles, accelerating the development of artificial intelligence in digital pathology.

Since pathology is one of the last medical specialties to be digitized (Goode *et al.*, 2013), there has been limited software and tools that helps pathologists manage and analyze extra-large digital scans of glass slides–namely, whole-slide images (WSIs). At the current stage of digital pathology, there are no central group developing and maintaining standards, guidelines, or best practices. Thus, researchers and pathologists are struggling to choose appropriate data infrastructure for their projects that must support WSIs. Furthermore, even fewer open-source software may be used for annotating and analyzing the WSIs. No existing software supports online multi-user annotation with detailed spatial resolution and semantic meaning. Here, we present OpenHI—Open Histopathological Image, a publicly available open-source framework for WSI. It is capable of achieving pixel-level precise boundary and semantic annotation, and supports online collaborative annotation. Eventually, the proposed framework will facilitate the large-scale histopathological image annotation and benefit machine-learning based phenotype extraction.

In recent years, there have been multiple implementations of machine learning methods to enhance histopathology image analysis workflow by either assisting pathologists in image analysis or by establishing automated pipeline to analyze, detecting and classifying the cancerous area, the WSI with high throughput and high precision to help reduce the workload of the pathologists (Kurc *et al.*, 2015; Xu *et al.*, 2017; Zhou *et al.*, 2017; Mercan *et al.*, 2018). However, these works were derived and tested on datasets with finite size and variability due to limited availability of public annotated datasets. With our proposed framework, rapid creation of such datasets could be accomplished, thus sophisticate computational approach that holistically analyze the WSIs may lead to a better grading decision (El-Gabry, Parwani and Pantanowitz, 2014).

Annotation of medical data is expensive since it demanded labor for manual data annotation from experienced medical personnel (Xu *et al.*, 2017). The scarcity of experts is not the only factor contributing to this problem. Pathologists around the work do not use the same grading standard to grade the sample. Pathologists in some countries must grade their tissue sample according to national standard provided by the local authorities. Measures must be taken to reduce the cost of biomedical data annotation. OpenHI was designed as a collaborative annotation tool, ensuring that the most efficient annotation is carried out by acquiring only the essential data from the annotators while maintaining high-quality annotation. Existing pathological routine was taken into consideration during the framework design. Collaboration between pathologists, as a source of expert knowledge, and data scientists, who will manage the acquired data, is necessary to complete large-scale histopathological image data annotation and cross-validate the quality, thus the annotation framework would need to minimize technical configuration at the annotator end (pathologist) and maximize configurability at the data scientist end. The ideal software for this application would minimize the annotators' effort to annotate the image while

capturing high information granularity including high spatial resolution and standardized semantic meaning. In the meantime, it should support online collaboration as well.

2   Related work

In cancer research, large-scale raw histopathology image repositories have been made publicly available by TCGA (Tomczak, Czerwińska and Wiznerowicz, 2015) project, GTEx (Keen and Moore, 2015) project, and some other projects. The repositories contain tremendous amount of WSIs in different cancer types and the amount of data is continuously growing. The main barrier of fully utilizing the data for automated methods is the lack of annotations. In recent years, there are enriched WSI datasets used in different image analysis competitions (B *et al.*, 2017; Veta *et al.*, 2018), however, they are limited by public availability, size, variability in cancer type, or spatially precise annotation. An annotation software capable of capturing precise visual annotation and semantic information enrichment is highly demanded.

WSI files are hard to be read efficiently because WSIs cannot be saved in standard image format due to its unusually high resolution. In response, different whole slide scanner venders have come up with their own proprietary standard and file format to save WSIs as tiled multi-scale images. The DICOM or Digital Imaging and Communications in Medicine standard is the only public and general standard (Clunie *et al.*, 2018) that has been proposed but has not been adopted or popularized (Singh *et al.*, 2011) despite the effort to demonstrate the use of DICOM in digital pathology (Herrmann *et al.*, 2018). In 2013, (Goode *et al.*, 2013) have introduced an open-source library to read the WSIs called OpenSlide which later become the only available vendor-neutral tool to read the WSIs to date. Many other WSI visualization and analysis software have adopted the library to create web-based application (Gutman *et al.*, 2013, 2017), stand-alone software such as QuPath (Bankhead *et al.*, 2017) and ASAP (Litjens, 2015), and extension, SlideJ (Della Mea *et al.*, 2017), to ImageJ (Schindelin *et al.*, 2015). Around the same time, OpenSeadragon (OSD) library ('OpenSeadragon Project', 2013) was introduced as a web-based JavaScript-based viewer for high-resolution zoomable image. It is capable of viewing the multi-scale images including WSIs. OSD is then used in web-based implementation of WSI viewer.

The web-based implementation of OpenSlide with OSD to help visualize the WSIs from TCGA project on the webpage could be seen in the US National Cancer Institute's Genomic Data Commons. It helps the users to visualize the WSIs without downloading the entire large WSI file, however, it lacks the functionality to modify, annotate, or analyze the WSIs. In 2017, QuPath (Bankhead *et al.*, 2017) was introduced as a cross-platform stand-alone software. It is a tool to view WSIs on local machine and it is capable of accomplishing many tasks including basic annotation of the WSIs and locally segmenting the image with superpixel algorithms. The most detailed annotation method that QuPath can achieve is selecting multiple points in the image to form a polygon, this approach is good for manually mark a small number of regions for human references, it is not detailed enough for computers. Furthermore, the annotation data made in QuPath has to be managed manually and it does not provide centralized system to manage the WSIs or annotation data. Some works use point-based selection (Dong *et al.*, 2014; Xing, Xie and Yang, 2016) to mark at the center of nuclei, but this method does not work in the situation where area containing several nuclei must be selected. It is clear that conventional hardware such as mouse and keyboard was not designed for effectively delineating precise boundaries in natural images, thus we also propose a method to automatically select precise region boundaries.

