A new approach assessing the quality of whole slide images by discovering specific issues

Yuchen Song
South China University of Technology
Ying Gao (gaoying@scut.edu.cn)
South China University of Technology
Changhong Liang
Guangdong General Hospital
Zaiyi Liu
Guangdong General Hospital

Research

Keywords: Whole slide images (WSI), Digital pathology, Quality assessment

Posted Date: March 31st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-20180/v1

A new approach assessing the quality of whole slide images(WSIs) by discovering specific issues

Yuchen Song[1], Ying Gao[1,*], Changhong Liang[2], Zaiyi Liu[2]

[1]School of Computer Science and Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China

[2]Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong 510080, China

***Corresponding Author**

Y. Gao, School of Computer Science and Engineering, South China University of Technology, Guangzhou 510632, China (E-mail: gaoying@scut.edu.cn).

Abstract

Background: Since whole slide images (WSIs) are widely used in clinical diagnosis, assessing the quality of these images has generated considerable interest. We divided the evaluation process into different sections focusing on specific issues and developed an approach based on computer technology to assess the quality of WSIs and provide details corresponding to specific issues.

Methods: We focused on three issues in this work: position deviation, cover slip misplacement and wrong focus. Our approach revolves around these issues. We collected two

datasets from Guangdong Provincial People's Hospital to develop the approach, discovered three issues and provided details corresponding to these issues.

Results: We designed experiments based on two datasets to validate the effectiveness of the approach. The experiments show that using our approach, existing issues can be discovered and corresponding details are provided at a high accuracy. In the training set, the accuracy of the discovered three issues reaches 0.998, 0.987 and 0.988. In the validation set, the accuracy of discovering all issues reaches 0.914.

Conclusions: We proposed an approach focusing on the following specific image issues: position deviation, cover slip misplacement and wrong focus. Using this approach, issues were discovered with high accuracy, and reports were produced that provide details regarding existing issues discovered by our approach. However, a dataset of WSIs with more issues, including stain concentration and poor microtomy, should be established, and methods discovering these issues must be developed in the future.

Keywords: Whole slide images (WSI), Digital pathology, Quality assessment

Background

Although pathology is considered the gold standard of clinical diagnosis, the use of microscopy has limitations in that microslides are difficult to store, archive and share. The newly developed whole slide image (WSI) is proposed to solve these limitations. WSIs can digitalize microslide information. In other words, we could use computer technology for clinical applications. Therefore, WSIs are widely used in clinical diagnosis [1-4]. To

45 guarantee the accuracy and efficacy of WSIs in place of the direct observation with a
46 microscope, assessing the quality of WSIs has generated considerable interest [5-8].
47 Initial attempts have focused on assessing the clarity of WSIs. The pathologist's
48 opinion on image quality is the most straightforward choice. Nevertheless, the difference
49 between pathologists in this work results in stability issues due to subjectivity. Furthermore,
50 owing to the massive number required to be evaluated, this method is time-consuming for
51 pathologists [7]. Consequently, automated methods for assessing image quality should be
52 proposed. Walkowski and Szymas [9] showed that the gray level co-occurrence matrix
53 (GLCM) could well represent the sharpness of WSIs. Hashimoto et al. [10] added noise into
54 the analysis model to assess image quality. Other studies have made great effort using
55 different features based on computer technology to represent the quality of WSIs [11-14]. In
56 addition, several tools have been developed to assess image quality automatically. Ameisen et
57 al. [15] developed a method that integrates features including saturation, contrast, and
58 brightness and then used a picture to display the distribution of blurred areas. Janowczyk et
59 al. [16] developed an open source tool to calculate image metrics and identify artifact-free
60 regions with supervised classifiers.
61 Previous studies demonstrated the effectiveness of methods based on computer
62 technology. However, one common issue of the methods mentioned above is that they fail to
63 address the fact that the clarity is not the unique standard to measure image quality. Even
64 though it is simple for a person to evaluate a single WSI by providing a numeric score or a
65 simple description of the quality, such as good or poor, it is a complex task for computer-
66 based methods. In addition, a numeric score or a simple description of the quality is not

meaningful enough compared to providing information regarding specific issues in WSIs. Hence, we divide the evaluation process into different sections focusing on different specific issues and develop an approach based on computer technology to not only judge the varieties of issues that exist in WSIs but also provide the details corresponding to specific issues.

