



## Original Research Article

## Whole slide image data utilization informed by digital diagnosis patterns



Kimberly Ashman<sup>a</sup>, Huimin Zhuge<sup>a</sup>, Erin Shanley<sup>a</sup>, Sharon Fox<sup>b</sup>, Shams Halat<sup>c</sup>, Andrew Sholl<sup>d</sup>, Brian Summa<sup>e,\*</sup>, J. Quincy Brown<sup>a,\*</sup>

<sup>a</sup> Tulane University, Department of Biomedical Engineering, New Orleans, LA 70118, USA

<sup>b</sup> LSU Health Sciences Center, Department of Pathology, New Orleans, LA 70112, USA

<sup>c</sup> Tulane School of Medicine, Tulane University Department of Pathology and Lab Medicine, New Orleans, LA 70112, USA

<sup>d</sup> Delta Pathology Group, Touro Infirmary, New Orleans, LA 70115, USA

<sup>e</sup> Tulane University, Department of Computer Science, New Orleans, LA 70118, USA

## ARTICLE INFO

## ABSTRACT

**Context:** Despite the benefits of digital pathology, data storage and management of digital whole slide images introduces new logistical and infrastructure challenges to traditionally analog pathology labs.

**Aims:** Our goal was to analyze pathologist slide diagnosis patterns to determine the minimum number of pixels required during the diagnosis.

**Methods:** We developed a method of using pathologist viewing patterns to vary digital image resolution across virtual slides, which we call variable resolution images. An additional pathologist reviewed the variable resolution images to determine if diagnoses could still be rendered.

**Results:** Across all slides, the pathologists rarely zoomed in to the full resolution level. As a result, the variable resolution images are significantly smaller than the original whole slide images. Despite the reduction in image sizes, the final pathologist reviewer could still provide diagnoses on the variable resolution slide images.

**Conclusions:** Future studies will be conducted to understand variability in resolution requirements between and within pathologists. These findings have the potential to dramatically reduce the data storage requirements of high-resolution whole slide images.

## Introduction

Digital pathology, or the acquisition, management, and analysis of digitized glass slides, provides many opportunities to improve the practice of pathology by enabling telemedicine, computer-aided diagnosis, and archiving of glass slides that may fade or break over time.<sup>1,2</sup> However, while digital radiology received FDA approval over 20 years ago<sup>3</sup> and rapidly became the standard of care, the adoption of digital pathology has lagged with the first digital pathology slide scanning system only receiving FDA clearance for primary surgical pathology diagnosis in 2017.<sup>4</sup> At present, digital pathology has not been broadly adopted for primary diagnosis in the US.

One reason for digital pathology's lack of widespread adoption is that uncompressed high-resolution pathology images can be 80 times the size of a digital radiograph or more.<sup>5</sup> In fact, for large specimens, such as radical prostatectomy specimens, an uncompressed 20x whole slide image (WSI) with 0.220  $\mu\text{m}/\text{pixel}$  resolution still can be over 90 000 x 110 000 pixels (10 gigapixels) and require nearly 30 gigabytes (GB) of storage. If such large specimens are digitized at the recommended 40x magnification (0.1  $\mu\text{m}/\text{pixel}$ ) for primary clinical diagnosis,<sup>6</sup> then the size of the uncompressed digital files can swell to 100 GB. In addition, because the number of slides

from a single radical prostatectomy case can range from 20 to over 100 slides,<sup>7</sup> this can cause data storage requirements to scale rapidly and possibly overwhelm existing data infrastructures. Large pathology centers may produce over 1 million digital slides<sup>8</sup> and around 1 petabyte (10<sup>6</sup> GB) of uncompressed data per year,<sup>9</sup> along with data redundancy and backup requirements.<sup>10</sup> Clearly, storage of these images introduces new logistical and infrastructure challenges to traditionally analog pathology labs.

One commonly used solution to this data problem is to store and view WSIs in compressed formats. There are 2 modes of compression: 'lossy' and 'lossless.' Lossless compression algorithms, such as LZW or lossless JPEG2000, are preferred in medical imaging because the information is not irreversibly changed during the conversion.<sup>11</sup> For instance, the FDA approved Philips system uses a modified lossless JPEG2000 compression scheme.<sup>12</sup> 'Lossy' algorithms, such as JPEG, can compress images to a greater degree, but the data, once converted, is unrecoverable. Given a particular compression quality level, the compression ratio (uncompressed image size over compressed image size) will vary based on image content.<sup>13</sup> Large amounts of compression can degrade image quality and create visual artifacts, such as blurry or 'blocky' pixels.<sup>14</sup> Compression ratios ranging from 10:1 to 20:1 for JPEG and 30:1 to 50:1 for JPEG2000 are commonly

\* Corresponding authors.

E-mail addresses: [bsumma@tulane.edu](mailto:bsumma@tulane.edu) (B. Summa), [jqbrown@tulane.edu](mailto:jqbrown@tulane.edu) (J.Q. Brown).

used, although there is no standard acceptable ratio for preserving diagnostic content.<sup>13,15,16</sup> Even after compressing 40x scans to 500 MB, Lujan et al. estimates that a non-redundant storage system for a lab with a throughput of 250 000 slides per year would cost around \$90k per year in data storage costs alone.<sup>6</sup> Also, unlike in digital radiology where raw images are produced using inherently digital acquisition systems, today in pathology glass slides still need to be created as the first step before digitization. Because the creation of glass slides is still part of the digital pathology workflow and primary diagnoses can be rendered on the glass slide directly using inexpensive microscopes, the use of digital imaging in pathology is an added expense. Such an added expense, driven primarily by instrumentation and data storage costs, may not ultimately be funded by hospital administration if it cannot easily be demonstrated to substantially improve the standard of care.