No open-source collaborative annotation software specifically made for histopathological image is publicly available at the moment. Besides, such software should allow online-collaboration to achieve high annotation throughput and maximize the accessibility for the users since they do not have to download the entire dataset and install additional software. It should also be able to manage highly detailed annotation with region-specific semantic enrichment. The challenges will be addressed with our proposed framework.

3   Material and method

Constructing pathologically meaningful annotations for WSI requires standardized procedures which should align with existing pathological routines (Zarella *et al.*, 2018). The way annotators use OpenHI was designed based on pathology grading standards. Currently, the platform can capture grades for well-known cancer type that use single grading scheme e.g. WHO/ISUP grading for renal cancer, Gleason grading (Gleason, 1992) for prostate cancer, Nottingham grading for breast cancer, and so on. The proposed platform was designed to be implemented on a web server therefore it could be accessed simultaneously via a web-browser by multiple users without geographical restrictions. To minimize effort of acquiring most detailed annotations with precise sub-region boundaries and semantic enrichment, our framework pre-segment the image into semantically meaningful sub-regions as demonstrated in Figure 1 by a widely used graph-based image segmentation method called SLIC superpixel (Achanta *et al.*, 2012). In this case, the annotators can quickly select the sub-regions by simply clicking or dragging mouse through them via the graphic user interface. Additionally, our framework has the ability to freely access any regions of WSI at will using different zooming level based on OpenSlide being utilized as the others did (Gutman *et al.*, 2013; Litjens, 2015; Bankhead *et al.*, 2017). OpenHI is developed on other well-established libraries and tools, thus the project should be easy to maintain and extend by other developers.

3.1   Annotation design

Tissue grading decision can be made at different levels: patient/case, slide, cluster of cells, or nuclei. While patient-level grade is the ultimate goal of histopathological examination, the underlying decision-making process relies on a more granular findings about different areas of different slides. Appropriate grade must be assigned to segments with precise spatial boundaries. The platform enables users to freely adjust to a particular grade and select applicable image segments. By achieving granular annotation, prediction systems developed using the data will be able to analyze digital slides in similar ways as human pathologists do.

Adjusting tumor grade setting in the platform can be done via the control panel where annotators can select one of the pre-defined grades according to chosen grading system. For example, annotators may choose either grade 1, 2, 3, 4, or healthy in a 4-tier grading system such as WHO/ISUP or Gleason guideline. After the grade has been decided,

annotators would select or mark relevant segments. The entire process can be summarized in to a simple annotation procedure: "observe-grade-mark", the procedure can be repeated until the annotation is finished.

3.2 Proposed framework

The platform consists of three main components along with a MySQL database server to store the annotation coordinates. The main components include image pre-processing of the WSIs, the server-end module, and the GUI. Data stored stored in the platform are WSI files, sub-region boundary matrices, annotation coordinates in MySQL database, and metadata of each slide image. Figure 2 illustrates each component and the data flow between them.

3.2.1 Image pre-processing

A WSI is a very large 2-dimensional array of data. It is too large to be handled by most (and in some cases all) of conventional image formats. Furthermore, the conventional formats are currently not a vendor natural standard for such kind of data (Cooper *et al.*, 2015). Conventional image segmentation technique has also faced a challenge in processing this kind of image as well. It is known that processing WSIs is memory demanding (Della Mea *et al.*, 2017). The SLIC Superpixel (Achanta *et al.*, 2012) segmentation on the WSIs also requires large amount of memory. Thus, it is reasonable to have memory-consuming processes deployed on a server with relatively larger amount of memory than typical computers. To our knowledge, there are currently no practical implementation on superpixel with sufficient segmentation precision on whole WSIs on a scale of personal computer, however, there are efforts to alter superpixel algorithm for implementation with large images. Developing a robust and fast image segmentation method is still a challenge in digital pathology informatics (Kothari *et al.*, 2013). By incorporating superpixel segmentation with the WSIs, we have established a new ROI selection method in digital pathology.

One lasting issue with superpixel algorithm in image segmentation is selecting a number of final segmented sub-region ($N_{superpixel}$). This number is crucial since it could lead to under- or over-segmentation. Tuning for good number of segments to avoid under- and over-segmentation in superpixel algorithm has been a challenge in utilizing the method. It is even more problematic to choose one number of segments for all WSIs since each of them has different dimensions. To tackle this problem, we shift the idea of number of sub-regions to average sub-region size (P_{sub}). We calculate the $N_{superpixel}$ by specifying desired P_{sub}, thus the size of the sub-region will be consistent throughout the annotation project as shown in (1) where P_{total} is resolution of each WSI. From our trial in WSIs with 0.25 micron/pixel resolution at 20x objective lenses, this number could range from around 6,000 to 50 pixels/sub-region to effectively cluster cells. The example of a portion of the image pre-segment with superpixel algorithm is shown in Figure 3.

$$N_{superpixel} = \frac{P_{total}}{P_{sub}} \qquad (1)$$

To avoid over-segmentation where an annotator has to choose unnecessarily large number of sub-regions to establish one cancerous region or under-segmentation where a cancerous region is not cleanly divided from normal regions (see Figure 4), our framework provides users an option to calibrate at multiple pre-segmentation levels and select the most suitable level for the annotation in specific area of the image. During the annotation session, the annotator could switch back and forth between different pre-segmentation level for a level that a cancerous region could be accurately separated at the same time with the consideration of the annotator's convenience and efficiency.