Methods

The most essential idea in this work is to divide the evaluation process into sections focusing on specific issues. Each section provides a result that contains whether a corresponding issue exists and details related to the issue. These results are integrated into a report that provides a reference to evaluate and enhance the quality of a single WSI. We claim three issues in this work: position deviation, cover slip misplacement and wrong focus [17]. Position deviation means that the scanner failed to obtain complete information due to the deviation between the virtual position and real position on a slide. Cover slip misplacement means that a black line exists and leads to a serious blur problem around the black line. Wrong focus means the failure of obtaining glass slides' texture information due to the wrong focus setting. Figure 1 shows an example of the 3 varieties of issues. Our approach addresses these three issues.

Datasets

We collected two datasets in this study. Both datasets are from Guangdong Provincial People's Hospital. The first dataset is composed of 276 gastric cancer WSIs, including 90 normal WSIs and 186 WSIs with abnormal quality issues. This dataset is used for training. Actually, multiple types of issues may exist in a WSI. In the training process, each WSI is

labeled by its most apparent issue. A total of 186 WSIs with abnormal quality issues consisted of 74 with position deviation, 38 with cover slip misplacement and 74 with wrong focus. The second dataset is composed of 58 colorectal cancer samples. This dataset is used for validation. As described above, the aim of this work is to generate result reports of issue details. Therefore, WSIs in the validation set are not divided into specific issue types. Datasets are defined as follows:

$I_{all} = \{I_{train}, I_{valid}\}$

$I_{train} = \{I_{PD}, I_{CSM}, I_{WF}, I_N\}$

where I_{train} represents the first dataset and I_{valid} represents the second dataset. In I_{train}, we use the first letter of each word to represent the problem. For instance, PD represents a position deviation issue.

Position Deviation

It could be easily observed that if a WSI has position deviation, at least one edge has tiles containing useful information for clinical diagnosis. Thus, the number of tiles with clinical information on one edge can represent whether position deviation exists on such edges. With the purpose of calculation, the following steps are required:

1. We trained a deep convolutional network to judge whether tiles contain useful information for clinical diagnosis.
2. We collected tiles on the four edges of WSIs from I_N and I_{PD}.
3. We calculated the count vector of each WSI and selected a suitable threshold to judge whether the WSI has a position deviation issue.

In the first step, our aim is to gain a model that can identify whether tiles split from the WSIs contain useful information for clinical diagnosis. We selected fifteen WSIs and split them into tiles using Openslide [18]. The tiles we collected for training this classifier model are 256×256 in size at 20× magnification. We chose samples from the split tiles and trained a deep convolutional network [19] to judge whether tiles contain useful information for clinical diagnosis.

Position deviation may exist in more than 1 orientation. Hence, we need to calculate a vector consisting of four values to represent whether position deviation exists and to determine its orientation. In our experiments, we used this method on all WSIs from I_N and I_{PD} to acquire the vectors mentioned above. Tiles we collected on edges are 64×64 in size at 10× magnification. Finally, we selected a suitable threshold that could distinguish whether position deviation occurs most accurately.

Cover Slip Misplacement

For a WSI with cover slip misplacement, an obvious black line appears. We use a method based on template matching to generate a mask displaying the position of the cover slip. We used the area threshold, the minimum area that we consider as the black region to adjust the masks segmented from the thumbnails of the WSIs to judge whether cover slip misplacement exists. In this study, we extended this method to obtain results that consist of classification outcomes and position information. To obtain the results of cover slip misplacement, the following steps are required:

1. We collected thumbnails of the WSIs from I_N and I_{CSM}.

2. We selected a thumbnail from I_{CSM} obtained in step 1 and artificially drew the cover slip area as the template mask.
3. We generated masks through the template mask and corresponding features.
4. We removed the edge pieces of masks generated in step 3 and calculated the mean of each mask represented by a matrix obtained by computer technology. Then, we selected a suitable threshold to judge whether a WSI indicates cover slip misplacement.

