OpenHI - An open source framework for annotating histopathological image

Pargorn Puttapirat[1,2], Haichuan Zhang[1,2], Yuchen Lian[1,2], Chunbao Wang[3], Xiangrong Zhang[4], Lixia Yao[5], Chen Li[12]*
[1] Shaanxi Province Key Laboratory of Satellite and Terrestrial Network Tech. R&D，
Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[2] School of Electronic and Information Engineering，Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[3] Department of Pathology, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China
[4] Institute of Intelligent Information Processing, Xidian University, Xi'an, Shaanxi 710071, China
[5] Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA
*Corresponding author: cli@xjtu.edu.cn

***Abstract*—Histopathological images carry informative cellular visual phenotypes and have been digitalized in huge amount in medical institutes. However, the lack of software for annotating the specialized images has been a hurdle of fully exploiting the images for educating and researching, and enabling intelligent systems for automatic diagnosis or phenotype-genotype association study. This paper proposes an open-source web framework, OpenHI, for the whole-slide image annotation. The proposed framework could be utilized for simultaneous collaborative or crowd-sourcing annotation with standardized semantic enrichment at a pixel-level precision. Meanwhile, our accurate virtual magnification indicator provides annotators a crucial reference for deciding the grading of each region. In testing, the framework can responsively annotate the acquired whole-slide images from TCGA project and provide efficient annotation which is precise and semantically meaningful. OpenHI is an open-source framework thus it can be extended to support the annotation of whole-slide images from different source with different oncological types. The framework may facilitate the creation of large-scale precisely annotated histopathological image datasets.**

I. Introduction

Large-scale histopathological images present rich information about the microenvironment of cancer which are crucial for interpreting the corresponding genotypes in omics data. Since pathology is one of the last medical specialties to be digitized [1], there has been limited software and tools that helps pathologists manage and analyze extra-large digital scans of glass slides – namely, whole-slide image (WSI). 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 annotation framework for WSI, which 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.

To make a cancer diagnosis in various oncological types, pathologists would need to analyze one or multiple tissue samples from a patient's biopsy and decide the grading of each specific samples. This tedious procedure is manually carried out with a bright field light microscope [2]. In recent years, there has 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, by detecting and classifying the cancerous area, the WSI with high throughput and high precision to help reduce the workload of the pathologists [3]–[6]. However, these works were derived and tested on datasets with finite size and variability due to limited availability of public datasets. With our proposed framework, rapid creation of such datasets could be accomplished, thus sophisticate computational approach that holistically analyze the WSIs [7] may lead to a better grading decision.

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 configuration at the annotator end (pathologist) and maximize configurability at the data scientists 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.

II. Related work

In cancer research, large-scale raw histopathology image repositories have been made publicly available by TCGA [8] and GTEx [9] project. 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 data for automated methods is the lack of annotations. In recent years, there are enriched WSI datasets used in competitions [10], [11], however, they are limited by public availability, size, variability in cancer type, or spatially precise annotation. An annotation software capable of precise visual annotation and semantic information enrichment is highly demanded.

WSI files are hard to read because WSIs cannot be saved in standard image format due to its unusually high resolution. In 2013, [1] 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 [12], [13], stand-alone software such as QuPath [14] and ASAP [15], and extension, SlideJ [16], to ImageJ [17]. Around the same time, OpenSeadragon (OSD) library [18] was introduced as a web-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 and analyze the WSIs. In 2017, QuPath [14] 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.

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.

III. Material and Methods

The proposed framework was designed to be implemented on a web server therefore it could simultaneously be accessed via a web-browser by multiple users. 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 Fig. 1 by a widely used image segmentation method called SLIC superpixel [19]. 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 at different zooming level based on OpenSlide being utilized as the others did [12], [14], [15].

A. The framework

The framework consists of three main components along with a MySQL server to store the annotation coordinates. The main components include image pre-processing of the WSIs, the web framework, and the GUI. Three types of data are stored in alongside the framework which is the WSI files, the sub-region boundary matrices, and the annotation coordinates. Fig. 2 illustrates each components and the data flow between them.

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 [20]. 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 [16]. The SLIC Superpixel [19] 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. To our knowledge, there is 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 [21]. By incorporating superpixel segmentation with the WSIs, we have established a new ROI selection method in digital pathology.