At the end of pre-processing, matrices containing sub-region boundary data are stored as a binary image using portable network graphic (PNG) format. This image with boundary data is large in dimensions, but not in file size, therefore it is practical to store them as one continuous image where it could be easily loaded into the memory when needed.

3.2.2 Server-end module

The second component of the framework is to interactively respond to the user request in real-time. To lower the computational expenses, only the area of a WSI being viewed is processed. The server-end module includes sub-region boundary generation, existing annotated sub-regions visualization, annotation coordinates recording and deletion, and virtual magnification factor calculation.

Composing the final viewing image which is sent to the graphic user interface requires three components: (1) original WSI, (2) sub-region boundary, and (3) existing annotation. Part of WSI can be easily obtained via OpenSlide by stating coordinates of requested viewing area. Similarly, sub-region boundary can be easily obtained by image cropping operation. Visualizing existing annotations requires data entries from the database server and sub-region boundary image. Corresponding data entries of annotated sub-regions will be fetched from the database server via SQL queries. The representation point of each sub-region will be treated as a seed point for a flood-fill operation, thus entire sub-region could be visualized, and corresponding color for each grade will be filled. Finally, the sub-region boundary layer is embedded into the WSI, then the visualized annotations layer is alpha-blended, preserving a certain degree of transparency so that annotators can still see through the annotation. This process is illustrated in Figure 5.

The server-end module also supplies the GUI with some indicator value to approximate the zooming power of certain digital slide zooming state, we call this virtual magnification. Some grading system such as WHO/ISUP kidney clear cell carcinoma annotation standard (Delahunt *et al.*, 2013) (see Table 1) require the annotators to take magnification power as a part of the final decision. This functionality is crucial for grading systems that rely on microscope magnification. It is also a worthy indicator for the pathologist who is new to digital pathology as well.

To accurately calculate virtual magnification, we need to understand that real magnification in the microscope is the combination of magnification from objective and eye piece lenses as in (2). In virtual slides, the objective magnification (M_{obj}) is restrained by magnification factor of the objective lenses used during the scan which is specified in the metadata of the WSI file, stated as a parameter named pixel size and often has a unit of micron/pixel. Tracing back from micron/pixel to objective lenses magnification power can be made by referring to the specification of whole-slide scanner. In our testing dataset, the WSIs with pixel size of 0.25 and 0.50 micron/pixel was scanned by 40x and 10x objective lenses respectively. The eye piece magnification (M_{eye}) is more complicated to calculate,

three parameters are needed for the calculation including scanner sensor pixel size, monitor's pixel size, and distance between the monitor and the user. T. Sellaro et al's method can be used to calculate virtual objective magnification and virtual eye piece magnification; there are calculation sample in their work (Sellaro *et al.*, 2013). Finally, the total virtual magnification (M_{total}) could be summarized by (2).

$$M_{total} = M_{obj} \times M_{eye} \qquad (2)$$

To avoid complications in our framework, we simplify the virtual magnification calculation and focus on three parameters which are WSI pixel size (S_{WSI}), monitor pixel size ($S_{monitor}$), and digital zooming factor ($M_{digital}$). The parameter image pixel size reflects the real-world size of the scanning object which we use as an anchor or original size and the magnification is the zooming factor based on this size. The monitor size can be calculated by considering the screen resolution and the monitor physical size. For example, the 24-inch computer display with 1920-by-1080 pixel resolution would have a monitor pixel size of approximately 270 micron. OpenHI can automatically detect these setting with the standard JavaScript protocol. To calculate the magnification factor by incorporating the image pixel and screen size can be done by (3). The digital zoom can be calculated by the size of the viewer shown in the GUI (S_{viewer}) and the size of the image being viewed in the viewer ($S_{viewing}$) as shown in (4). Finally, the virtual magnification in OpenHI can be calculated by (5).

$$M_{WSI-monitor} = \frac{S_{monitor}}{S_{WSI}} \qquad (3)$$

$$M_{digital} = \frac{S_{viewer}}{S_{viewing}} \qquad (4)$$

$$M_{total} = M_{WSI-monitor} \times M_{digital} \qquad (5)$$

3.3.3 Graphic user interface

The web-based GUI of our proposed framework is comprised of the main WSI viewer, virtual magnification indicator, control panel for annotation configuration, and slide information panel as shown in Figure 5.

As the use of OSD with OpenSlide has been demonstrated in (Goode *et al.*, 2013) in 2013, we extend the usage for WSI annotation. Our framework has attained the capabilities of OSD and OpenSlide to view a WSI with smooth zooming and panning experience similar to other image viewers while adding the customized annotation capability to the framework. This allows the annotators to easily annotate any part of WSI they want. When annotating a number of sub-regions with the same grade, an annotator may select each individual sub-region by a mouse click. Rather than clicking the sub-regions one by one, the framework provides an option of clicking then pressing mouse over several sub-regions to do bulk annotation since the adjacent sub-regions tend to contain the same swamp of cancerous cells.

Control panel allows users to change some configurations to make the annotation easier including the pre-segmentation level, tumor gradings, undo button, and option to show or hide sub-region boundaries along with existing annotations. An alternative way to adjust these parameters is the usage of keyboard hotkeys that could speed up the annotation once users get familiar with the system. Virtual magnification indicator—zooming factor and resolving power—helps annotators to get a sense of how much zooming power they are using.

Slide information panel allows annotators to look at basic information about the tissue slide. The information may come from electronic medical record or slide preparation process. The framework supports data written in XML format recorded in the US Genomic Data Commons (GDC) standard. The parsed data is simplified into key-value pairs. Examples of keys are gender, age, race, etc.

