Intelligent Oncology 1 (2025) 256–264

Contents lists available at ScienceDirect

Intelligent Oncology

journal homepage: www.keaipublishing.com/en/intonc/

Full Length Article

A fully automated quantitative analysis method based on deep learning algorithms for immunohistochemical staining expression intensities

Yongjian Deng[a,#], Bojun Cai[b,#], Xiaomei Wang[a,b,*]

[a] Hangzhou MediGenPro Biotechnology Co., Ltd., Hangzhou Zhejiang 311200, China
[b] Hangzhou PathoAI Technology Co., Ltd., Hangzhou Zhejiang 311100, China

ARTICLE INFO

Keywords:
Deep learning
Immunohistochemistry analysis
Image processing algorithm
Optical density separation
Quantification of whole-slide images

ABSTRACT

This paper focuses primarily on exploring the application of deep learning techniques and image processing algorithms in immunohistochemistry analysis, specifically targeting automated quantitative methods for nuclear, membrane, and cytoplasmic expressions of animal cells in whole-slide images. Cell nuclei, membranes, and cytoplasm were precisely identified and quantified by employing optical density separation techniques to differentiate between hematoxylin and 3,3'-diaminobenzidine staining components in combination with the CellViT nuclear segmentation algorithm and the region growing algorithm. Experimental validation demonstrates that the proposed algorithm performs excellently in terms of accuracy and recall. Compared to traditional manual interpretation, this algorithm achieve greater accuracy in specific quantitative metrics.

1. Introduction

The precise identification and quantification of tumor markers are critical tasks in modern medical research, with an immeasurable value in tumor diagnosis, the selection of precision treatment plans, and prognosis assessment.[1] Immunohistochemistry (IHC), a detection technique based on highly specific antigen-antibody reactions, has become an indispensable tool in the field of tumor marker identification due to its ability to accurately reveal the distribution and expression levels of proteins in cells and tissues.[2] IHC plays a crucial role in various important aspects, such as clinical pathological diagnosis, tumor classification, and making prognoses, providing valuable diagnostic information and treatment guidance for doctors.

However, traditional IHC analysis methods have significant limitations. This technique relies heavily on pathologists' visual observation and manual evaluation, which are not only time-consuming and labor-intensive but also highly susceptible to subjective factors.[3] It is even impossible to perform a fully quantitative analysis of markers with varying intensities across all cells in a whole slide. Different pathologists—or even the same pathologist at different times—may achieve differing assessment results when evaluating the same tissue section, which undoubtedly reduces the accuracy and consistency of analysis results. Such subjectivity and inconsistency not only affect the accuracy of diagnosis but may also misinform the formulation of treatment plans, further impacting the patient's treatment outcome and prognosis.[4,5]

Breakthroughs in the field of medical image processing in recent years due to the development of deep learning (DL) technologies make it possible to address this concern. DL algorithms, especially advanced image recognition and processing techniques such as convolutional neural networks (CNNs), have demonstrated exceptional capabilities in image segmentation and recognition.[6] By training on a large amount of histopathological image data, DL models can automatically and accurately identify and segment tumor tissues from normal tissues,[7,8] further precisely recognizing the expression of specific tumor markers.[9] The application of DL has greatly improved the accuracy and efficiency of IHC analysis and provided a more objective and reliable means for identifying tumor markers.[10]

DL algorithms also exhibit significant advantages in terms of quantitative analysis. Traditional IHC methods rely on manual assessment of staining intensity and distribution, which is both time-consuming and highly subjective, making it difficult to ensure consistent results.[11] However, DL algorithms can achieve precise quantification of tumor marker expression through intelligent analysis of staining features in images.[12] This intelligent analysis method improves analytical objectivity, reduces the influence of human factors, and provides more accurate and reliable data support for tumor grading, staging, and

* Corresponding author.
E-mail address: christinewang@pathoai.com (X. Wang).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.intonc.2025.06.001
Received 11 March 2025; Received in revised form 8 June 2025; Accepted 14 June 2025
Available online 30 June 2025
2950-2616/© 2025 The Author(s). Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

prognosis assessment.[13–15] This holds immense significance in formulating personalized treatment plans, enhancing treatment efficacy, and improving patient prognoses.

Therefore, DL algorithms offer broad aspects for application in tumor marker identification and quantification. They overcome the limitations of traditional IHC analysis, increasing accuracy and efficiency, and provide a more scientific and objective basis for the clinical diagnosis and treatment of tumors.[16] With continuous technological advancements and an expanding scope of application, DL is expected to play an increasingly important role in the field of oncology, bringing more precise and effective treatment plans and better prognoses to tumor patients.[17] The present study developed an IHC assistant software based on existing DL algorithms that automatically quantifies the expression intensity of multiple antibodies through interaction with users. Its effectiveness and reliability were verified through experiments.

In summary, IHC clinical interpretation is a dynamic and evolving field that continues to play a pivotal role in the diagnosis and management of various diseases, and advances in technology and the standardization of protocols are expected to further improve the accuracy and reliability of IHC results. However, challenges remain in terms of standardization, antibody quality, and interpretation variability, which require ongoing research and collaboration among pathologists and other healthcare professionals.