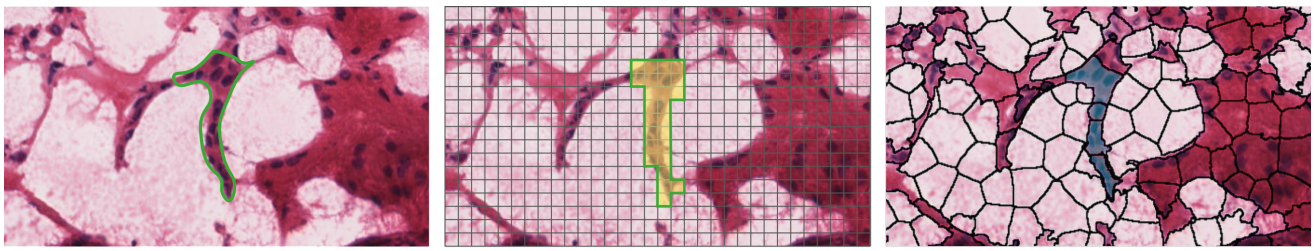

Fig. 1. Comparison of how the intended region for selection (a) can be selected by a tiled-based (b) and superpixel-based (c) segmentation. Sub-region selection in (c) is more efficient since it needs less selection and achieve more accurate annotation boundary.

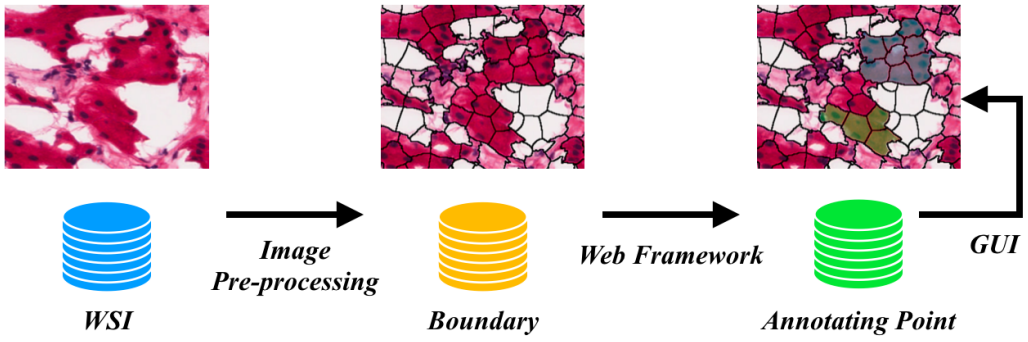

Fig. 2. Structure of the framework with WSI data flow from original image to pre-segmentation and annotation.

The long lasting issue with superpixel algorithm in image segmentation is of making a decision about the number of final segmented sub-region which 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 for all images. To tackle this problem, we calculate the number of superpixeled sub-regions ($N_{superpixel}$) by specifying desired average sub-region size (P_{sub}), thus the size of the sub-region will be consistent throughout the annotation project as shown in (1) where P_{total} is the WSI resolution. In practice, this number could range from around 6000 to 50 pixels/sub-region providing that the WSI has been scanned with 20x magnification lenses and it should increase or decrease with the magnification. The example of a portion of the image pre-segment with superpixel algorithm is shown in Fig. 3.

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

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, our framework provides users an option to calibrate at multiple pre-segmentation levels. 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.

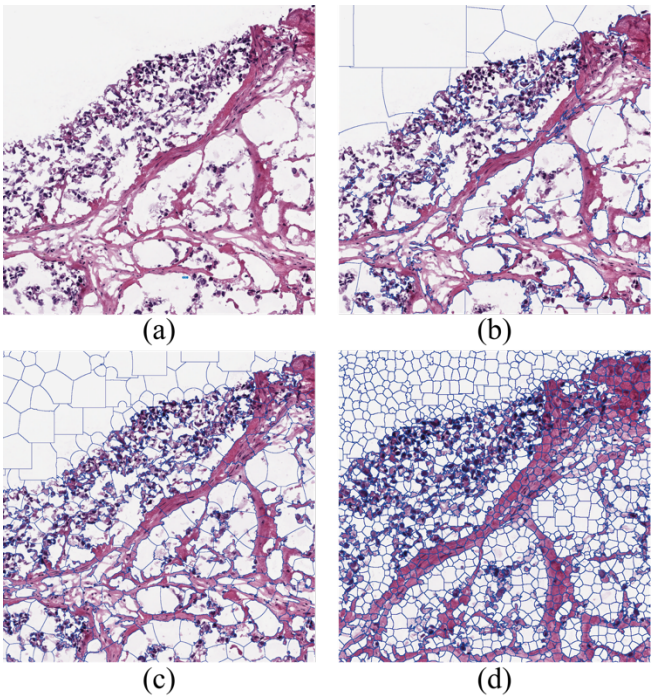

Fig. 3. A part of a WSI is pre-segmented using superpixel algorithm at 6100, 610, and 60 pixel/sub-region.