3.3.4 Database server

Due to simultaneous access—read and write—to annotation data on any part of any WSIs from multiple users, MySQL database server is chosen to cope with this job. Each user interaction to annotate the WSI is recorded as an entry. The number of entries could grow up to several thousand for a single WSI. To maximize the granularity of the annotation while keeping the utility of collected data, properties in each entry includs exact chosen coordinate x and y, assigned tumor grading, pre-segmentation level, time of annotation, slide ID, and annotator ID.

The annotation coordinates are stored based on the highest resolution of a WSI to preserve the precision. Each of the coordinates is paired to the tumor grading which the annotator has assigned during the annotation of that point. The point is also bound to the pre-segmentation level which represents the boundary of each sub-regions. Retrieving existing annotation according to the area being viewed is easily done by simple SQL queries.

To maximize the extendibility of this annotation software, it is designed to be highly customizable. Some parameters are designed for the annotators (pathologists) and can be switched during annotation processes. The other parameters are made for the data scientist so that they can achieve the sufficient annotation quality for further uses.

3.3 Dependencies

The image pre-processing module was implemented by MATLAB version 9.4.0 (R2018a) with image processing toolbox version 10.2 to perform SLIC superpixel algorithm (Achanta *et al.*, 2012). Noted that superpixel segmentation could also be implemented by other open source image processing library such as OpenCV or scikit-image (van der Walt *et al.*, 2014).

The server-end module is written in Python 3. We have adopted Flask 1.0.2 as our web framework to manage information flowing between the database and the annotators. Within the web framework, OpenSlide is used to read WSI files and OpenCV with Numpy are used to generate images to be sent to the client. To connect to MySQL server, the module 'mysql-connector' are used.

In the web-based GUI, OSD is used as an image viewer and annotating coordinates and user interaction to navigate around the WSI are acquired through OSD. JQuery are used for handling HTTP requests.

4 Result

The scalability and extendibility of the proposed framework would make large-scale collaborative annotation of WSIs possible. It does allow online cooperation between a group of researchers from different geographical locations to perform annotation on an online platform where data scientists can reconfigure the framework for the need of each specific project using different parameters as shown in Table 2, thus the annotator's effort to reconfigure the software is minimized. It allows users to view WSI scans from multiple vendors and let them navigate around smoothly. Annotation with high information granularity can be accomplished. The proposed framework can ultimately be a foundation of crowed-sourcing WSI annotation platform. Enabling online real-time collaboration converges effort from pathologists to annotate and cross-validate large-scale images.

4.1 Functionality

The proposed framework offers the freedom of choosing from pre-defined segmentation levels and switching during annotation. Meanwhile, the pre-defined segmentation levels are kept consistent on different images. For example, a sub-region in one image would contain approximately the same number of cells as another sub-region does in a different image, as long as the two sub-regions comply with the same pre-defined segmentation level. The consistency in the sub-region average size is achieved because the number of segments in superpixel segmentation was calculated by the desired sub-region size and the image size. The ability to support multiple pre-defined segmentation level which let the user to interactively switch between them during the annotation session can increase annotation efficiency.

The proposed framework offers two methods of selecting annotation sub-regions which is a click to select individual sub-region or hold-and-drag over multiple sub-regions to select as illustrated in Figure 7 (b) and (c) respectively. The latter option is useful since a swamp of cancerous cells is likely to occupy several consecutive sub-regions. Selecting sub-regions in this manner could be relatively much faster than using relatively more mouse clicks to form an equivalently precise polygon as shown in Figure 7 (a). The comparison can be seen in Figure 7. Besides the easy sub-region selection, the user can also select gradings easily via the control panel next to the viewer as shown in Figure 6 or by hotkeys. Additionally, in the case that the annotator mistakenly selects the region, the framework can revert to previous step or the annotator can deselect some sub-regions.

4.2 Virtual magnification

The magnification is crucial to some grading standard such as WHO/ISUP grading system for renal cell carcinoma where it requires the uses of microscope magnification to make a grading decision. Viewing the tissue image in virtual environment could be challenging for pathologists since there are more factors contributing to the actual size of image that pathologists would see such as monitor size and density which varies between monitors, unfixed distance between the eyes and the monitor, and continuous zooming instead of discrete zooming heads in conventional microscope. Nevertheless, there are a few studies about the differences between diagnosis from conventional microscopy and virtual slides (Jukić *et al.*, 2011; Mukhopadhyay *et al.*, 2018). These studies have ensured us that WSI

is noninferior to microscopy for primary diagnosis. Digital pathology has been virtualizing all aspects of conventional microscope including zooming, panning, and magnifying. The only aspect of microscope that has not been perfected is magnifying because all previous viewers lack a precise virtual magnification to choose in virtual slide viewing due to the fact that there are many factors affecting virtual magnification. Furthermore, grading for this type of cancer also has major discordance between WSI and conventional microscopy in (Mukhopadhyay *et al.*, 2018) therefore we think that it is important to improve this aspect of virtual slide for renal clear cell carcinoma gradings. Besides, we have added accurate virtual magnification in our viewer to overcome this problem.

4.3 Extendibility

In term of digital slide formats, since we utilize OpenSlide, the proposed framework can support various WSI formats from different scanner vendor including Aperio (.svs, .tif), Hamamatsu (.vms, .vmu, .ndpi), Leica (.scn), MIRAX (.mrxs), Philips (.tiff), Sakura (.svslide), Trestle (.tif), Ventana (.bif, .tif), and Generic tiled TIFF (.tif) (Goode *et al.*, 2013).