With an eye toward a “smarter” approach for data reduction to lower the barriers to adoption, we have developed a hypothesis based on the insight that during a diagnostic review, pathologists typically do not observe the entire specimen at the highest magnification available on the microscope, or the highest possible zoom-level available in the digital image viewer. In both analog (glass slide) and digital (WSI) review workflows, pathologists first view slides globally using low power objectives or low zoom levels as they navigate around the specimens and only switch to higher power objectives or zoom into smaller, specific areas based on targeted recognition of features of interest. We hypothesized that if we could capture the pathologist’s behavior while reviewing digital slides, that we could enable a form of digital slide archival that maintains image quality in a way that directly represents the information needed by the pathologist to render the clinical diagnosis, while relegating less diagnostically important regions as targets for lower resolution archival, or adaptive, targeted compression.

We further hypothesized that such adaptively archived images could be rendered subsequently as “variable resolution images” (VRIs) and still be diagnostically sufficient. Therefore, our goal for this study was to use a custom-developed web-based WSI viewer to track search path, visitation history, and resolution utilization spatiotemporally in digital images during clinical diagnosis, to use the resulting information to create VRIs that preserve the highest quality in diagnostically important regions, and to evaluate the reduction in permanent file size storage that would result with such an approach. To evaluate the feasibility of this approach for primary pathologist-directed WSI archival and secondary pathologist review, we then conducted a pilot study to evaluate the impact on diagnostic outcomes that would result from such an approach.

## Methods

### Case selection and image acquisition

We chose to test our approach in prostate cancer surgical pathology. Prostate cancer is the second most commonly diagnosed cancer in men worldwide, with over 1 million new cases diagnosed in 2018.<sup>17</sup> Treatment often involves the surgical removal of the prostate gland and tumor, called a radical prostatectomy. A typical radical prostatectomy case consists of 25 to upwards of 100 slides, each of which typically contains large tissue sections. Due to the subsequently large file sizes and storage requirements, radical prostatectomy represents a challenging case for routine digitization that would benefit from the data reduction approach described in this work. In this feasibility study, we recorded pathologists reviewing one radical prostatectomy case, which consisted of 25 slides. Using a slide scanner (Aperio, Leica Biosystems), we scanned the slides at 20x magnification (0.504  $\mu\text{m}/\text{pixel}$ ). The average uncompressed WSI size was 35 000 x 46 000 pixels or 1.6 gigapixels. The minimum WSI size was 0.64 gigapixels, and the maximum was 3.7 gigapixels.

### Digital slide interaction data collection

We converted each WSI into an “image pyramid,” or multi-resolution representation consisting of successively downsampled image tiles. By

displaying an image pyramid in the viewer, only image tiles in the field of view (FOV) at a particular zoom level, which we will describe later in relation to magnification, were rendered at one time. In other words, at a low zoom level, only a low-resolution image was loaded, and at the highest zoom, the original high-resolution tiles were loaded.<sup>18</sup>

We added mouse and scroll event handlers to a custom OpenSeadragon web-based multi-resolution viewer to capture the zoom level, image tiles, mouse coordinates, viewport bounds, and viewport coordinates as the user interacts with the image. These data are then saved to a log file along with their timestamps, which enables reconstruction of the search path, zoom-level utilization, and visitation history spatiotemporally with millisecond resolution.<sup>19</sup>

### Resolution analysis

Image resolution is the minimum resolvable distance between 2 objects. Resolution is limited either by diffraction of the imaging system or by the size and spacing of the pixels in the camera used to collect the digital image. The final displayed resolution of a digital WSI thus depends on the image display monitor (number and density of pixels the monitor can display), zoom level, scanning optical magnification and numerical aperture, and camera sensor resolution.<sup>18</sup> Perceived resolution depends on the human visual system and how far away the pathologist is from the monitor.<sup>14</sup> For this study, we only considered image zoom level and assumed other factors affecting resolution were consistent.

As mentioned above, the custom image viewer logged the viewport coordinates (i.e., the coordinates of the part of the image visible to the user on the screen), viewport size, image zoom level, mouse coordinates, and time during each session. Log entries were added when any interaction with the viewer, such as scroll or mouse movement, was detected. Image zoom is the ratio of the displayed image width to the original image width, which we can relate to image magnification. For instance, if the image zoom is 0.5 of an original image that is at 20x magnification, the image resolution that the pathologist views is 10x. We binned the resolution to 2.5x, 5x, 10x, or 20x according to the utilized image zoom per pixel for each log entry. Image zoom between the bin levels was rounded to the next highest level.

Next, to create “resolution maps” that visualize the highest utilized magnification or zoom level over the image, we created new arrays scaled to 1/10 of the original WSI size for each resolution level to encode and display visited viewport dimensions and coordinates. To create the VRIs, we first found the outlines of the visited areas. Using the point in polygon test,<sup>20</sup> for each 256 x 256-pixel tile pertaining to the original image size, we determined whether any of the 4 coordinates of the tile were within the viewport outlines. For each tile, we only stored the maximum resolution level visited (Fig. 1).

### Variable-resolution images

After determining each tile’s highest visited resolution level, we eliminated tiles in the image pyramids greater than the required resolution. For instance, if the pathologist viewed an area at a maximum of 10x resolution, we removed the corresponding four 20x image tiles in the pyramid hierarchy. If high-resolution image tiles are missing, the OpenSeadragon package automatically serves and upsamples the image tiles lower in the image pyramid.