At the end of pre-processing, matrices containing sub-region boundary information are stored using a binary image format. This image with boundary information 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.

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 processing includes the generation of sub-region boundary, visualization of the annotated sub-regions, and annotation coordinates recording and deletion.

As the use of OSD with OpenSlide has been demonstrated in [1] in 2013, we extend the usage for WSI annotation. Our framework has attained the capability of OSD and OpenSlide to view a WSI with smooth zooming and panning experience 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. Aside from conveniently choosing the sub-region, the GUI should provide some indicator to approximate the zooming power of current digital image state, this is also called virtual magnification. Since the gradings of some cancer type such as WHO/ISUP kidney clear cell carcinoma annotation standard [22] (see Table 2) require the annotators to take magnification power as a part of decision. This functionality is crucial for the grading system 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. 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. 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. The scanner sensor pixel size is often not included in the metadata of the WSI file while magnification power is specified, thus it could be referenced back to the scanner manufacturer about scanner's resolving power correspond to the specific magnification factor, and the sensor pixel size could be calculate by T. Sellaro et al method where there are calculation sample in their work [23]. To obtain the objective magnification (M_{obj}). Finally, the total virtual magnification (M_{total}) could be calculated by (2).

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

3) Graphic user interface

The web-based GUI of our proposed framework is comprised of the main WSI viewer, virtual magnification indicator, and menus for annotation configuration as shown in Fig. 4. The viewer, based-on OSD, enables an annotator to navigate around the WSI and the menu allows the user 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 and existing annotations. An alternative way to adjust these parameters is the usage of keyboard hotkeys that could speed up the annotation once the user gets familiar with the system.

The GUI also has indicators which could assist the annotator to make a grading decision such as current virtual magnification factor, current resolving power, and the basic information about the tissue slide.

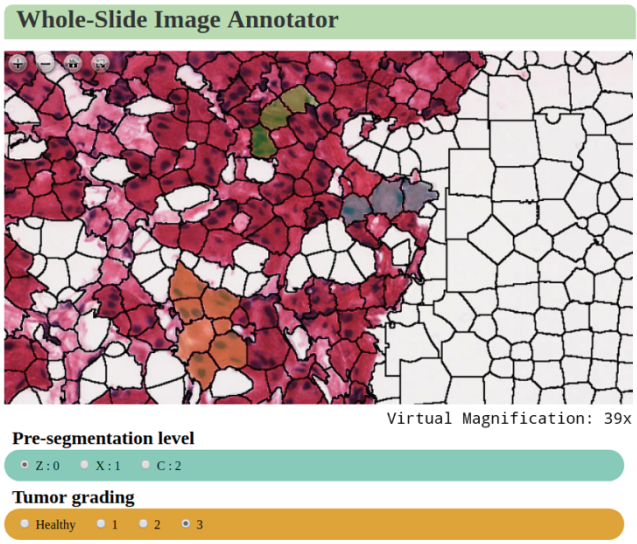

Fig. 4. The GUI of the framework with a sample image showing virtual magnification indicator and options for pre-segmentation level and tumor grading switch.

4) The framework's data model

To maximize the granularity of the annotation and keeping the utility of collected data. The framework will orient the data collection around the annotation coordinates by their properties including the exact chosen coordinate x and y, assigned tumor grading, pre-segmentation level, time of annotation, and the annotator's 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.

To maximize the extendibility of this annotation software, it is designed to be highly customizable. Some parameters are designed for the annotators (pathologist) 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.

B. Dependencies

The image processing component of the framework was implemented by MATLAB with image processing toolbox to perform SLIC superpixel algorithm [19]. Noted that it could also be done by other open-source image processing library such as OpenCV or scikit-image [24]. The web framework is written in Python. We have adopted Flask 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 is used in visualization. In the web-based GUI, OSD is used as the viewer and annotating coordinates and user interaction to navigate around the WSI are acquired through OSD. Lastly, the annotation information is stored in a MySQL server.

IV. Results

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 1, 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. The enabled online real-time collaboration converges effort from pathologists to annotate and cross-validate large-scale images.