The existence of black areas on the border of WSIs may influence the result. Figure 2 is an example in which black areas on the border are detected on the border. Therefore, it is necessary to remove the border pieces to exclude the effect of the existence of black areas on the border. The masks are all binary images. The masks of WSIs with cover slip misplacement contain more white pixels than masks of normal WSIs. Therefore, we chose a threshold to judge whether WSIs have cover slip misplacement issues. In addition, the result returned information on the position of the cover slip location to provide a reference for correcting cover slip misplacement.

Wrong Focus

Figure 3 shows a comparison of a tile from a normal WSI and a WSI with wrong focus. We could speculate that if a WSI is composed of tiles with lower information richness, it is more likely to be a WSI with wrong focus. With the purpose of judging whether a WSI has wrong focus, the following steps are required:

1. We collected thumbnails of WSIs from I_N and I_{WF}.

2. We split thumbnails collected in step 1 into pieces.
3. We calculated each piece's Laplacian gradient function to represent the information richness. For each WSI, we selected the ten largest values of the Laplacian gradient function and appended them to a vector.
4. We calculated the mean of each vector obtained in step 3 and chose a suitable threshold to judge whether the WSI had the wrong focus.

In this part, it is very important to use an appropriate feature to represent information richness. Some methods of image quality assessment in natural images may be helpful in this section [20, 21]. In this work, we use the Laplacian gradient function to represent information richness of pieces from the thumbnails of the WSIs. The formula of the Laplacian gradient function is as follows:

$$D(f) = \sum_y \sum_x |G(x,y)| \ \ s.t. G(x,y) > T$$

$$G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)}$$

where $G_x^2(x,y)$ is the convolution of the Laplacian horizontal edge detection operator at pixel (x,y) and $G_y^2(x,y)$ is the convolution of the Laplacian vertical edge detection operator at pixel (x,y). T is a constant representing the edge detection threshold.

We calculated the mean of the ten largest values of the Laplacian gradient function of each WSI and then chose a suitable threshold to judge whether the WSIs have the wrong focus issue.

Results

We designed experiments on two datasets to validate the effectiveness of the approach. We

implemented the methods of each section and integrated these sections into an application.

Training Set

Position Deviation

An essential premise in this section is the accuracy of the classifier. We collected 1200 tiles, including 900 for training and 300 for testing. The accuracy of the training set reaches 0.995, and the accuracy of the testing set reaches 0.989. Such accuracy is sufficient to guarantee the proceeding of the left experiment in this section. An essential factor affecting the result is the size of the tiles that we collected on the edge. Table 1 illustrates the effect of the tile size on the results of this section.

Table 1. The relationship between the tile size and the performance of the method.

	Tile_Size=32		Tile_Size=64		Tile_Size=128		Tile_Size=256	
	N	PD	N	PD	N	PD	N	PD
Actual								
N	89	1	90	0	87	3	85	5
PD	2	72	2	72	9	65	13	61
Accuracy	0.982		0.988		0.927		0.890	
Time	2175		1726		1649		3013	

If the size of tiles is too large, then the accuracy will decrease because it may be more possible that tiles we collected contain tissue areas. In contrast, if the size of tiles is too small, then the accuracy may no longer ascend with a higher time consumption. The best value of the tile size is 64.

Cover Slip Misplacement

Due to the disturbance of the existence of other black areas, the selection of the area threshold in segmenting the black line area from the thumbnail of the WSIs mentioned in the Methods section is significant. Table 2 indicates the relationship between the area threshold and the accuracy of judging whether WSIs have the cover slip misplacement issue.

Table 2. The relationship between the area threshold and the accuracy.

Area Threshold	500	1000	2000	3000	5000
Accuracy	0.896	0.934	0.974	0.987	0.842

If the value of the area threshold is too small, then the prediction may be affected by other black areas of thumbnails. In contrast, if the value of the area threshold is too large, then the actual cover slip misplacement issue could be ignored. The best area threshold is 3000.

Wrong Focus

In this section, we experimented with three features, including the Brenner gradient function, variance function and entropy function [22], compared with the Laplacian gradient function. Table 3 illustrates the comparison of the accuracy using different features to judge whether wrong focus exists in the WSIs.