2. Related work

DL has facilitated the interpretation of immunohistochemical staining and led to more precise assessment of tumor markers, an achievement that significantly supports clinical research on post-treatment outcomes for patients with tumors. In 2019, Qaiser et al. used an attention-based model to assess the expression of human epidermal growth factor receptor 2 (HER2) in invasive breast cancer. Based on deep reinforcement learning, their novel model treats HER2 IHC scoring as a sequential learning task. The algorithm performed better than other DL CNN algorithms on public datasets for HER2 and Glyoxalase I.[18]

The complexity of whole-slide images (WSIs) indeed poses significant challenges to computer-aided diagnosis. The images contain billions of pixels and are often accompanied by various artifacts, such as bubbles, pen annotations, folded tissue, and blurred areas, which can interfere with the accuracy of algorithms, thereby affecting diagnostic outcomes. In 2023, Gallagher-Syed et al. used a U-Net to implement an automatic segmentation method for IHC-stained synovial tissue in rheumatoid arthritis. This method is trained on a carefully curated, heterogeneous, multicenter real-world clinical dataset called R4RA, which encompasses multiple types of IHC staining. The model achieved a Dice score of 0.865 and successfully segmented different types of IHC staining while also addressing variations in color, intensity, and common WSI artifacts from different clinical centers.[19] In 2024, Jabar et al. introduced a fully automated supervised DL pipeline for WSI quality assessment. This pipeline employs a fusion model that combines CNN and vision transformers (ViT) to detect and exclude WSI regions with artifacts, ensuring that only qualified WSI regions are used to build DL-based computational pathology applications. The pixel-based segmentation model classifies WSI regions as qualified or unqualified based on the presence of artifacts. The model was trained on a large and diverse dataset and validated using internal and external data from various human organs, scanners, and hematoxylin and eosin (H&E) staining protocols. Both quantitative and qualitative evaluations demonstrated the superiority of the proposed model.[20] Ganz et al. employed DL to quantify the expression of IHC marker PHH3, which in turn was used to quantify mitotic figures in H&E samples from the same patient, serving as a measure of tumor cell proliferation. This method addressed the issue of low inter-rater agreement in identifying mitotic figures on H&E slides and simultaneously mitigated the problem of noise-contaminated annotations, which can affect DL algorithms trained directly on H&E images.[21] Fanous et al. developed a rapid scanning optical microscope named BlurryScope, which utilizes continuous image acquisition and advanced DL to offer a compact cost-effective solution for automated inspection and analysis of tissue sections. BlurryScope streamlines the entire process, from scanning images to stitching and cropping areas of interest, and even includes automated HER2 score classification. This innovative microscope enhances efficiency and practicality in the automated detection and analysis of tissue samples.[22]

The field of IHC-stained cell quantification maintains extremely high precision requirements for cell detection and segmentation. Issues such as focusing, cell overlap, and density can lead to unclear cell contours, causing an algorithm to produce incorrect results. In 2020, Weigert et al. developed StarDist, an algorithm for 3D object detection and segmentation in microscopy using star-convex polyhedra.[9] Accurate detection and segmentation of cell nuclei in 3D fluorescence microscopy is crucial in biomedical research, but existing automated methods often struggle with low signal-to-noise ratios and dense nuclei packing. Walter et al. introduced an instance segmentation method specifically designed for overlapping objects utilizing star-convex polygons, serving as an extension to StarDist. The key novelty of their approach lies in identifying pixels where objects overlap and leveraging that information to enhance proposal sampling and circumvent the suppression of proposals for genuinely overlapping objects.[23] In 2024, Tommasino et al. proposed NuLite, a faster, more lightweight algorithm for nuclear instance segmentation that employs a U-Net-like architecture specifically designed on top of Fast-ViT, a state-of-the-art lightweight CNN.[24]

Inspired by those approaches to applying DL, we extended the CellViT for cell segmentation and classification. CellViT employs a pre-trained ViT encoder that excels in various computer vision tasks and enhances segmentation performance. Its U-Net architecture, featuring an encoder-decoder network structure, efficiently and accurately segments cell instances and effectively extracts both high-level and low-level features, improving segmentation accuracy. Our challenges included finding efficient polyhedra representations, adapting to anisotropic voxel sizes, and computing intersections. We focused on accurately detecting and segmenting cell nuclei. CellViT achieves outstanding performance on the PanNuke dataset,[25] with an F1 score of 0.83. The PanNuke dataset is a challenging benchmark for cell instance segmentation and includes cell nuclei from various tissue types, covering 19 different anatomical structures, including tissues such as the kidney, liver, breast, colon, esophagus, lung, and many others. The dataset contains a total of 205,343 labeled nuclei, with each nucleus belonging to one of the following five categories: (1) tumor cells, (2) inflammatory cells, (3) connective tissue cells (fibroblasts and other connective tissue-related cells), (4) dead cells (necrosis or apoptosis), and (5) epithelial cells. In 2024, Koyuncu et al. used CellViT to perform DL and image analysis on lung sections stained with CD20 IHC, achieving precise detection of CD20+ cells.[26] Experimental results demonstrate that this model matches the original CellViT [27] (another state-of-the-art method) in terms of panoptic quality and detection, while being smaller in terms of giga floating-point operations per second and other parameters.