### Evaluation of Variable Resolution Images (VRIs) for accurate secondary review

We recorded 2 expert board-certified pathologists with genitourinary surgical pathology expertise, referred to as Pathologists A and B, as they reviewed all 25 slides in the test case. From these reviews, we generated variable resolution images for each pathologist separately that corresponded to each of their specific viewing sessions. The pathologists were asked to conduct 4 tasks during the review of the 25 slides in the case: (1) evaluating the specimen margin for positive or negative margins, (2) checking for

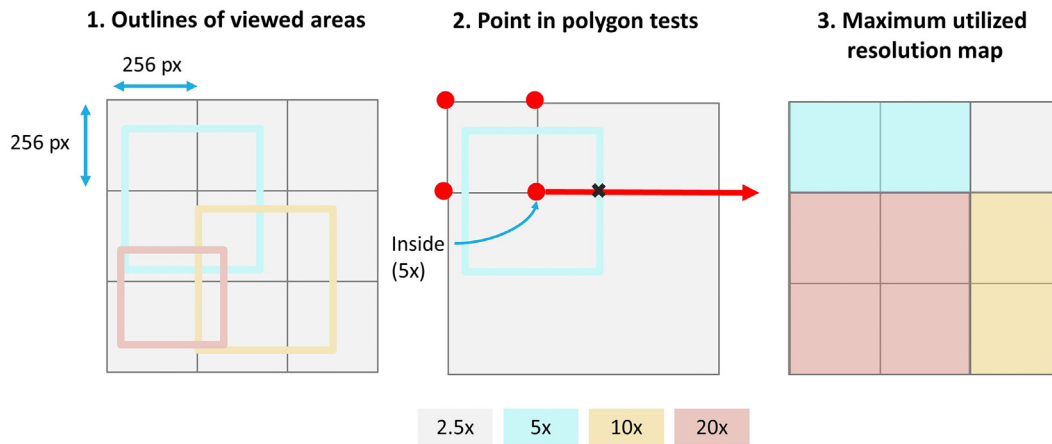


Fig. 1. Method for creating resolution maps according to the 256 x 256 image tiles using the point in polygon method

extraprostatic extension (EPE), (3) quantifying the tumor volume by outlining the tumor with a digital annotation pen, and (4) Gleason grading.

To perform an initial evaluation of this method for downstream diagnostic sufficiency, we then asked a third pathologist, referred to as Pathologist C, to review the VRIs. Our goal was to understand whether the VRIs contained enough information for an accurate secondary review despite variation in resolution due to Pathologist A or B’s specific viewing behaviors. First, we asked Pathologist C to review the 25 VRIs produced from only Pathologist A’s recordings and perform 3 tasks per slide: (1) evaluate the specimen margin for positive or negative margins, (2) check for EPE, and (3) provide the Gleason grade. If during the review of a VRI, Pathologist C felt that the VRI did not contain sufficient information, on their request, we provided to Pathologist C the VRI that combined the highest resolution tiles used by both Pathologist A and B, called the ‘union’ VRI. If Pathologist C still felt the resolution was insufficient to render an accurate diagnosis, on their request, we provided them with the full-resolution WSI (Fig. 2).

For each slide, we recorded which of the 3 images (Pathologist A’s VRI, the union VRI, or the original WSI) that Pathologist C ultimately used for each task. We recorded each task’s results and produced resolution maps from Pathologist C’s sessions to analyze the visitation history of Pathologist C for every image viewed. To compare diagnostic results, we defined major diagnostic discordance as instances where Pathologist C’s diagnosis of margin status, EPE, or Gleason grade differed from both Pathologist A and B’s diagnosis.

## Results and discussion

### VRI generation and analysis

In this study, we digitized slides at 20x magnification and asked expert pathologists to review the WSIs to determine the extent to which pathologists utilize the native resolution. Resolution maps were created for each reviewer for each slide, which provide information on specific resolution utilization for each reviewer/slide combination. While certain regions were viewed at the native resolution, we found that large areas of the slides were never viewed at the highest available resolution by any of the three

reviewers. The resolution maps derived from digital slide review were then used to resample the WSIs to create VRIs based on each pathologist’s viewing history. Fig. 3 shows an example of VRI resolution maps overlaid onto a thumbnail of the WSI, as well as a comparison of pixel zoom utilization across both primary pathologists for a single slide from the case (Slide 0). For this slide, the greatest proportion of the slide area was viewed at 2.5x or lower resolution by both pathologists, and the size of the VRI compared to the WSI was 4% for Pathologist A and 8% for Pathologist B. The effective compression ratio (defined as the number of utilized pixels over the total number of pixels in the original WSI) for this example was therefore 23:1 and 12:1 for Pathologists A and B, respectively. Fig. 4 and the supplemental video shows an example of a VRI at a low zoom level and increased zoom levels, where transitions in the image can be observed at boundaries between utilized resolution levels. This specimen contained a positive surgical margin, the location of which is indicated by the black box in Fig. 3a. The resolution maps clearly show areas where both pathologists required higher resolution. For instance, higher resolution was used at the boundary of the specimen by both reviewers, reflective of the task to assess the surgical margin for tumor cells touching the inked boundary.