Table 3. The contrast of the influence of different features on the prediction of wrong focus.

Prediction of Different Features	Actual Normal	Actual Wrong Focus
Brenner		
Normal	87	1
Wrong Focus	3	72
Laplacian		
Normal	88	0
Wrong Focus	2	73
Variance		
Normal	79	15
Wrong Focus	11	58

Entropy		
Normal	84	14
Wrong Focus	6	59

Validation Set

This dataset includes 10 WSIs with position deviation, 17 WSIs with cover slip misplacement, 24 WSIs with wrong focus and 20 normal WSIs. A WSI may have multiple types of issues. Figure 4 illustrates the comparison between the results using the approach we proposed in this work and the facts. In addition, the approach could produce a report to provide results describing whether three issues exist and details related to existing issues.

Discussion

Previous studies have paid more attention to evaluating the quality of WSIs through clarity, but they ignored the fact that other factors could also influence the quality. We proposed an approach dealing with specific issues including the position deviation, cover slip misplacement and wrong focus to evaluate the quality of WSIs and provided the details of issues as references. We confirmed that issues could be discovered at a high accuracy. The most significant consideration of this work is that we focus more on specific quality issues rather than look for a single numeric value to represent the quality of WSIs. This work is the first research study that regards the evaluation of WSIs as a project of discovering several specific issues. This approach can be more meaningful than using only the clarity to represent the quality of WSIs.

However, more work should be done. Limited specialized datasets with issues exist

due to the lack of attention on the quality evaluation of WSIs. For instance, inappropriate stain concentrations could affect the efficiency of clinical diagnostics. Previous studies proposed some methods of evaluating the color reproducibility of WSIs [23-25]. However, it is difficult to discover areas of improper stain concentrations in the WSIs. In addition, methods of discovering areas with poor microtomy have not yet been proposed. Therefore, we will collect WSIs with inappropriate stain concentrations and poor microtomy. In addition, we will make an effort to develop methods to detect these issues.

Conclusions

In this paper, we proposed an approach to deal with several specific image issues including position deviation, cover slip misplacement and wrong focus. The approach could discover issues with high accuracy and produce reports providing details of the existing issues discovered. However, the limitation of specialized datasets with issues and methods of discovering such issues restrict the development of a quality evaluation. A dataset of WSIs with more issues including stain concentration and poor microtomy should be established, and methods discovering these issues should be developed in the future.

List of Abbreviations

WSI: whole-slide imaging; PD: position deviation; CSP: cover slip misplacement; WF: wrong focus

Declarations

253

254 ***Ethics Approval and Consent to Participate***

255 The present study was approved by the Ethics Committee of Guangdong Provincial People's

256 Hospital.

257

258 ***Consent for Publication***

259 Written informed consent for publication was obtained for each participant.

260

261 ***Availability of Data and Materials***

262 All data generated or analyzed during this study are included in this published article.

263

264 ***Competing Interests***

265 The authors declare that they have no competing interests.

266

267 ***Funding***

268 This work was supported in part by the Guangzhou People's Livelihood Science and

269 Technology Project under Grant 201803010097, in part by the Guangzhou City Industrial

270 Technology Major Research Project with Grant 201802010035, and in part by the Industrial

271 Internet Innovation and Development Project in 2018 Grant MIZ1824020-2.

272

273 ***Authors' Contributions***

274 YS and YG designed the experiments and developed methods by coding. YS wrote the

manuscript. CL and ZL collected the data required in the experiments and provided necessary background knowledge of the quality assessment of WSIs. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Authors' Information

[1]School of Computer Science and Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China

[2]Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong 510080, China

References

1. Al-Janabi S, Huisman A, Van Diest PJ. Digital pathology: current status and future perspectives. Histopathology. 2012;61:1-9.
2. Ghaznavi F, Evans A, Madabhushi A, Feldman M. Digital imaging in pathology: whole-slide imaging and beyond. Annu Rev Pathol. 2013;8:331-59.
3. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. Med Image Anal. 2016;33:170-5.
4. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol. 2019;20:e253-61.