The proposed software is free and open-source, it is available at GitLab (https://gitlab.com/BioAI/OpenHI) under GNU General Public License v3.0, therefore it can be modified to suit the need in different purposes. The framework is also compatible with general LAMP stack which is widely available on the could computing platforms or local server environment.

4.4 Data acquisition

In software development and testing, we use WSIs directly downloaded from TCGA data repository (Tomczak, Czerwińska and Wiznerowicz, 2015). Thus the testing environment, WSI format used in our proposed framework is Aperio (.svs) file. The images are tissue slides scanned with 20x magnification with resolving power of 0.5 micron/pixel (Cooper *et al.*, 2015). In our sample set of data, the average resolution of WSIs is 920 megapixel with the maximum at 11,282 megapixel. The file contains three levels of multi-scale representation, and the average file size is 202 MB with the maximum of ~2GB.

4.5 Performance

Annotation outcome and processing speed are the two main factors that contribute to the performance of the framework: annotation quality and annotation speed. The level of precision in annotation could be categorized into bounding boxes and circles, polygons, and pixel-level annotation as shown in Figure 8 (a-c) respectively. The mentioned annotation methods are ordered from greater simplicity with less precision to more complexity with more precision. Comparing OpenHI framework to other similar software, our proposed method can support pixel-level annotation precision while the other software (Gutman *et al.*, 2013, 2017; Litjens, 2015; Bankhead *et al.*, 2017) cannot. In comparison, OpenHI could significantly improve the annotation quality thus achieve better annotation outcome. Other than the annotation quality, the processing speed which contribute directly to the interactivity of the framework can reflect annotation efficiency. In our proposed method, the framework will consume more processing time when large area of the image

is needed to be viewed. To demonstrate the robustness of the proposed method, it was tested based on the real-world usage as described later in this section.

The framework was tested on two environments: server and personal computer. In the server environment, an Intel(R) Xeon(R) CPU E5-2650 v4 (2.20GHz), 1266 MHz with a total of 48 cores and 256 gigabytes of RAM. The testing repository occupies 500 gigabytes of storage. The host operating system is Ubuntu 16.04 LTS. However, for a single user, the environment at the scale of personal computer is enough, the minimum requirement for the host that we have tested with is Intel Core i7 (1.7GHz) with 2 cores and 8 gigabytes of RAM excluding image pre-processing module due to memory limitation (see image pre-processing section).

A 250 MB WSI was used during the processing speed testing. The test was performed on a machine with single user requirement. The response time is varied by processing resolution which is specified by the zooming level that the user has queried. For instance, if the user request to view a small area of the WSI or use high magnification, the processing resolution will be low. To view the image and sub-regions clearly, the user will need to magnify the WSI so that only less than 15 megapixels of resolution are needed to be processed. The examples of the viewing image on 800-by-460 pixel viewer at different processing resolution is illustrated in Figure 9. In most cases, comfortable viewing magnification is at 3 to 8 megapixels processing resolution where it is suitable for annotation task. The response time is shown in Figure 10 where the average processing time will take around 300 ms with the maximum at 580 ms which is almost unnoticeable and responsive enough to perform annotation task efficiently.

To conclude about the performance of the framework, in term of annotation outcome, the framework can achieve superior annotation precision compare to other similar software while maintaining decent processing speed. Thus, OpenHI can be used to perform efficiently in large-scale annotation of WSIs.

5 Discussion

While OpenHI do support multiple user during the annotation on the same tissue slide, interpretation of multi-expert annotation has not been included in the framework yet. Methods to effectively combine multi-expert annotation should be investigated. There are two aspects consolidate multiple annotations: spatial and grading degree. Similarly, methods to validate inter-rater reliability should be established as well. Using collaborative platform as an infrastructure, real-time validation should be possible. Moreover, discordance between pathologists do happen even in glass slides setting (Joseph P Houghton *et al.*, 2014; Joseph P. Houghton *et al.*, 2014; Carney *et al.*, 2016). Some are caused by pathologists' experiences and some do not have clinical consequences, being able to distinguish genuine pathological disagreements from the ones influenced by the annotation framework should be useful.

Further acceleration of annotation can be made by building on OpenHI using different approaches: provide machine-learning based suggestion, improve segmentation with nuclei localization, provide basic analysis the slide region, etc. Machine-learning based suggestion system is a promising approach to improve annotation speed since annotators would only need to confirm majority of annotations and correct some. But this approach may induce biases and affect annotation quality. The current pre-segmentation method

could be improved by localizing areas with nuclei and only pre-segment those areas. It could save some efforts since annotators can only focus on areas with cells. Basic information about the viewing area could be provided such as nuclei density, nuclei size, or ratio of H&E color.

Generally, pixels in WSI files produced from whole-slide scanner originate from a light source, passing through objective lenses and recorded with an image sensor array. There are many parameters during each step of the light pathway, and they could be between scanners from different manufacturer. OpenHI calculates virtual magnification factor based on image pixel size, 0.25 micron/pixel in our testing dataset. This parameter was not provided in the metadata of every WSI file format. A recent regulatory guideline for whole-slide imaging system (Zarella *et al.*, 2018) by US Food and Drug Administration does not specify that WSI files should provide this parameter. Instead, it focuses on a closed whole-slide imaging system spanning from scanners to computer displays which is not good for data interoperability. Furthermore, simulating virtual magnification should bear in mind that image presented on computer displays and perceived by annotators may be different.