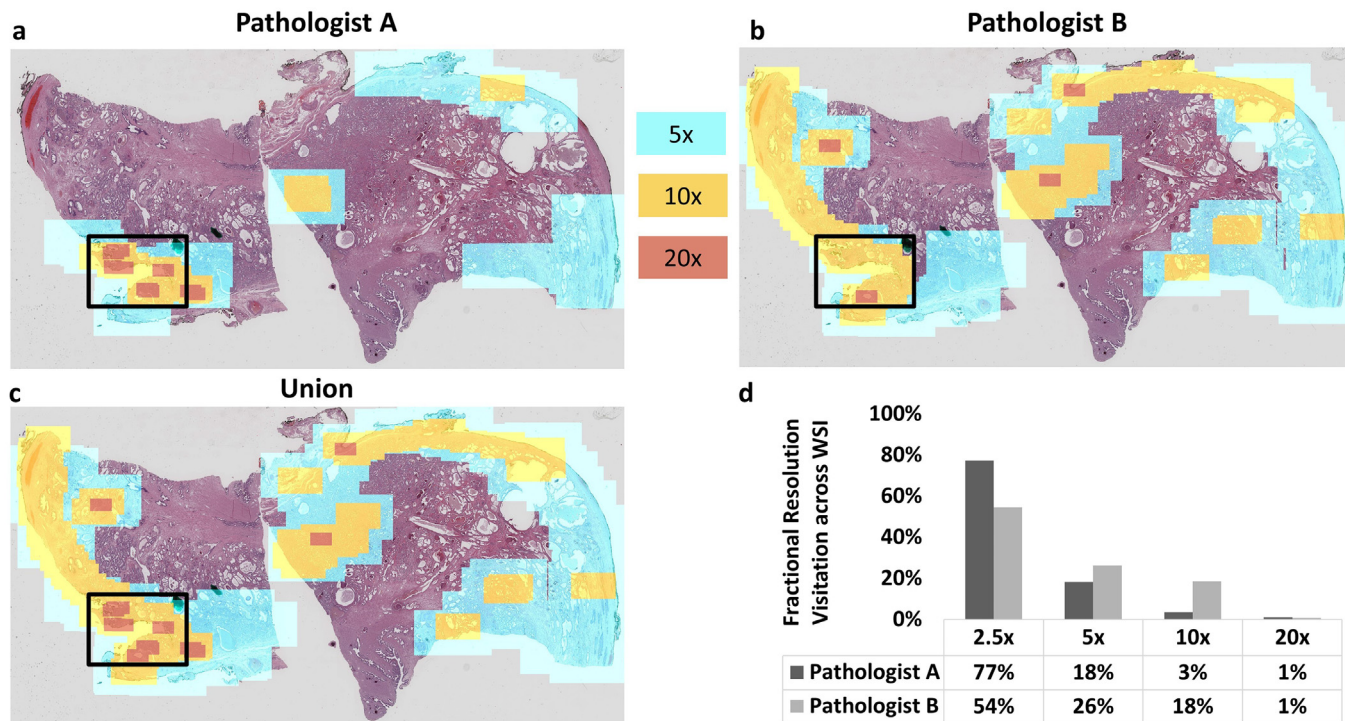
In aggregate, for all 25 slides in this radical prostatectomy case, the number of pixels contained in the VRIs was on average 4% of the original WSI pixel sizes across both primary reviewing pathologists, with standard deviations of 1% and 2% for Pathologists A and B, respectively. The average effective compression ratio was therefore 31:1. The VRIs produced from Pathologist A’s session ranged from 3% to 5% of the original WSI pixel sizes, whereas the VRIs produced from Pathologist B’s session ranged from 2% to 8% of the original WSI pixel sizes (Fig. 5b). A paired Student’s *t*-test confirmed the sizes of Pathologist A and B’s VRIs were not significantly different with *P*-value > 0.05. The average size of the union VRIs was also only 5% of the original WSI size.

In terms of each pathologist’s range of zoom or resolution levels, Pathologist A viewed at least part of the slide image at 20x resolution for 80% of the slides. Pathologist B utilized 20x resolution in at least part of 68% of the slides. Whether the slide image contained cancer did not greatly impact data size reduction. Specifically, the average VRI to WSI size ratio was 3% for specimens with cancer and 4% for specimens without cancer. Although

### VRI Evaluation Process



Fig. 2. VRI evaluation workflow for secondary pathologist review. For each slide, Pathologist C first reviewed the VRIs produced from Pathologist’s A session, and provided a diagnosis. If Pathologist C felt the resolution was too low, the pathologist moved on to the union VRI, and updated their diagnosis as necessary. Finally, the pathologist had the option to review the full resolution WSI and to update their diagnosis accordingly.



**Fig. 3.** Resolution utilization example for a radical prostatectomy slide for two primary reviewers. This figure shows the highest utilized resolution per pixel for: (a) Pathologist A and (b) Pathologist B as they reviewed Slide 0, a specimen with a positive surgical margin. Areas in red correspond to 20x resolution, areas in yellow correspond to 10x resolution, areas in blue correspond to 5x resolution, and all other areas correspond to 2.5x or lower resolution. The area inside the black box contains the positive surgical margin. (c) This image corresponds to the union of both pathologist reviewer’s resolution maps. (d) The bar graph displays the fractional resolution visitation history compared to the full resolution WSI across the specimen area for each pathologist.

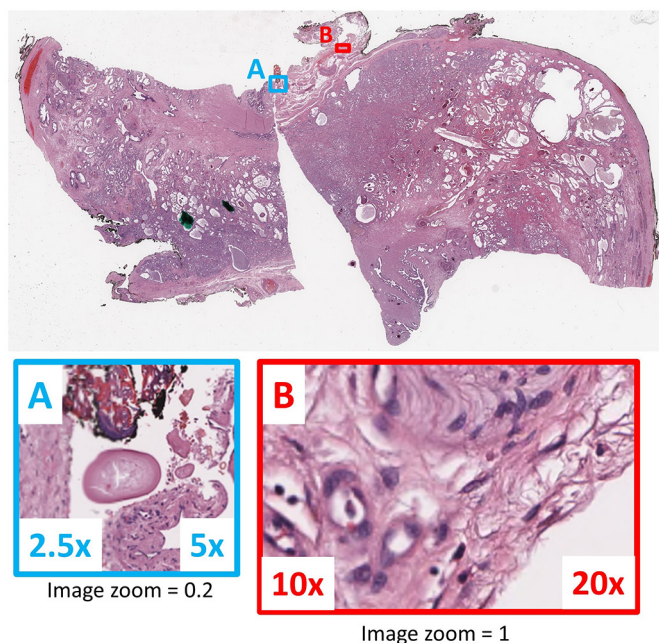
the sample size is too small to make a definitive conclusion, this result suggests that image file size can be dramatically reduced by pathologist-directed data reduction, whether or not cancer is present. Whereas overall data reduction was similar, we also analyzed the total area viewed at 20x in the union VRIs across specimens with and without tumor. We found a

significant difference (two-tailed Student’s *t*-test *P*-value = 0.01) in total area covered with 20x resolution between specimens with and without tumor. As expected, the pathologists covered more area with 20x resolution for specimens with tumor.