The application of DL algorithms in pathological research—and even the use of trained algorithms to achieve cell quantification and more complicated tasks—has become increasingly sophisticated. However, CellViT was trained on H&E-stained datasets, whereas IHC utilizes hematoxylin-diaminobenzidine (H-DAB) staining, so applying CellViT directly to IHC images will significantly reduce its accuracy. IHC tasks require interpreting different types of cell expression based on H-DAB staining and then quantifying the results to obtain the slide outcomes. Additionally, different antibodies exhibit specificity in their expression sites within cells, and the intensity of some antibodies needs to be re-quantified, which can be subjective. Therefore, interpreting IHC using traditional computer-aided diagnosis methods poses distinct

difficulties. In our work, we converted the images from red, green, and blue (RGB) color to hue, saturation, and value, which better aligns with the understanding of color. We then employed color deconvolution techniques to separate the images into sub-images of those stained with hematoxylin and those stained with DAB. The hematoxylin-stained sub-image visually resembles H&E images, preventing a decrease in the accuracy of CellViT. The DAB-stained sub-image contains information on the intensity of antibody expression. By quantifying this sub-image and combining it with the cell location information from the hematoxylin sub-image, we achieved quantification of antibody expression in cells. Users can interactively adjust the thresholds of the quantification results to distinguish cells with different expression intensities and obtain the final slide interpretation outcomes.

3. Material and methods

This section mainly introduces our pipeline for processing IHC WSIs. Figure 1 outlines the digital image processing approach, including color deconvolution, nuclear segmentation, membrane and cytoplasm segmentation, and quantification of expression intensity.

3.1. Color deconvolution

In IHC staining, commonly used stains include hematoxylin and DAB. Hematoxylin, appearing blue, is typically used to stain cell nuclei whereas DAB is used to label the target antigen, appearing brownish-yellow.[28] The combined use of these two stains allows for clear observation of the cell structure and the distribution of the target antigen under a microscope. The color deconvolution algorithm proposed by Ruifrok et al. indeed has significant application value in stain separation and quantification for IHC-stained RGB digital images.[29] The principle of this algorithm is as follows: In the RGB color model, the camera captures the intensity of red light, green light, and blue light for each pixel. According to Lambert-Beer's law, each pure color dye has specific absorption factors for light in the three RGB channels. The relative light intensity in each channel has a nonlinear relationship with the dye concentration, so directly using the light intensity values makes it difficult to accurately determine the contribution of each dye to the light intensity.[30] The orthogonal normalization of optical density (which describes the degree to which light is absorbed as it passes through a material) can overcome the problem of the nonlinear relationship between light intensity and dye concentration.[31] Orthogonal normalization of optical density can address the issue of the nonlinear relationship between light intensity and dye concentration, thereby enabling the separation and measurement of the concentration of each dye.[32] After color deconvolution processing, the IHC image can be separated into two sub-images stained with hematoxylin and DAB, respectively.

3.2. Nuclear segmentation

Because hematoxylin primarily stains the cell nuclei, the hematoxylin-stained channel image obtained through color normalization is processed to avoid the influence of DAB staining on nuclear segmentation results. CellViT is used for nuclear segmentation. Due to the immense size of WSIs and computer's limited graphics memory, the WSIs must be cropped into multiple tiles, as DL algorithms cannot segment cell nuclei across the entire WSI in one pass. Additionally, the input images' resolution, which describes the physical size that each pixel in the WSI image represents in the real world, can influence the results of CellViT. Our pathology slide scanner, the Sqray SQS-120P (Sqray, China), had a scanning resolution of 0.105 μm/pixel at 40× magnification, meaning that each pixel in the digital image scanned by the 40× objective lens represents a 0.105 × 0.105 μm^2 area on the actual specimen. In our pipeline, WSIs scanned by the SQS-120P scanner at 40× objective magnification were cropped into image tiles of 2,048 × 2,048 pixels. Each tile corresponded to a physical area of 215.04 × 215.04 μm^2 and was subsequently upsampled to 512 × 512 pixels. This tile is separated into two sub-images stained with hematoxylin and DAB after color deconvolution. The sub-image stained with hematoxylin is used as the input for CellViT, which then outputs the contours of the cell nuclei in the image. Each contour represents an independent instance of a cell nucleus, containing the coordinate information of the outer boundary of that nucleus within the image.