TABLE I. 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 in pro-processed
Number of grading level	Annotation	In different cancer type and grading standards. The level of grading is different.
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 cancer sub-type has 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.

A. 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 Fig. 5 (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 Fig. 5 (a). The comparison can be seen in Fig. 5. Besides the easy sub-region selection, the user can also select gradings easily via the GUI next to the

viewer as shown in Fig. 4 or by hotkeys. Additionally, in the case that the annotator mistakenly select the region, the framework can revert to previous step or the annotator can deselect some sub-regions.

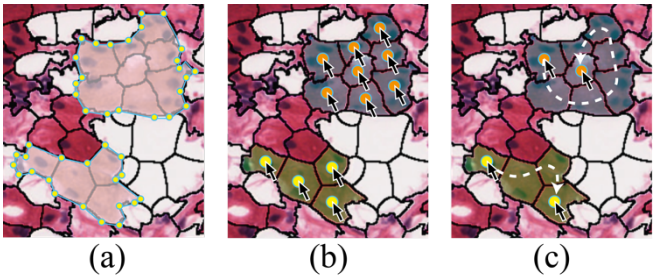

Fig. 5. Comparison of different sub-region selection method where (a) is done by using polygons, (b) is region-by-region selection, and (c) demonstrate continuous selection across multiple sub-regions suing hold-and-drag method.

B. *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 [25], [26]. 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 viewer lack a precise virtual magnification to choose in virtual slide viewing due to many factors. Furthermore, grading for this type of cancer also has major discordance between WSI and conventional microscopy in [25] 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.

TABLE II. WHO/ISUP GRADING STANDARD

Grade	Description
Grade 1	Having inconspicuous or absent nucleoli at x400 magnification
Grade 2	Nucleoli should be distinctly visible at 400, 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