Most of the processing time is spent on retrieving and compiling different layers of images to visualize the current state of a particular part of the WSI. Another problem with the current WSI file formats is that we can only read from them. Tackling these problems, we suggest that a new file format with support to store annotation data may be needed. This could be a completely new file format based on multi-scale image techniques or deep zooming technology.

6 Conclusion

Digitalized histopathological images are increasing in a fast pace with continuous health informatics development around the world. The images present phenotypes of tumors at cellular level and may support the association study with genotypes from genome sequencing data. OpenHI may accelerate precise creation of phenotype annotations with semantic meaning in the images. Additionally, the framework utilizes web technology, therefore is capable of scaling and collaborative annotation which is a foundation of crowed-sourcing to create large-dataset. As a result, large-scale datasets with precise and semantically rich annotations which is suitable for training computational model could be efficiently created. The framework is open source and could be easily extended and implemented into a clinical decision-making workflow (Kothari *et al.*, 2013). It also can be easily configured at the server-end module for data scientists to adapt different diagnosis standards, e.g. various cancer sub-types or gradings.

This work does not only propose a framework for annotating WSI which overcomes technological challenges to read and store extremely large histopathological image, but also discuss procedures to combine semantic meaning within existing cancer grading systems onto digital slides. The proposed process along with the framework complies with usual pathological routine thus pathologists can effectively, efficiently annotate using the framework to achieve high quality biomedical dataset.

Large-scale datasets with precise annotations may be efficiently created by the framework. Artificial intelligent methods, for example, based on statistical machine learning, could benefit from the rich features in the data and move forward to practically assist the pathologist's routine laboratory work. Such pipeline could also provide a solution to imminent issue such as misgrading which could lead to misdiagnosis and to provide a

good foundation for the future development of phenotype-genotype or multi-omics associations (Cooper *et al.*, 2015).

References

Achanta, R. *et al.* (2012) 'SLIC superpixels compared to state-of-the-art superpixel methods', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), pp. 2274–2281. doi: 10.1109/TPAMI.2012.120.

B, E. B. *et al.* (2017) 'Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer', *JAMA*, 318(22), pp. 2199–2210. Available at: http://dx.doi.org/10.1001/jama.2017.14585.

Bankhead, P. *et al.* (2017) 'QuPath: Open source software for digital pathology image analysis', *Scientific Reports*. doi: 10.1038/s41598-017-17204-5.

Carney, P. A. *et al.* (2016) 'Identifying and processing the gap between perceived and actual agreement in breast pathology interpretation', *Modern Pathology*. Nature Publishing Group, 29(7), pp. 717–726. doi: 10.1038/modpathol.2016.62.

Clunie, D. *et al.* (2018) 'Digital Imaging and Communications in Medicine Whole Slide Imaging Connectathon at Digital Pathology Association Pathology Visions 2017', pp. 1–4. doi: 10.4103/jpi.jpi.

Cooper, L. A. D. *et al.* (2015) 'Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images', *Laboratory Investigation*. Nature Publishing Group, 95(4), pp. 366–376. doi: 10.1038/labinvest.2014.153.

Delahunt, B. *et al.* (2013) 'The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters', *American Journal of Surgical Pathology*, 37(10), pp. 1490–1504. doi: 10.1097/PAS.0b013e318299f0fb.

Dong, F. *et al.* (2014) 'Crowdsourcing Image Annotation for Nucleus Detection and Segmentation in Computational Pathology: Evaluating Experts, Automated Methods, and the Crowd', pp. 294–305. doi: 10.1142/9789814644730_0029.

El-Gabry, E. A., Parwani, A. V. and Pantanowitz, L. (2014) 'Whole-slide imaging: widening the scope of cytopathology', *Diagnostic Histopathology*. Elsevier, 20(12), pp. 456–461. doi: 10.1016/j.mpdhp.2014.10.006.

Gleason, D. F. (1992) 'Histologic grading of prostate cancer: A perspective', *Human Pathology*, 23(3), pp. 273–279. doi: 10.1016/0046-8177(92)90108-F.

Goode, A. *et al.* (2013) 'OpenSlide: A vendor-neutral software foundation for digital pathology', *Journal of Pathology Informatics*, 4(1), p. 27. doi: 10.4103/2153-3539.119005.

Gutman, D. A. *et al.* (2013) 'Cancer digital slide archive: An informatics resource to support integrated in silico analysis of TCGA pathology data', *Journal of the American Medical Informatics Association*. Oxford University Press, 20(6), pp. 1091–1098. doi:

10.1136/amiajnl-2012-001469.

Gutman, D. A. *et al.* (2017) 'The digital slide archive: A software platform for management, integration, and analysis of histology for cancer research', *Cancer Research*, 77(21), pp. e75–e78. doi: 10.1158/0008-5472.CAN-17-0629.

Herrmann, M. D. *et al.* (2018) 'Implementing the DICOM Standard for Digital Pathology.', *Journal of pathology informatics*. Medknow Publications and Media Pvt. Ltd., 9(1), p. 37. doi: 10.4103/jpi.jpi_42_18.

Houghton, Joseph P *et al.* (2014) 'Concordance between digital pathology and light microscopy in general surgical pathology: a pilot study of 100 cases.', *Journal of clinical pathology*. BMJ Publishing Group, 67(12), pp. 1052–5. doi: 10.1136/jclinpath-2014-202491.

Houghton, Joseph P. *et al.* (2014) 'Concordance between digital pathology and light microscopy in general surgical pathology: A pilot study of 100 cases', *Journal of Clinical Pathology*, 67(12), pp. 1052–1055. doi: 10.1136/jclinpath-2014-202491.