Across all slides, the necessary resolution for the primary pathologists to render their final clinical decisions was heavily weighted towards lower zoom levels (Fig. 5a). In fact, on average, only 0.6 (0.5%) and 0.3 (0.3%) of the slides’ areas were observed at the raw native resolution (i.e., 20x) for Pathologists A and B, respectively. Given the pathologists’ limited need to view the digital slides at the full native acquisition resolution, the VRIs produced from the viewing sessions for both pathologists were significantly smaller than the original WSIs. While it is understood that radical prostatectomy slide review does not typically require a large degree of high magnification usage, this does represent an example use case where the size of the data of a full-resolution WSI is large compared to the necessary image information needed to render the diagnosis; such a situation would likely be common across surgical pathology applications in other organ sites, or in cytopathology or hematopathology. From additional preliminary work with kidney biopsy specimens (data not shown), we have reason to believe that our hypothesis on resolution variation during slide viewing can be applied in other settings. Fig. 5 thus supports our hypothesis that only a fraction of the slide is viewed at the native resolution of the whole slide imaging system, and that pathologist-directed adaptive resampling using VRIs could be an efficient means for data size reduction for archival purposes.

*VRI evaluation results*

After transforming the WSIs into VRIs based on Pathologist A and B’s review sessions, Pathologist C then reviewed Pathologist A’s VRIs, with the opportunity to review the union VRIs and full-resolution WSIs if deemed necessary by the pathologist. We chose to use Pathologist A’s VRIs as the primary data source for secondary review, since this pathologist had the



**Fig. 4.** Example of resolution transitions in a variable resolution image. A and B show examples of border transitions between variable resolution image tiles at image zoom 0.2 and 1, respectively.

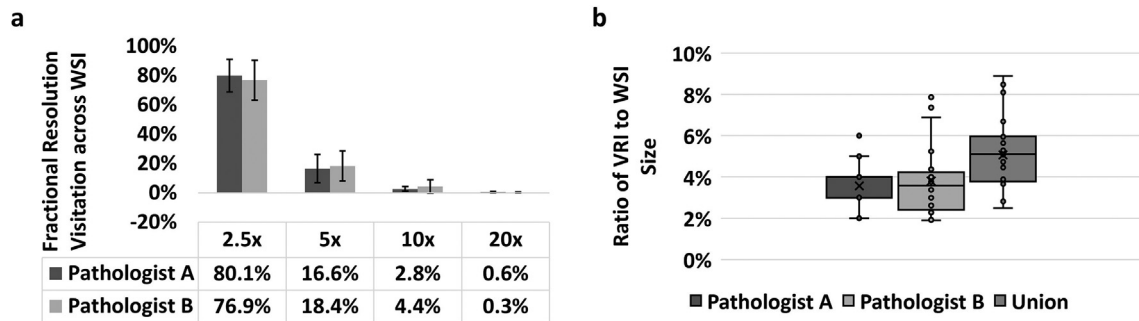


Fig. 5. Comparison of pixel utilization and file size between pathologist-directed VRIs and full-resolution WSIs. (a) The highest fraction of resolution utilization for both pathologists was at 2.5x, and dropped off precipitously at higher zoom/resolution levels. (b) The ratio of VRI to WSI pixel sizes for Pathologists A and B and the union of their VRIs.

smallest VRIs on average across all slides (i.e., lowest resolution utilization over the whole slide area). Pathologist C reached a definitive diagnosis on 84% of the VRIs derived from Pathologist A’s diagnostic session only. Additionally, Pathologist C asked to review the union VRIs for only 16% of the slides (Table 1) and did not ask to review any of the full resolution WSIs. For 2 out of the 4 higher-resolution union VRIs reviewed, Slides 1 and 6, Pathologist C’s diagnosis did not change from their initial diagnosis rendered on Pathologist A’s VRIs, after reviewing the higher resolution union VRIs. For Slide 20, where Pathologist C referred to the union VRI, Pathologist C mentioned the specimen would require recutting because the specimen was partially missing. Pathologist C decided not to refer to the full resolution WSI (Fig. 6), likely because resolution was not the issue limiting diagnostic decision in this case.

In the case of Slide 7, the glass slide contained 2 serial sections from the block. In our instructions to all 3 reviewers, if the slide contained more than 1 section, we did not dictate which section should be reviewed. In this

particular case, Pathologists A and Pathologist B reviewed opposite sections. Subsequently, Pathologist C chose to review a different section than Pathologist A (Fig. 7) which was rendered at only 2.5x in the Pathologist A VRI. As shown in Fig. 7, by observing where Pathologist C was looking on the section, if Pathologist C had reviewed the left tissue section instead of the right, there may have been adequate resolution to come to a diagnosis from Pathologist A’s result, due to the fact that the areas of interest for Pathologist A and C were in similar areas of the specimen. However, because we chose to present Pathologist A’s VRI first, Pathologist C subsequently asked for the union VRI, which contained additional resolution on the particular section in question due to the fact that this was the one that Pathologist B reviewed. A limitation of this study’s method is that pathologists may choose to review different sections if multiple are presented on 1 slide. However, in the future, we may present the resolution map as a thumbnail (see supplemental video) to indicate areas where the resolution is higher to aid the pathologists’ secondary reviews.

Table 1

Diagnostic results from Pathologists A, B, and C. We defined diagnostic concordance as instances where Pathologist C agreed with either Pathologists A and B on EPE and margin status and Gleason grade, for specimens with more than 5% tumor content.