297 5. Kayser K, Gortler J, Metze K, Goldmann T, Vollmer E, Mireskandari M, et al. How to
298 measure image quality in tissue-based diagnosis (diagnostic surgical pathology).
299 Diagn Pathol. 2008;3:S11.
300 6. Zerbe N, Hufnagl P, Schluns K. Distributed computing in image analysis using open
301 source frameworks and application to image sharpness assessment of histological
302 whole slide images. Diagn Pathol. 2011;6:S16.
303 7. Shrestha P, Kneepkens R, van Elswijk G, Vrijnsen J, Ion R, Verhagen D, et al.
304 Objective and subjective assessment of digital pathology image quality. AIMS Med
305 Sci. 2015;2:65-78.
306 8. Shrestha P, Kneepkens R, Vrijnsen J, Vossen D, Abels E, Hulsken B. A quantitative
307 approach to evaluate image quality of whole slide imaging scanners. J Pathol Inform.
308 2016;7:56.
309 9. Walkowski S, Szymas J. Quality evaluation of virtual slides using methods based on
310 comparing common image areas. Diagn Pathol. 2011;6:S14.
311 10. Hashimoto N, Bautista PA, Yamaguchi M, Ohyama N, Yagi Y. Referenceless image
312 quality evaluation for whole slide imaging. J Pathol Inform. 2012;3:9.
313 11. Lopez XM, D'Andrea E, Barbot P, Bridoux AS, Rorive S, Salmon I, et al. An
314 automated blur detection method for histological whole slide imaging. PLoS One.
315 2013;8:e82710.
316 12. Avanaki ARN, Espig KS, Xthona A, Lanciault C, Kimpe TRL. Automatic image
317 quality assessment for digital pathology. In: Lang K, Tingberg A, Timberg P, editors.
318 Breast imaging - 13th international workshop, IWDM 2016, proceedings. Malmo,

Sweden: Springer Verlag; 2016. p. 431-8.

13. Jiménez A, Bueno G, Cristóbal G, Déniz O, Toomey D, Conway C. Image quality metrics applied to digital pathology. In: Schelkens P, Ebrahimi T, Cristóbal G, Truchetet F, Saarikko P, editors. Proceedings of SPIE 9896. Optics, photonics and digital technologies for imaging applications IV. Brussels, Belgium: SPIE; 2016. p. 1-18.

14. Campanella G, Rajanna AR, Corsale L, Schuffler PJ, Yagi Y, Fuchs TJ. Towards machine learned quality control: a benchmark for sharpness quantification in digital pathology. Comput Med Imaging Graph. 2018;65:142-51.

15. Ameisen D, Deroulers C, Perrier V, Bouhidel F, Battistella M, Legres L, et al. Towards better digital pathology workflows: programming libraries for high-speed sharpness assessment of whole slide images. Diagn Pathol. 2014;9:S3.

16. Janowczyk A, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: an open-source quality control tool for digital pathology slides. JCO Clin Cancer Inform. 2019;3:1-7.

17. Pantanowitz L. Digital images and the future of digital pathology. J Pathol Inform. 2010;1:15.

18. Goode A, Gilbert B, Harkes J, Jukic D, Satyanarayanan M. OpenSlide: a vendor-neutral software foundation for digital pathology. J Pathol Inform. 2013;4:27.

19. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012. p. 1097-105.

20. Harlick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973;3:610-21.

21. Eskicioglu AM, Fisher PS. Image quality measures and their performance. IEEE Trans Commun. 1995;43:2959-65.

22. Sklansky J. Image segmentation and feature extraction. IEEE Trans Syst Man Cybern. 1978;8:237-47.

23. Yagi Y. Color standardization and optimization in whole slide imaging. Diagn Pathol. 2011;6:S15.

24. Cheng WC, Keay T, O'Flaherty N, Wang J, Ivansky A, Gavrielides MA, et al. Assessing color reproducibility of whole-slide imaging scanners. In: Gurcan MN, Madabhushi A, editors. Proceedings of SPIE 8676. Medical imaging 2013: digital pathology. Florida, US: SPIE Medical Imaging; 2013. p. 1-5.