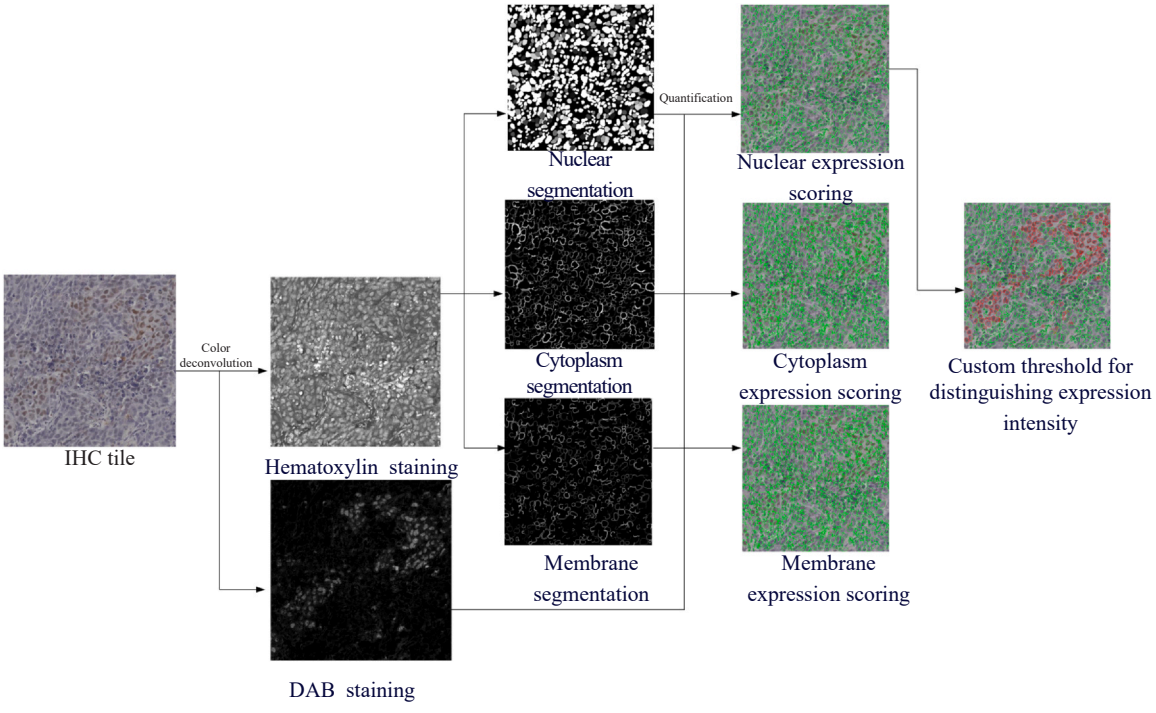


Figure 1. Overview of the IHC WSI analysis scheme.
Abbreviations: 1. IHC: immunohistochemistry 2. DAB: diaminobenzidine.

CellViT relies on open-source codes, and we used the publicly available pre-trained weights CellViT-SAM-H. Detailed information and source codes can be obtained through GitHub (codes available at: https://github.com/TIO-IKIM/CellViT).

3.3. Membrane and cytoplasm segmentation

The H-DAB staining fails to clearly delineate the contours of the cytoplasm and cell membrane, primarily due to the following reasons: (1) Chromatin within the nucleus is primarily composed of DNA, which adopts a double-helix structure with outward-facing phosphate groups that confer a negative charge and acidic properties, facilitating binding to positively charged hematoxylin and thus staining the nucleus; (2) The cytoplasm exhibits alkalinity and has a strong affinity for acidic dyes, hindering binding to hematoxylin and resulting in unstained cytoplasm. Thus, it is difficult to visually discern the complete structure of membranes and cytoplasm in non-expressing and weakly positive cells, making it challenging for algorithms to accurately locate these regions.

The cell cytoplasm segmentation is performed using a seeded region growing algorithm,[33] with the cell nuclei serving as the initialized seeds, which are independent of another. Iteratively applying morphological dilation to the seed pixels multiple times depicts the cell membrane and cytoplasm. By extending the regions around the nuclei, the corresponding cell membranes and cytoplasm can be delineated, ensuring that they do not deviate excessively from the nuclei. The refined pseudocode for the region growing algorithm is as follows:

Algorithm 1. Region growing

```
Input: Image (pixel matrix), seed point set, similarity criterion function, growth stopping condition function
Output: Segmented image
for each seed point do:
    add the seed point to the queue/stack and mark its position in the label matrix with the current region's label (e.g., an incrementing integer starting from 0)
    while the queue is not empty do
        dequeue a pixel from the queue/stack (referred to as the current pixel)
        for each neighboring pixel of the current pixel (based on 4-connectivity or 8-connectivity) do:
            if the neighboring pixel is unmarked in the label matrix (i.e., it is −1) then
                calculate the similarity between the current pixel and the neighboring pixel using the similarity criterion function
                if the similarity satisfies the growth criterion (i.e., the similarity criterion function returns true) then
                    add the neighboring pixel to the queue/stack
                    mark the position of the neighboring pixel in the label matrix with the current region's label
                end if
            end if
        end for
        if the neighboring pixel is already marked but with a different label than the current pixel (indicating potential region merging, which may not be considered in some applications) then
            handle region merging according to requirements, e.g., update the label matrix to merge region labels
        end if
    end while
end for
return the label matrix as the segmented image
```