Slide ID	Margins assessment				EPE assessment				Tumor volume		Gleason grade				Highest resolution image reviewed	Diagnostic concordance
	Path. A	Path. B	Path. C (Path. A VRI's)	Path. C (Union VRI's)	Path. A	Path. B	Path. C (Path. A VRI's)	Path. C (Union VRI's)	Path. A	Path. B	Path. A	Path. B	Path. C (Path. A VRI's)	Path. C (Union VRI's)	Path. C	Path. C
0	Positive	Positive	Negative	Negative	Negative	Negative	Negative	Negative	50%	40%	4+3	4+3	3+4		Pathologist A VRI	No
1	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	40%	30%	3+4	3+4	3+4	3+4	Union VRI	Yes
2	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	25%	30%	3+4	3+4	3+4		Pathologist A VRI	Yes
3	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	5%	5%	3+3	3+3	3+3		Pathologist A VRI	Yes
4	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	1%	<5%	3+3	3+3	3+3		Pathologist A VRI	Yes
5	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
6	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Union VRI	Yes
7	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	1%	<5%	3+4	3+3		3+4	Union VRI	Yes
8	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	<5%		3+3			Pathologist A VRI	Yes
9	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	<5%		3+3			Pathologist A VRI	Yes
10	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
11	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	5%		3+3			Pathologist A VRI	Yes
12	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	8%	10%	4+3	4+3	4+3		Pathologist A VRI	Yes
13	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	1%	0%	3+4		3+3		Pathologist A VRI	Yes
14	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
15	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
16	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
17	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes
18	Negative	Negative	Negative	Negative	Negative	Positive	Negative	Negative	60%	60%	3+4	3+4	3+4		Pathologist A VRI	Yes
19	Negative	Negative	Negative	Negative	Negative	Positive	Negative	Negative	50%	50%	3+4	3+4	3+4		Pathologist A VRI	Yes
20	Negative	Negative	Negative	Negative	Negative	Positive		Negative	10%	25%	3+3	3+3		3+3	Union VRI	Yes
21	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	1%	5%	3+3	3+4	3+3		Pathologist A VRI	Yes
22	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	8%	10%	4+3	4+3	3+4		Pathologist A VRI	Yes
23	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	25%	25%	4+3	3+4	4+3		Pathologist A VRI	Yes
24	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	0%	0%					Pathologist A VRI	Yes

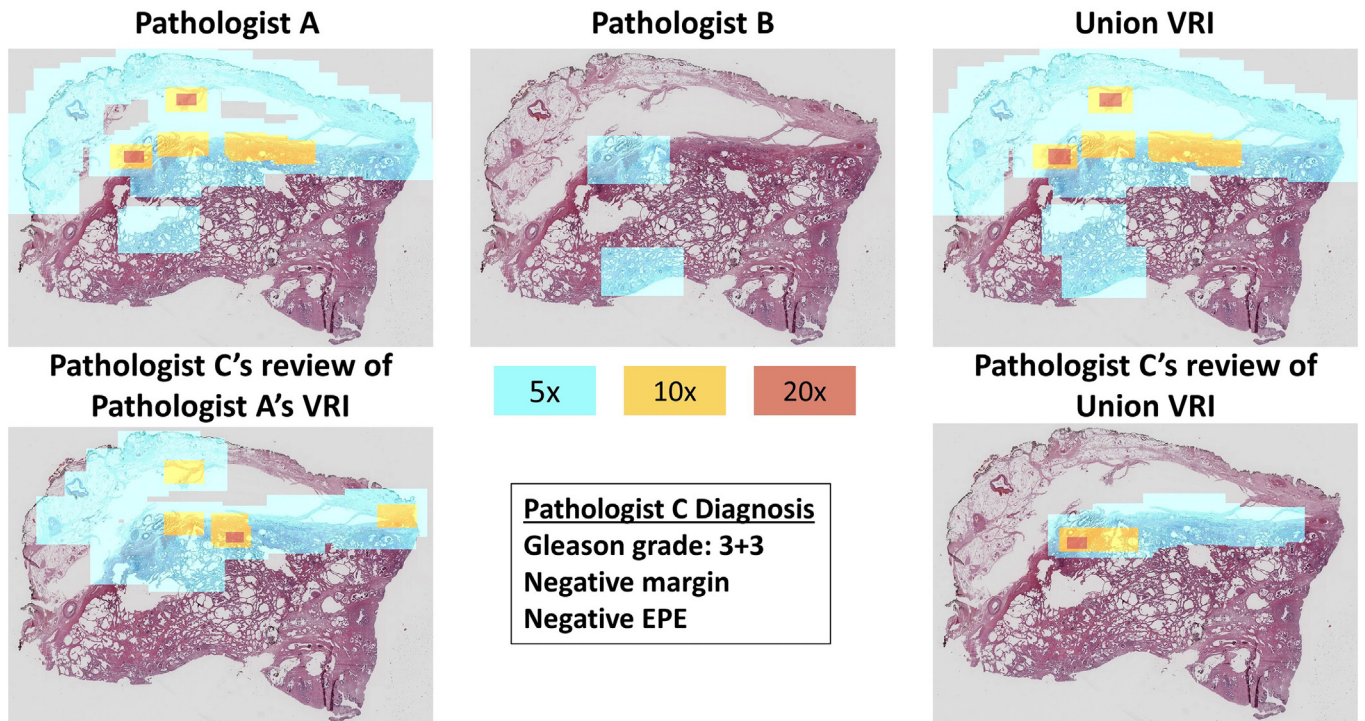


Fig. 6. This figure shows the resolution maps of Pathologist A and B and the union VRI of Slide 20 (top row). The bottom row features Pathologist C's resolution maps, produced from reviewing Pathologist A's VRI and the union VRI. Pathologist C agreed with Pathologist A and B on the Gleason grade and margin status. Pathologist A and B disagreed on EPE. Pathologist C noted that the slide preparation quality was low, which made EPE difficult to identify.

The most significant differences in diagnostic results were that Pathologist C reported a negative margin on Slide 0 after reviewing Pathologist A's VRI but did not review the union or full-resolution WSI. In contrast, Pathologists A and B both reported a positive margin. Pathologist B reported EPE on Slides 18, 19, and 20, but Pathologist A and C did not report EPE. Pathologist C felt EPE was difficult to determine on Slide 18, but did not refer to the full resolution WSI, indicating the difficulty in determining this was not due to image resolution. As previously mentioned, Pathologist C also

would have asked to recut Slide 20, which indicates a slide preparation issue rather than an image resolution issue. Other differences in diagnostic reporting were slight, such as not grading tumor when only a few glands were present and slight variations in Gleason grade (which is known to be variable across reviewers), none of which were diagnostically significant. Overall, Pathologist C mentioned that the difficulty in providing diagnoses for some slides was due to poor slide preparation or complexity of the diagnostic features present rather than the available resolution.