25. Shrestha P, Hulsken B. Color accuracy and reproducibility in whole slide imaging scanners. J Med Imaging (Bellingham). 2014;1:027501.

Figure Legends

Figure 1. Examples of three issues claimed in this work. Figure 1(a) shows an example of position deviation. Figure 1(b) shows an example of cover slip misplacement. Figure 1(c) shows an example of wrong focus.

Figure 2. An example of a normal WSI that contains black areas on the border. Figure 2(a) is a normal thumbnail, and Figure 2(b) is a mask generated by the method. The white area on the border may influence the classification results. WSIs with white areas on the

363 border may be wrongly classified as WSIs with cover slip misplacement. Therefore, it is
364 necessary to remove the edge of the original mask.

365 **Figure 3. Comparison of pieces from the thumbnails of normal WSIs and WSIs with**
366 **Wrong Focus.** Figure 3(a) shows a piece from a normal WSI thumbnail, and Figure 3(b)
367 shows a piece from a thumbnail of the WSI with wrong focus. It is apparent that pieces from
368 normal WSIs contain more useful information for clinic than pieces from WSIs with wrong
369 focus.

370 **Figure 4. The contrast of the prediction provided by our approach and the facts.**

Figures

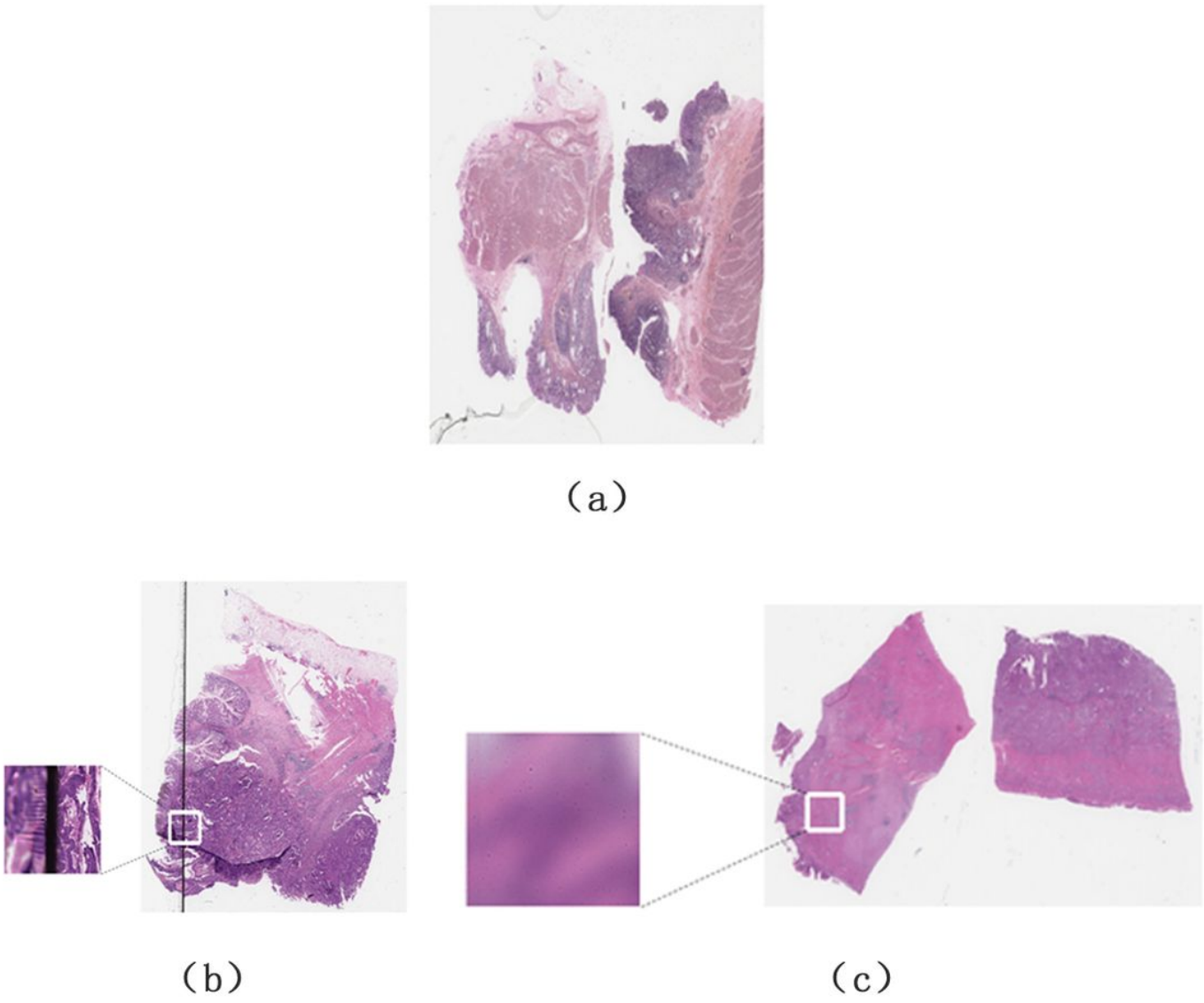

Figure 1

Examples of three issues claimed in this work. Figure 1(a) shows an example of position deviation. Figure 1(b) shows an example of cover slip misplacement. Figure 1(c) shows an example of wrong focus.

Figure 2

An example of a normal WSI that contains black areas on the border. Figure 2(a) is a normal thumbnail, and Figure 2(b) is a mask generated by the method. The white area on the border may influence the classification results. WSIs with white areas on the border may be wrongly classified as WSIs with cover slip misplacement. Therefore, it is necessary to remove the edge of the original mask.