C. *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) [1].

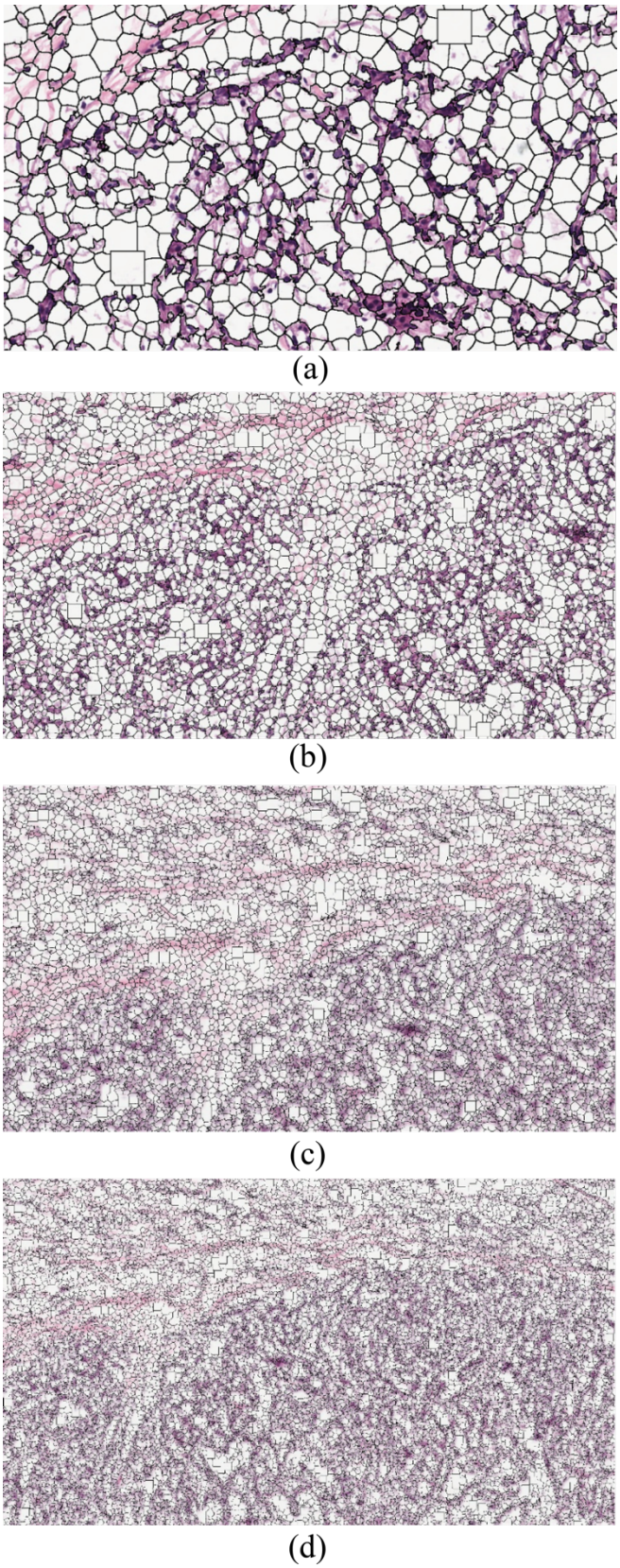

Fig. 6. Example of viewing image in different processing solution ranging from 3.3 (a), 5.1 (b), 9.4 (c), and 15 (d) megapixel.

Our software is free and open-source, It is available at 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.

D. Data acquisition

In software development and testing, we use WSIs directly downloaded from TCGA data repository [8]. Thus the testing environment, WSI format used in our proposed framework is Aperio (.svs) file. The images are scanned with 20x magnification with resolving power of 0.5 micron/pixel [20]. 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.

E. Performance

The framework was tested on 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 current repository occupies 500 gigabytes of storage. The host operating system is Ubuntu 16.04 LTS. However, for single user, 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 due to memory limitation (see image pre-processing section).

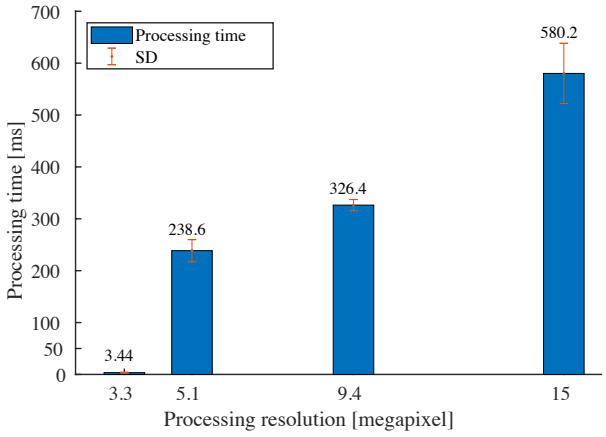

Fig. 7. *Processing time in different processing resolution ranging from 3.3 to 15 megapixel, corresponding to images (a) to (d) in Fig. 6*

The 250 MB WSI is used during the performance testing where we test the response time of the framework on a machine with minimum requirement. The response time is varied by processing resolution which is specified by the zooming level that the user has requested. 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 megapixel 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 Fig. 6. In most cases, we find viewing the image at 3 to 8 megapixel processing resolution is suitable for annotation task. The response time is shown in Fig. 7 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.

V. Conclusion

Digitalized histopathological images are increasing in a fast pace with continuous health informatics development around the world. The images presents phenotypes of tumors at cellular level and may support the association study with genotypes from sequence 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 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 [21]. It also can be easily configurable at the back-end for the data scientist to adapt different diagnosis standards, e.g. various cancer sub-types or gradings.

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-omic associations [20].

Acknowledgment

This work has been supported by "The Fundamental Theory and Applications of Big Data with Knowledge Engineering" under the National Key Research and Development Program of China with grant number 2016YFB1000903; National Natural Science Foundation of China (Grant NO: 61772409); the Fund of Ministry of Education and China Mobile; Project of China Knowledge Centre for Engineering Science and Technology; Innovation team of Ministry of Education (IRT_17R86); Innovative Research Group of the Nation Natural Science Foundation of China (61721002); Ministry of Education-Research Foundation of China Mobile Communication Corp(MCM20160404); Professor Chen Li's Recruitment Program for Young Professionals of "the Thousand Talents Plan".

References

[1] A. Goode, B. Gilbert, J. Harkes, D. Jukic, and M. Satyanarayanan, "OpenSlide: A vendor-neutral software foundation for digital pathology," J. Pathol. Inform., vol. 4, no. 1, p. 27, 2013.

[2] J. P. Houghton et al., "Concordance between digital pathology and light microscopy in general surgical pathology: A pilot study of 100 cases," J. Clin. Pathol., vol. 67, no. 12, pp. 1052–1055, 2014.

[3] T. Kurc et al., "Scalable analysis of Big pathology image data cohorts using efficient methods and high-performance computing strategies," BMC Bioinformatics, vol. 16, no. 1, pp. 1–21, 2015.