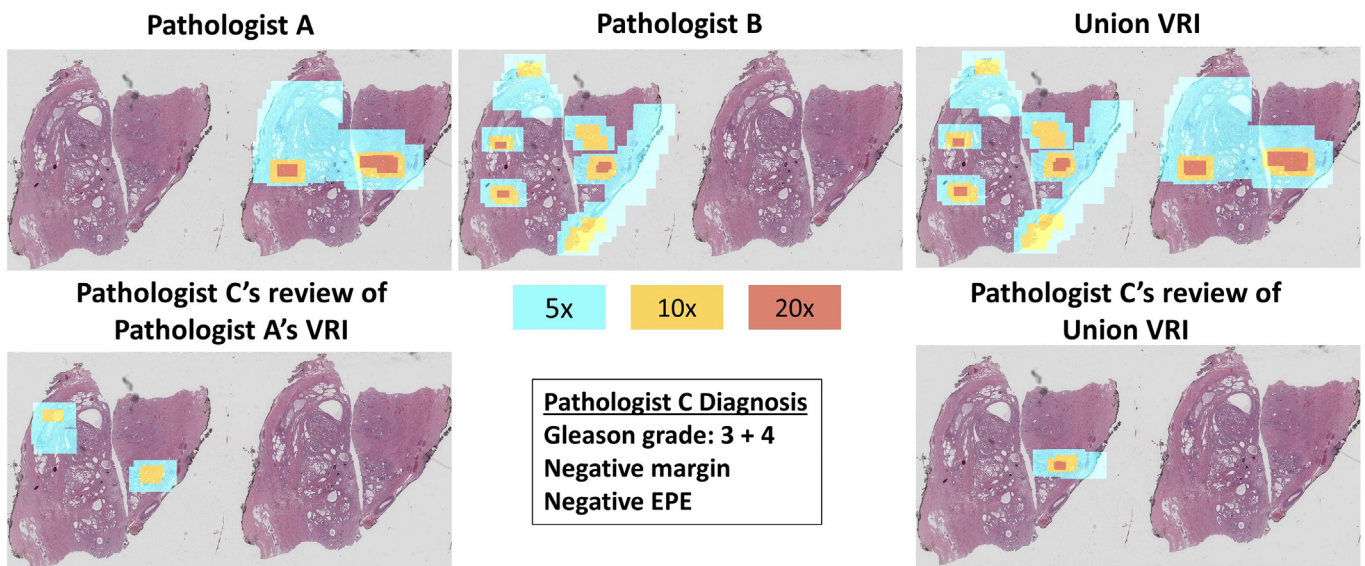


Fig. 7. This figure shows the resolution maps of Pathologist A and B and the union VRI of Slide 7 (top row). The bottom row features Pathologist C's resolution maps produced from reviewing Pathologist A's VRI and the union VRI. Pathologist C agreed with Pathologist A on the Gleason grade and with both Pathologists A and B on margin status and EPE.

### Comparisons to related work

The concept of reducing whole slide images data size based on regions of interest has been appreciated in the past. Specifically, Varga et al. altered image compression based on annotated histopathological images.<sup>21,22</sup> In contrast, our method does not rely on annotations but is based on the actual viewing sessions of the pathologists, directly informing which areas require certain resolutions. There have also been studies on user interaction in digital pathology with multi-resolution viewers<sup>19,23–25</sup> and eyetracking,<sup>26</sup> however, to our knowledge, this is the first application of using pathologists' interactions with digital slides to adaptively alter whole slide image resolution for storage and subsequent secondary review. There has also been previous work on compositing several images of differing zoom factors into a single VRI,<sup>27</sup> but this is the first application of such variable resolution imaging in pathology. With our approach, pathologist diagnostic behavior determines the image resolution across the specimen area and allows us to reduce whole slide image data sizes while maintaining the relevant information to the diagnosis. In the future, this method may reduce data storage costs or enable more efficient sharing of WSIs.

### Future work

While these pilot results are promising, before the VRIs can be implemented in clinical practice for archival or secondary review, we will need to conduct additional appropriately powered studies to understand inter- and intra-pathologist variability in resolution requirements. Additionally, our results from radical prostatectomy specimens will need to be confirmed across other specimen types and with more complex presentations of prostate cancer, such as specimens that include benign mimickers of prostate cancers. Interestingly, even if pathologists did not view the same areas at the same resolution levels, there was diagnostic concordance amongst reviewers for some specimens. For Gleason grading in particular, pathologists can review different areas of the tumor and come to the same conclusion on grade. In this study, we combined resolution levels viewed by different pathologists to create 'union' VRIs to account for variation in zoom behavior. In the future, we could potentially train a machine-learning model to predict which portions of the digital slides require certain resolution levels by one or more pathologists based on particular image features of interest, and require that these areas also be stored at the highest resolution.

Although we did not focus on it here, this method can also be combined with standard web file format or compression schemes, such as JPEG or JPEG2000, to reduce data sizes further. Additionally, images and redundant backups may be stored in tiers with cloud storage, where cost varies based on ease of access and performance requirements. With our approach, tiered cloud storage could be informed by the actual viewing patterns of the pathologist, which would help determine which images to store in higher-cost storage. Additionally, our approach could seamlessly reduce data storage requirements for clearly unusable slide images, such as out of focus slides or those with other artifacts, that slip through the institution's quality control process.

### Conclusions

We investigated the resolution requirements of 2 pathologists reviewing 25 digital radical prostatectomy slides. The pathologists reviewed each slide in a web-viewer that logged image zoom and viewport size and coordinates over time. We created resolution maps from the logs and used this information to reduce WSI pyramids on a per-tile basis. We found that pathologists performing diagnoses only view a small percentage of the WSI pixels, and most of the specimens were viewed with low resolution primarily. By altering the WSIs to reflect the pathologist-utilized resolution levels, the VRIs are optimally digitized, meaning the image file maintains at least the minimum number of pixels required to review their diagnoses. We were able to reduce the size of the WSIs by altering resolution without the use of compression by an average of 96%. Despite significantly reducing file sizes,

a third pathologist reviewer was able to use the VRIs to evaluate diagnostic features and provide diagnoses for each slide.