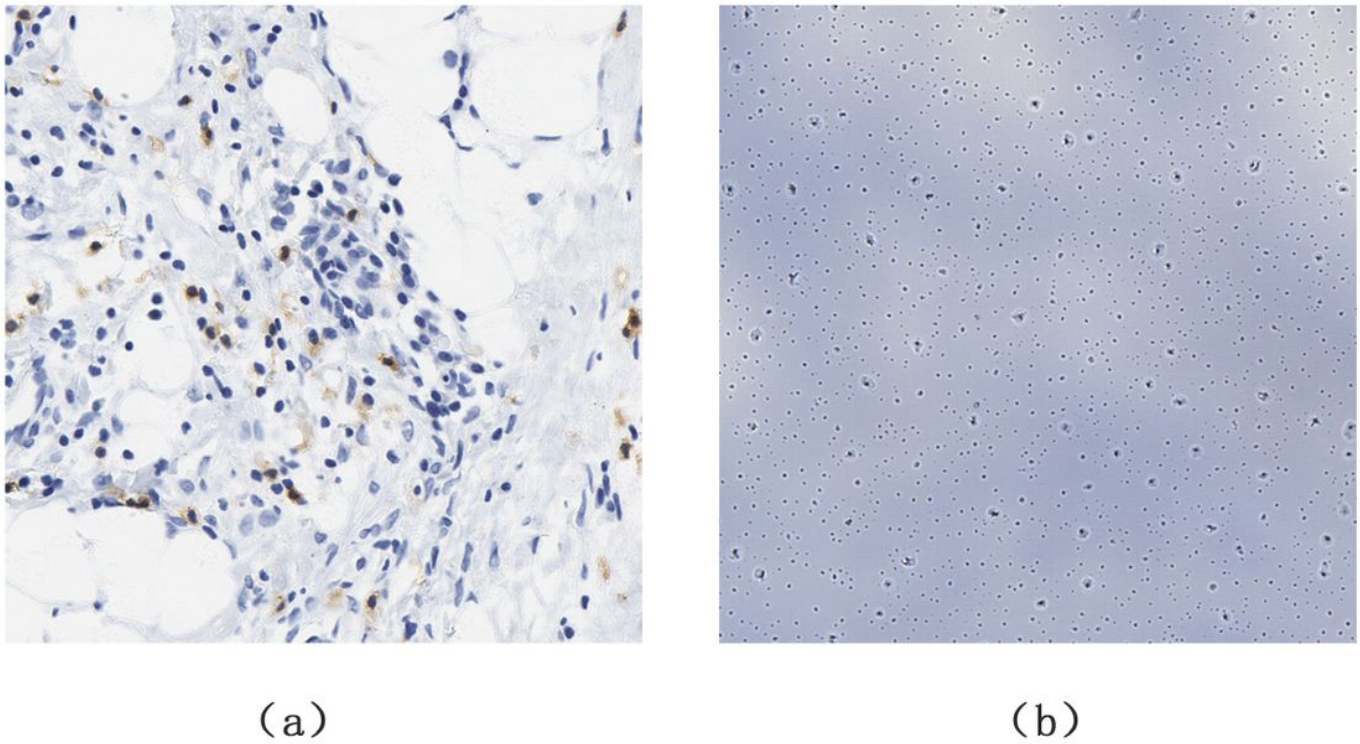

Figure 3

Comparison of pieces from the thumbnails of normal WSIs and WSIs with Wrong Focus. Figure 3(a) shows a piece from a normal WSI thumbnail, and Figure 3(b) shows a piece from a thumbnail of the WSI with wrong