3.4. Quantification of expression intensity

Protein expression can be quantified by combining segmentation results with the DAB channel separated through deconvolution staining.[34] The process of slide preparation and scanner imaging means that even areas on the slide that are not stained with DAB on the slide will not show a value of 0 in the DAB staining channel.[35] Thus, one must calculate the median grayscale value of the DAB staining channel for the entire slide image, record that median value, and set the grayscale values of all patches in the DAB staining channel that are below that value to 0 to remove noise. Two main types of segmentation patterns emerge depending on the location of antibody expression: nuclear expression results in a patchy pattern, whereas membrane and cytoplasmic expression results in a ring-like pattern. Each independent connected domain in the segmentation results represents a separate cell, and the expression intensity is calculated individually for each cell. The independent cell segmentation results are combined with the DAB staining channel using a bitwise AND operation to obtain the expression status at the corresponding location of the cell. The formula for calculating the expression intensity of nuclear expression E_{cell} is as follows:

$$E_{cell} = \frac{mean(D_{cell})}{max(D_{wsi})} \tag{1}$$

For membrane and cytoplasmic expression E_{cell}, the formula is:

$$E_{cell} = \frac{\max(D_{cell})}{max(D_{wsi})} \tag{2}$$

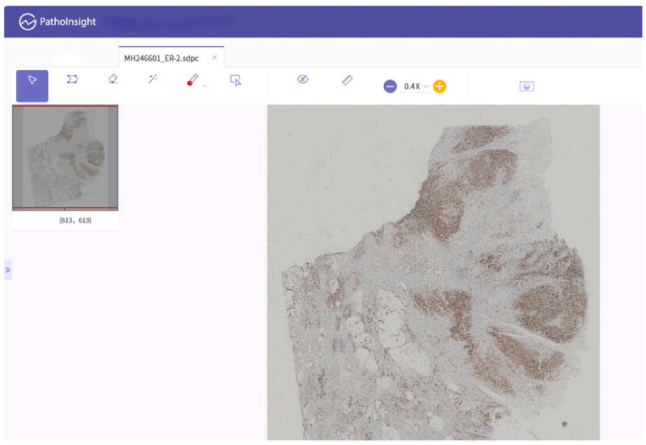


Figure 2. User-defined thresholds for grading antibody expression intensity.

D_{cell} represents the grayscale value of DAB staining for that cell, and D_{wsi} represents the grayscale value of DAB staining for the slide. The maximum grayscale value of DAB staining for the cell is used in the calculation for membrane and cytoplasmic expression, because compared to nuclear expression, membrane and cytoplasmic expression are more complex than nuclear expression, and often only parts of the cell membrane or cytoplasm exhibit staining expression within the cell.

4. Results

4.1. IHC quantification analysis via developed software

We developed image analysis software based on C++ and OpenSlide (version 2.1),[36] an open-source WSI processing library, with algorithms deployed using PyInstaller. Our software is deployed on a server host with the following hardware environment: i7-13700 CPU, 32 GB RAM, RTX4060-8GB GPU, and 512 GB HDD. The software environment used the Windows 10 operating system, with CUDA version 11.7, Python version 3.10.12, TensorFlow version 2.12.0, and PyTorch version 1.13. For slides with a scanning resolution of 0.105 μm/pixel and a size of approximately 200,000 × 200,000 pixels, the average analysis time was approximately 12.46 min. The software was enhanced with a threshold option. For samples requiring the assessment of staining intensity, users can customize and set multiple thresholds to differentiate the levels of antibody expression, ranging from weak to strong (Figure 2); the software marks cells of different intensities with dots of different colors.

4.2. Slide interpretation

To interpret IHC, a pathologist examines the stained tissue slides under a microscope, evaluating various aspects of the staining (e.g., the intensity, pattern, and location of the stain) to derive meaningful information about the tissue's biology and pathology. Sixteen slides including nuclear and membrane expression were selected to examine both staining methods. These surgical resection tissues were obtained from Guangzhou Huitong Yidian Pathological Diagnosis Center Co., Ltd., comprising lesion samples submitted by 16 breast cancer patients. The patients authorized the company to use samples for disease diagnosis and future scientific research, including HER2 and estrogen receptor immunohistochemistry quantification/qualification, with stipulation against commercial use of individuals' disease status/data. The slides were scanned into WSIs to be interpreted twice by the same pathologist with the assistance of our image analysis software. Figure 3 shows the results after correction following manual slide interpretation. The interpretation results of these slide are shown in Tables 1 and 2.

We conducted a statistical analysis to compare the quantitative method with manual review in terms of the sensitivity and specificity in distinguishing between cells and non-cells (Table 3), tumor cells and stromal cells (Table 4), assessing cellular intensity for nuclear/membrane expression (Tables 5 and 6).