Although we would need a larger data set with additional specimen types to make any conclusive arguments on the ability to retain full diagnostic quality, this work has provided the framework and identified methods for demonstrating user-directed viewer interaction to reduce data size. While some pathologist raters more routinely used higher zoom levels for prostate cancer diagnosis and pathology report workup than others, there were few major differences in diagnostic results between the 3 raters. In the future, by collecting the results of many pathologists, we may be able to train an algorithm to detect all possible areas where pathologists may want to zoom, such as the full tumor area, to create VRIs, and combine our method with standard image compression to further reduce data size. With additional development, our method could have enormous implications for data and cost reduction in digital pathology.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpi.2022.100113>.

### Source(s) of support

NIH R01 CA222831, DMS 1664848, NSF IIS 2136744, NSF DGE-1144646

### Conflicts of interest

Co-contributor Brown is a founder and officer of Instapath, Inc. Instapath, Inc. did not contribute any financial or material support of this work. Co-contributors Ashman, Brown, Huimin, and Summa are listed as co-inventors on a patent application submitted by Tulane University related to this work.

### References

- Pantanowitz L, Sharma A, Carter A, Kurc T, Sussman A, Saltz J. Twenty years of digital pathology: an overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. *J Pathol Inform* 2018;9(1):40.
- Aeffner F, Zarella M, Buchbinder N, et al. Introduction to digital image analysis in whole-slide imaging: a white paper from the digital pathology association. *J Pathol Inform* 2019;10(1):9.
- US Food and Drug Administration. *TAU Corporation FDA approval letter for TigerScan/TigerView (K930527/A)*. Oct. 1995.
- US Food and Drug Administration. *Philips IntelliSite Pathology Solution FDA approval letter (DEN160056)*. Oct. 2017.
- Dandu RV. Storage media for computers in radiology. *Ind J Radiol Imag* Oct 2008;18:287–289.
- Lujan G, Quigley J, Hartman D, et al. Dissecting the business case for adoption and implementation of digital pathology: a white paper from the digital pathology association. *J Pathol Inform* 2021;12(1):17.
- Vainer B, Toft BG, Olsen KE, Jacobsen GK, Marcussen N. Handling of radical prostatectomy specimens: total or partial embedding?: handling of RPS. *Histopathology* Jan 2011;58:211–216.
- Hanna MG, Reuter VE, Samboy J, et al. Implementation of digital pathology offers clinical and operational increase in efficiency and cost savings. *Arch Pathol Lab Med* Dec 2019;143:1545–1555.
- Schüffler PJ, Geneslaw L, Yarlagadda DVK, et al. Integrated digital pathology at scale: a solution for clinical diagnostics and cancer research at a large academic medical center. *J Am Med Inform Assoc* Aug 2021;28:1874–1884.
- Clunie DA, Dennison DK, Cram D, Persons KR, Bronkalla MD, Rimo H. Technical challenges of enterprise imaging: HIMSS-SIIM collaborative white paper. *J Digit Imag* Oct 2016;29:583–614.
- Kalinski T, Zwönitzer R, Grabellus F, et al. Lossless compression of JPEG2000 whole slide images is not required for diagnostic virtual microscopy. *Am J Clin Pathol* Dec 2011;136:889–895.
- U.S. Food and Drug Administration. *Evaluation of Automatic Class III Designation for Philips IntelliSite Pathology Solution (PIPS)*. Oct 2017.
- Helin H, Tolonen T, Ylänen O, Tolonen P, Nääpänkangas J, Isola J. Optimized JPEG 2000 compression for efficient storage of histopathological whole-slide images. *J Pathol Inform* 2018;9(1):20.
- Cross S, Furness P, Igali L, Snead D, Treanor D. Best practice recommendations for implementing digital pathology. *Tech. Rep. The Royal College of Pathologists*; Jan 2018.
- "Digital Imaging and Communications in Medicine (DICOM) – Supplement 145: Whole Slide Microscopic Image IOD and SOP Classes," *Tech. Rep., DICOM Standards Committee, Working Group 26, Pathology*.
- Tuominen VJ, Isola J. The application of JPEG2000 in virtual microscopy. *J Digit Imag* June 2009;22:250–258.

17. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* Nov 2018;68:394–424.
18. Zarella MD, Bowman D, Aeffner F, et al. A practical guide to whole slide imaging: a white paper from the digital pathology association. *Arch Pathol Lab Med* Feb 2019;143:222–234.
19. Zhuge H, Summa B, Brown JQ. *A web-based multiresolution viewer to measure expert search patterns in large microscopy images: towards machine-based dimensionality reduction*. International Society for Optics and Photonics. Mar 2019.
20. Shimrat M. Algorithm 112: position of point relative to polygon. *Commun ACM* Aug 1962;5:434.
21. Varga MJ, Dicksbury PG, Callagy G. Application of content-based image compression to telepathology San Diego, CA. May 2002:587–595.
22. Varga MJ, Dicksbury PG, Callagy G. Performance visualization for image compression in telepathology(San Diego, CA). Apr 2002:145–156.
23. Krupinski EA, Graham AR, Weinstein RS. Characterizing the development of visual search expertise in pathology residents viewing whole slide images. *Human Pathol* Mar 2013;44:357–364.
24. Mercan E, Shapiro LG, Brunyé TT, Weaver DL, Elmore JG. Characterizing diagnostic search patterns in digital breast pathology: scanners and drillers. *J Digit Imag* Feb 2018;31:32–41.
25. Molin J, Fjeld M, Mello-Thoms C, Lundström C. Slide navigation patterns among pathologists with long experience of digital review. *Histopathology* Aug 2015;67:185–192.
26. Shin D, Kovalenko M, Ersoy I, et al. PathEdEx – Uncovering high-explanatory visual diagnostics heuristics using digital pathology and multiscale gaze data. *J Pathol Inform* 2017;8(1):29.
27. Licorish C, Faraj N, Summa B. Adaptive compositing and navigation of variable resolution images. *Comput Graphics Forum* Oct 2020;p. cgf.14178.