focus. It is apparent that pieces from normal WSIs contain more useful information for clinic than pieces from WSIs with wrong focus.

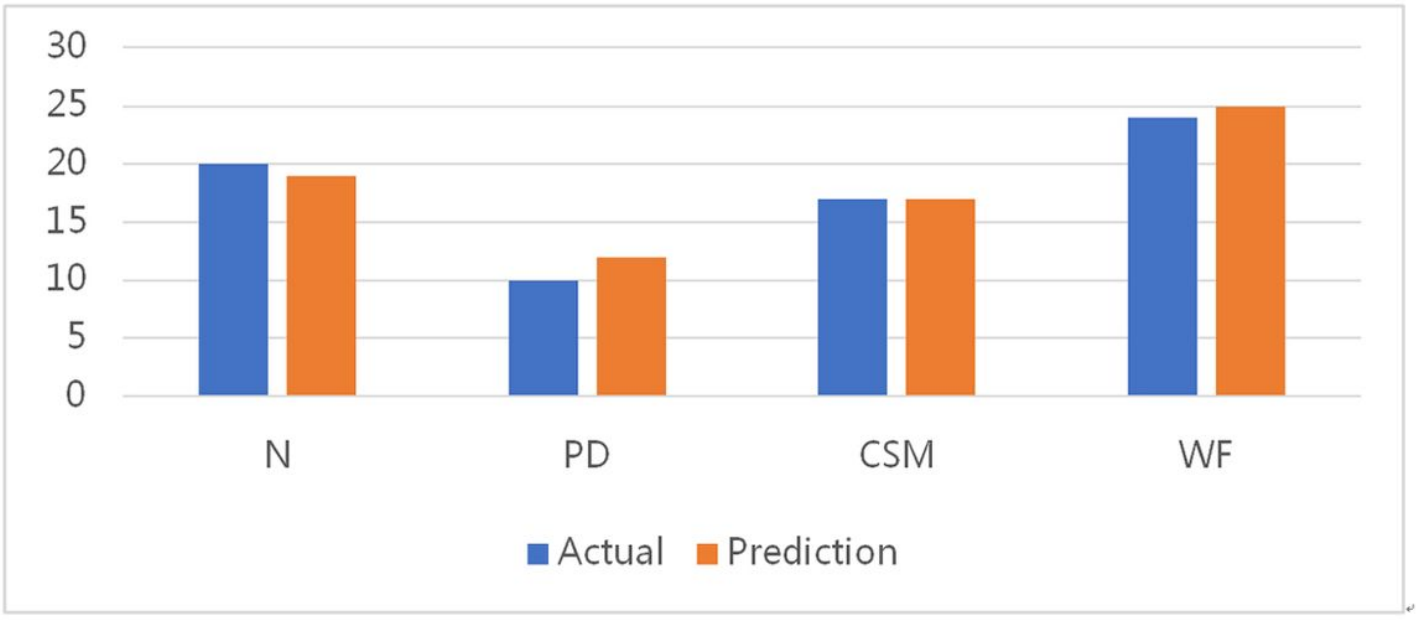

Figure 4

The contrast of the prediction provided by our approach and the facts.