Jukić, D. M. *et al.* (2011) 'Clinical examination and validation of primary diagnosis in anatomic pathology using whole slide digital images', *Archives of Pathology and Laboratory Medicine*, 135(3), pp. 372–378. doi: 10.1043/2009-0678-OA.1.

Keen, J. C. and Moore, H. M. (2015) 'The Genotype-Tissue Expression (GTEx) Project: Linking Clinical Data with Molecular Analysis to Advance Personalized Medicine.', *Journal of personalized medicine*. Multidisciplinary Digital Publishing Institute (MDPI), 5(1), pp. 22–9. doi: 10.3390/jpm5010022.

Kothari, S. *et al.* (2013) 'Pathology imaging informatics for quantitative analysis of whole-slide images', *Journal of the American Medical Informatics Association*. Oxford University Press, 20(6), pp. 1099–1108. doi: 10.1136/amiajnl-2012-001540.

Kurc, T. *et al.* (2015) 'Scalable analysis of Big pathology image data cohorts using efficient methods and high-performance computing strategies', *BMC Bioinformatics*. BMC Bioinformatics, 16(1), pp. 1–21. doi: 10.1186/s12859-015-0831-6.

Litjens, G. (2015) 'Automated Slide Analysis Platform (ASAP)'.

Della Mea, V. *et al.* (2017) 'SlideJ: An ImageJ plugin for automated processing of whole slide images', *PLoS ONE*, 12(7), pp. 1–9. doi: 10.1371/journal.pone.0180540.

Mercan, C. *et al.* (2018) 'Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images', *IEEE Transactions on Medical Imaging*, 37(1), pp. 316–325. doi: 10.1109/TMI.2017.2758580.

Mukhopadhyay, S. *et al.* (2018) 'Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology of 1992 Cases (Pivotal Study)', *American Journal of Surgical Pathology*, 42(1), pp. 39–52. doi: 10.1097/PAS.0000000000000948.

'OpenSeadragon Project' (2013).

Schindelin, J. *et al.* (2015) 'The ImageJ ecosystem: An open platform for biomedical image analysis', *Molecular Reproduction and Development*, 82(7–8), pp. 518–529. doi: 10.1002/mrd.22489.

Sellaro, T. *et al.* (2013) 'Relationship between magnification and resolution in digital pathology systems', *Journal of Pathology Informatics*, 4(1), p. 21. doi: 10.4103/2153-3539.116866.

Singh, R. *et al.* (2011) 'Standardization in digital pathology: Supplement 145 of the DICOM standards', *Journal of pathology informatics*. Medknow Publications Pvt Ltd, 2, p. 23. doi: 10.4103/2153-3539.80719.

Tomczak, K., Czerwińska, P. and Wiznerowicz, M. (2015) 'The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge.', *Contemporary oncology (Poznan, Poland)*. Termedia Publishing, 19(1A), pp. A68-77. doi: 10.5114/wo.2014.47136.

True, L. D. and Jordan, C. D. (2008) 'The cancer nuclear microenvironment: Interface between light microscopic cytology and molecular phenotype', *Journal of Cellular Biochemistry*, 104(6), pp. 1994–2003. doi: 10.1002/jcb.21478.

Veta, M. *et al.* (2018) 'Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge', pp. 1–22. Available at: https://arxiv.org/abs/1807.08284.

van der Walt, S. *et al.* (2014) 'scikit-image: image processing in Python', *PeerJ*, 2, p. e453. doi: 10.7717/peerj.453.

Xing, F., Xie, Y. and Yang, L. (2016) 'An automatic learning-based framework for robust nucleus segmentation', *IEEE Transactions on Medical Imaging*. IEEE, 35(2), pp. 550–566. doi: 10.1109/TMI.2015.2481436.

Xu, Y. *et al.* (2017) 'Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features', *BMC Bioinformatics*. BioMed Central, 18(1), p. 281. doi: 10.1186/s12859-017-1685-x.

Zarella, M. D. *et al.* (2018) 'A Practical Guide to Whole Slide Imaging: A White Paper From the Digital Pathology Association', *Archives of Pathology & Laboratory Medicine*. the College of American Pathologists, p. arpa.2018-0343-RA. doi: 10.5858/arpa.2018-0343-RA.

Zhou, N. *et al.* (2017) 'Large scale digital prostate pathology image analysis combining feature extraction and deep neural network', pp. 1–14.

Tables

Table 1   WHO/ISUP grading standard for renal cell carcinoma

Grade	Description
Grade 1	Having inconspicuous or absent nucleoli at x400 magnification
Grade 2	Nucleoli should be distinctly visible at x400, but inconspicuous or invisible at x100 magnification
Grade 3	Nucleoli should be distinctly visible at x100 magnification
Grade 4	Tumors should encompass tumors with rhabdoid or sarcomatoid differentiation or those containing tumor giant cells or showing extreme nuclear pleomorphism with clumping of chromatin

Table 2   The configurable parameters of the framework for extending to different image types and situations

Parameter	Time of configuration	Description
Pre-segmentation level	Annotation	This parameter determine the size of the sub-region as pre-processed in the pre-processing step.
Number of grading level	Annotation	In different cancer type and grading standards. The number of possible grades is different. Annotators can choose appropriate grade for each sub-region acoording to available grades.
Pre-segmentation sub-region pixel density	Experiment setup	This parameter could be mutual agreement between the annotator and the analyst before the pre-segmentation begins. Tuning of this parameter will allow efficient annotation process.
Grading tier	Experiment setup	Different oncological types have different grading system. The annotation software could support different grading level.
Viewer size	Experiment setup	The viewer size could be fixed to a specific size or make it adaptive to the annotator's monitor screen size.