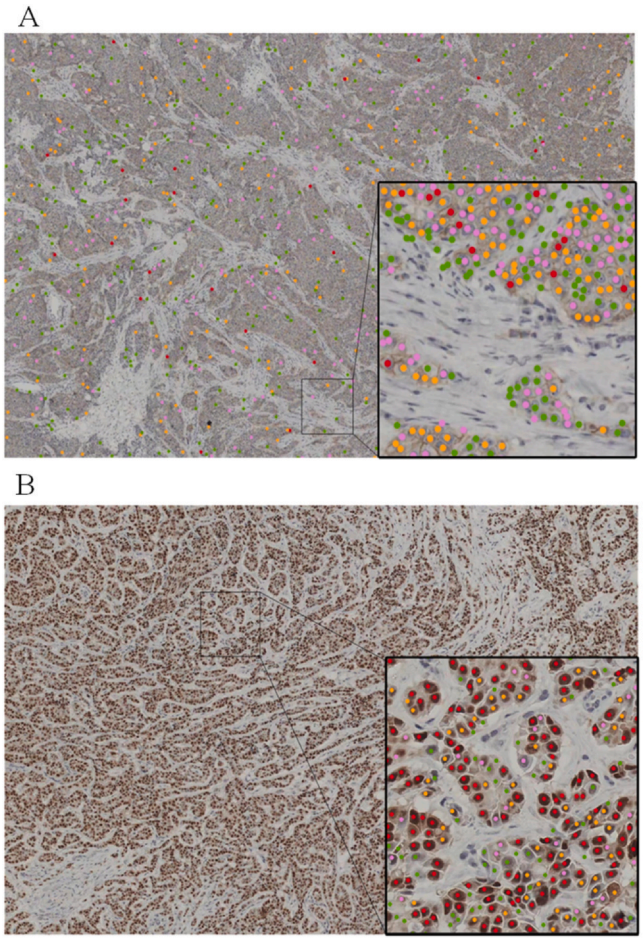


Figure 3. Overview of manual slide interpretation. The IHC specimen samples are a HER2 (A) and an ER (B) both of invasive ductal carcinoma of the breast and primarily displaying the tumor cell markers with varying expression intensities after correction. Red dots represent an expression intensity of 3+; orange dots, an expression intensity of 2+; yellow, an expression intensity of 1+; and green, a negative expression.
Abbreviations: 1. HER2: human epidermal growth factor receptor 2 2. ER: estrogen receptor.

5. Discussion

In the clinical interpretation of IHC slides, the diversity of cellular morphologies (Figure 4) and variations in staining intensity (Figure 5) pose major challenges that arise from cells' complexity and the need for precise and accurate diagnosis based on IHC results. Besides, variability in antibody quality and staining protocols can lead to inconsistencies in IHC results, resulting in difficulty in comparing results between different laboratories or studies. Therefore, the effectiveness and reliability of quantification methods for different cell morphologies and staining techniques are crucial in evaluating their clinical value and significance.

5.1. Effectiveness of IHC quantification method

The IHC quantification method's effectiveness is reflected primarily in its accuracy in counting cell numbers, differentiating between tumor and stromal cells, distinguishing between different expression patterns, and assessing expression intensity under varying nuclear-to-cytoplasmic ratios. According to Table 3, the method demonstrates excellent performance in

Table 1
Results of 16 IHC slides with nuclear and membrane expression as interpreted by physician with the assistance of the software.

Sample ID	Tumor area (mm^2)	Expression	Stromal cells	Non-cellular components	AI-labeled tumor cells	Corrected total cells	Strong positive	Medium positive	Weak positive	Negative
1	0.726	2+	464	8	4,771	4,768	10	1,772	1,256	1,730
2	0.711	0	276	12	4,317	4,336	0	0	13	4,323
3	0.315	1+	457	17	1,495	1,504	1	112	201	1,190
4	0.454	1+	725	5	1,689	1,721	0	5	469	1,247
5	0.668	0	1,872	0	2,209	2,209	0	0	0	0
6	0.354	3+	347	39	1,846	1,845	915	125	494	494
7	0.363	2+	231	16	2,169	2,163	61	482	650	970
8	0.855	3+	1,339	12	3,122	3,125	1,831	147	297	850
9	0.889	2–3+	540	30	5,470	5,490	1,758	1,807	959	966
10	1.425	1+	2,346	69	4,427	4,435	0	82	939	3,414
11	0.845	3+	139	42	7,625	7,329	2,903	1,259	1,263	1,904
12	0.703	1–2+	976	136	2,976	3,037	522	895	785	835
13	2.020	-	6,242	98	3,806	3,806	0	0	0	3,806
14	1.787	2–3+	1,872	81	9,083	9,077	2,451	1,114	2,057	3,455
15	1.380	3+	3,842	247	2,762	2,769	2,419	171	123	56
16	0.812	1–2+	574	166	6,482	6,482	974	977	1,100	3,431

Table 2
Pathologist comments on slide interpretation using our software.

Sample ID	Pathologist comments
1	Cancer cells are medium-sized, with few stromal components and high accuracy in cancer cell labeling.
2	Cancer cells are medium-sized with high consistency, with few stromal components and high accuracy in cancer cell labeling.
3	Rich stromal fibrosis; stromal fibrosis and stromal cells are easily misclassified as cancer cells.
4	Rich stromal fibrosis; stromal fibrosis and stromal cells are easily misclassified as cancer cells.
5	Rich inflammatory cells in the stroma; inflammatory cells are misclassified as cancer cells.
6	Abundant tissue gaps; rich stromal fibrous tissue; fibrous tissue and fibroblasts are misclassified as cancer cells.
7	High accuracy in distinguishing cancer cells of different intensities.
8	Biopsy tissue with rich inflammatory cells and fibrous tissue in the stroma; inflammatory cells, fibrous tissue, and fibroblasts are misclassified as cancer cells.
9	Cancer cells are medium-sized, with few stromal components and high accuracy in cancer cell labeling.
10	Rich tumor stroma with extensive inflammatory cell infiltration.
11	High accuracy in distinguishing cancer cells of different intensities.
12	Rich stromal fibrosis; collagen fibers are easily misclassified as cells.
13	Rich tumor stroma with extensive inflammatory cell infiltration.
14	Large cancer cells with a tendency for multiple labeling.
15	Small cancer cells with rich stromal fibrosis, large tissue gaps, stromal fibrosis, and stromal fibrosis; stromal cells are easily misclassified as cancer cells.
16	Patchy necrosis in cancer tissue; necrotic cells are misclassified as cancer cells.