[4] N. Zhou, A. Fedorov, F. Fennessy, R. Kikinis, and Y. Gao, "Large scale digital prostate pathology image analysis combining feature extraction and deep neural network," pp. 1–14, 2017.

[5] Y. Xu et al., "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," BMC Bioinformatics, vol. 18, no. 1, p. 281, Dec. 2017.

[6] C. Mercan, S. Aksoy, E. Mercan, L. G. Shapiro, D. L. Weaver, and J. G. Elmore, "Multi-Instance Multi-Label Learning for Multi-Class

Classification of Whole Slide Breast Histopathology Images," IEEE Trans. Med. Imaging, vol. 37, no. 1, pp. 316–325, 2018.

[7] E. A. El-Gabry, A. V. Parwani, and L. Pantanowitz, "Whole-slide imaging: widening the scope of cytopathology," Diagnostic Histopathol., vol. 20, no. 12, pp. 456–461, Dec. 2014.

[8] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge.," Contemp. Oncol. (Poznan, Poland), vol. 19, no. 1A, pp. A68-77, 2015.

[9] J. C. Keen and H. M. Moore, "The Genotype-Tissue Expression (GTEx) Project: Linking Clinical Data with Molecular Analysis to Advance Personalized Medicine.," J. Pers. Med., vol. 5, no. 1, pp. 22–9, Feb. 2015.

[10] M. Veta et al., "Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge," pp. 1–22, 2018.

[11] E. B. B, M. Veta, J. van D. P, and et al, "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," JAMA, vol. 318, no. 22, pp. 2199–2210, Dec. 2017.

[12] D. A. Gutman et al., "Cancer digital slide archive: An informatics resource to support integrated in silico analysis of TCGA pathology data," J. Am. Med. Informatics Assoc., vol. 20, no. 6, pp. 1091–1098, Nov. 2013.

[13] D. A. Gutman et al., "The digital slide archive: A software platform for management, integration, and analysis of histology for cancer research," Cancer Res., vol. 77, no. 21, pp. e75–e78, 2017.

[14] P. Bankhead et al., "QuPath: Open source software for digital pathology image analysis," Sci. Rep., 2017.

[15] G. Litjens, "Automated Slide Analysis Platform (ASAP)." 2015. (https://github.com/computationalpathologygroup/ASAP)

[16] V. Della Mea, G. L. Baroni, D. Pilutti, and C. Di Loreto, "SlideJ: An ImageJ plugin for automated processing of whole slide images," PLoS One, vol. 12, no. 7, pp. 1–9, 2017.

[17] J. Schindelin, C. T. Rueden, M. C. Hiner, and K. W. Eliceiri, "The ImageJ ecosystem: An open platform for biomedical image analysis," Mol. Reprod. Dev., vol. 82, no. 7–8, pp. 518–529, 2015.

[18] "OpenSeadragon Project." 2013. (https://github.com/openseadragon/openseadragon)

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2281, 2012.

[20] L. A. D. Cooper, J. Kong, D. A. Gutman, W. D. Dunn, M. Nalisnik, and D. J. Brat, "Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images," Lab. Investig., vol. 95, no. 4, pp. 366–376, 2015.

[21] S. Kothari, J. H. Phan, T. H. Stokes, and M. D. Wang, "Pathology imaging informatics for quantitative analysis of whole-slide images," J. Am. Med. Informatics Assoc., vol. 20, no. 6, pp. 1099–1108, Nov. 2013.

[22] B. Delahunt et al., "The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters," Am. J. Surg. Pathol., vol. 37, no. 10, pp. 1490–1504, 2013.

[23] T. Sellaro et al., "Relationship between magnification and resolution in digital pathology systems," J. Pathol. Inform., vol. 4, no. 1, p. 21, 2013.

[24] S. van der Walt et al., "scikit-image: image processing in Python," PeerJ, vol. 2, p. e453, 2014.

[25] S. Mukhopadhyay et al., "Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology of 1992 Cases (Pivotal Study)," Am. J. Surg. Pathol., vol. 42, no. 1, pp. 39–52, 2018.

[26] D. M. Jukić, L. M. Drogowski, J. Martina, and A. V. Parwani, "Clinical examination and validation of primary diagnosis in anatomic pathology using whole slide digital images," Arch. Pathol. Lab. Med., vol. 135, no. 3, pp. 372–378, 2011.