Figures

Figure 1   Comparison of how the intended region for selection (a) can be selected by a tiled-based (b) and superpixel-based (c) segmentation. Sub-region selection in (c) is more efficient since it needs less selection and achieve more accurate annotation boundary.
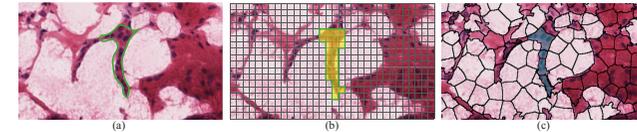
Figure 2   Structure of the framework with WSI data flow from original image to pre-segmentation and annotation.
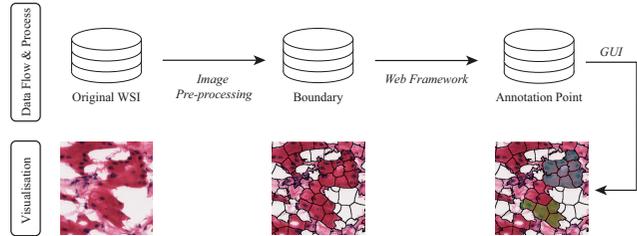
Figure 3   A part of a WSI (a) is pre-segmented using SLIC superpixel algorithm at 6100 (b), 610 (c), and 60 (d) pixel/sub-region.
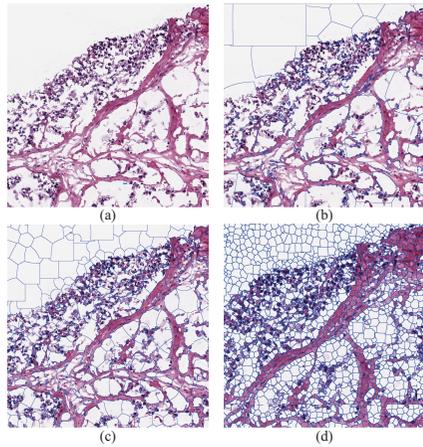
Figure 4   Illustration of superpixel segmentation resulting in (a) over-segmentation, (b) under-segmentation, and (c) appropriate sub-region size.

(a) over-segmentation   (b) under-segmentation   (c) appropriate sub-region size
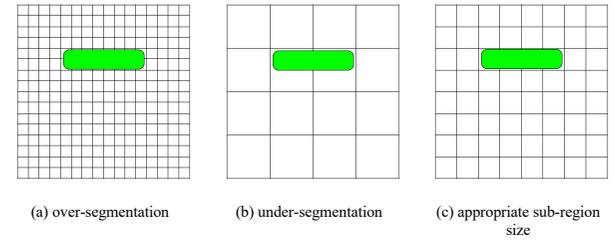
Figure 5   Process for creating viewing image by compiling original WSI, sub-region boundary, and existing annotation.

(b) WSI

(a) database server   (c) sub-region boundary   Server-end module   (e) viewing image

(d) visualized annotation
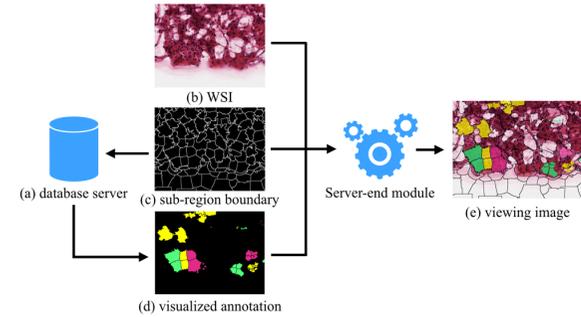
Figure 6   The GUI of the framework with a sample image showing WSI viewer (left), virtual magnification indicator (bottom-left), control panel (top-right) containing tumour grading selector, pre-segmentation level selector, and slide information panel where accompanying information of a slide is shown.
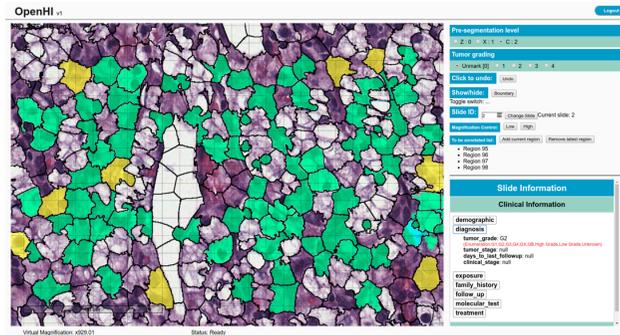
Figure 7   Comparison of different sub-region selection method where (a) is done by using polygons, (b) is region-by-region selection, and (c) demonstrate continuous selection across multiple sub-regions suing hold-and-drag method.

Figure 10   Processing time in different processing resolution ranging from 3.3 to 15 megapixel, corresponding to images (a) to (d) in Figure 9
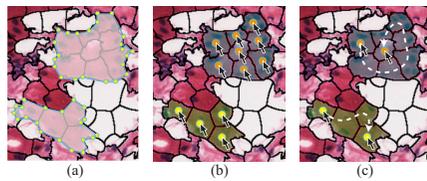
Figure 8   Different level of annotation precision from different image annotation methods from the least precise to the most precise boundary including (a) bounding box, (b) polygon, and (c) pixel-level annotation.
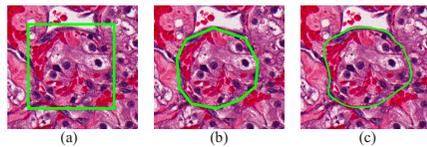
Figure 9   Example of viewing image in different processing solution ranging from 3.3 (a), 5.1 (b), 9.4 (c), and 15 (d) megapixel.