Table 3
The confusion matrix of the quantitative method in distinguishing between cells and non-cells.

Ground truth	Prediction		Sum
	Cell	Non-cell	
Cell	84,313	2,178	86,491
Non-cell	106	872	978
Sum	84,419	3,050	87,469
Sensitivity		0.9987	
Specificity		0.8916	

Table 4
The confusion matrix of the quantitative method in distinguishing between tumor cells and stromal cells.

Ground truth	Prediction		Sum
	Tumor	Stromal	
Tumor	61,767	304	62,071
Stromal	2,482	19,760	22,242
Sum	64,249	20,064	84,313
Sensitivity		0.9614	
Specificity		0.8884	

identifying cell-like objects, with a sensitivity of 0.9987, indicating that it accurately detects almost all actual cells. However, the specificity is relatively low (0.8916), suggesting a higher rate of false positives in identifying non-cells. Figure 6 shows a sample analyzed by CellViT in which instances are observed of cell points being missed and non-cell components being incorrectly identified as cells. This poor performance in identifying non-cell components (including necrotic tissue, cellular debris, and secretions) is attributed to the fact that the images input into CellViT were hematoxylin stained, lacking the staining information from eosin, which prevented CellViT from achieving its optimal performance. Furthermore, the cells that CellViT overlooked often have lighter staining compared to the cells surrounding them, further explaining the algorithm's failure to detect them.

In distinguishing between tumor cells and stromal cells, the method shows high sensitivity (0.9614) and specificity (0.8884), as shown in Table 4. This indicates that the method can accurately identify tumor cells while minimizing false positives for stromal cells. The method also demonstrates high sensitivity and specificity for both nuclear and membrane expression across different staining sites, indicating stable performance (Tables 5 and 6).

5.2. Reliability of IHC quantification method

A reliable method should demonstrate stable performance across different scenarios. The method in this research shows high sensitivity and specificity for both nuclear and membrane expression, indicating

Table 5
The confusion matrix of the quantitative method in distinguishing cellular intensity for nuclear expression.

Ground truth	Prediction				Sum
	Negative	Weak	Moderate	Strong	
Negative	33,842	13	15	6	33,876
Weak	47	7,633	0	3	7,683
Moderate	7	18	6,276	59	6,360
Strong	0	12	13	11,012	11,037
Sum	33,896	7,676	6,304	11,080	58,956
Sensitivity of negative		0.9984	Specificity of negative		0.9987
Sensitivity of weak		0.9979	Specificity of weak		0.9970
Sensitivity of moderate		0.9883	Specificity of moderate		0.9980
Sensitivity of strong		0.9939	Specificity of strong		0.9983

Table 6
The confusion matrix of the quantitative method in distinguishing cellular intensity for membrane expression.

Ground truth	Prediction				Sum
	Negative	Weak	Moderate	Strong	
Negative	16,259	178	19	0	16,456
Weak	70	3,058	48	257	3,433
Moderate	2	23	2,591	33	2,649
Strong	0	0	12	2,807	2,819
Sum	16,331	3,259	2,670	3,097	25,357
Sensitivity of negative		0.9879	Specificity of negative		0.9919
Sensitivity of weak		0.8906	Specificity of weak		0.9909
Sensitivity of moderate		0.9780	Specificity of moderate		0.9966
Sensitivity of strong		0.9063	Specificity of strong		0.9869

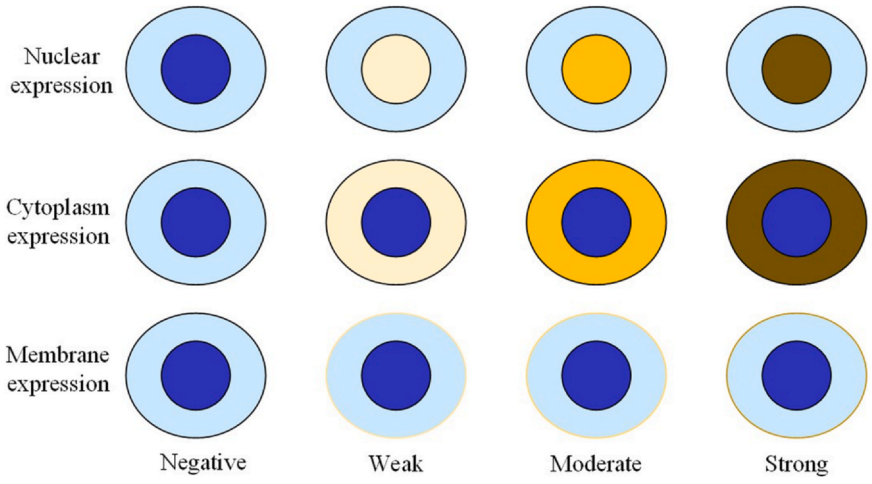


Figure 4. Illustrative diagrams of cellular morphology with varying nuclear-to-cytoplasmic ratios and morphologies.

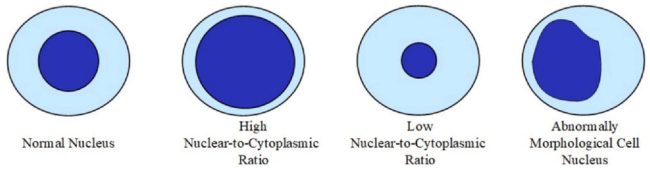


Figure 5. Illustrative diagram of varying intensities of expression in different regions of a cell.

stable performance, but its performance is slightly weaker in sensitivity to weak membrane expression (0.8906; Tables 5 and 6). The performance differences between membrane and nuclear expression indicate that the reliability of this method on membrane expression data is slightly lower than on nuclear expression. The reasons for these limitations include the current algorithm's difficulty in distinguishing between negative, weak, moderate, and strong positive statuses, as well as factors related to the slide preparation process and the brightness during scanning.

5.3. *Efficiency in clinical use*

When the software is used in clinical deployment, it struggles to keep up with the high volume of slide reading demands, primarily due to the limited computational power of the consumer-grade GPU (RTX4060) and the slow data transfer speed of the mechanical hard drive. These factors lead to frequent data I/O operations, which consume a significant amount of analysis time. Additionally, the amount of

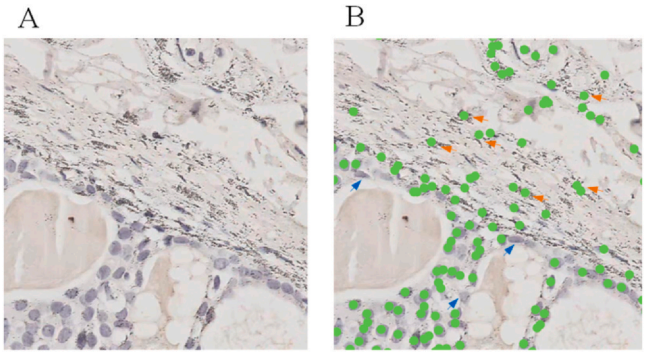


Figure 6. Sample images showing errors in CellViT analysis. Figure 6A shows the original image prior to algorithmic analysis; Figure 6B, the outcomes of CellViT identification. Orange arrows indicate examples of non-cell components that the algorithm mistakenly identified, and blue arrows show cell points that CellViT failed to detect.

GPU memory also determines the number of subprocesses that can be used for algorithmic analysis. The 8 GB of GPU memory and the large parameter size of the CellViT-SAM-H weights limited our system to using only six subprocesses for acceleration.

5.4. Future work

The current IHC quantification method based on DL algorithms and image processing techniques offers a highly effective and reliable solution for automated quantitative analysis of IHC staining, but some areas still need improvement. Currently, the method cannot distinguish between benign and malignant tumor cells, and its identification of tumor cells, stromal cells, and inflammatory cells is not precise enough. It also misidentifies necrotic cells in non-cell components as tumor cells and is unable to accurately define the boundaries of cancerous tissue. Therefore, future work will focus on addressing these limitations to make the method more comprehensive and enhance the efficiency of IHC research. Specifically, to tackle the issues of missed cells, distinguishing tumor cells and ensure our method's reliability, we will endeavor to collect as many IHC slides produced by different types of scanners as possible to create our own annotated dataset for validation and fine-tuning. We will then fine-tune and train CellViT or other more effective algorithms to strengthen the software's performance and achieve the goal of fully automated diagnosis.

To address the issue of difficulty in distinguishing between weakly positive and negative cells, it is necessary to adjust the existing cell expression intensity quantification scheme. Currently, the expression intensity is calculated based on algorithmic analysis. The results from the current algorithm will be collected, and the classifications of weakly positive and negative cells will be corrected by pathologists. This corrected dataset will serve as a fine-tuning dataset for CellViT. With CellViT trained on this dataset, the model will be able to directly output cell categories, expression intensities, and cell locations. Reducing the number of preprocessing steps can also promote the processing speed of the algorithm. It is necessary to enhance the system's cell discrimination accuracy to enable continuous full-quantitative analysis of IHC slides.

To move forward, we plan to upgrade to a solid-state drive to improve data transfer speeds. We also intend to deploy the algorithm in the cloud and use a GPU array to process the data. Moreover, we will implement the preprocessing steps that currently run on the CPU (such as color deconvolution and region-growing algorithms) using PyTorch operators and accelerate these processes with the GPU. Pruning and quantizing the deployed model weights is also crucial for enhancing the efficiency of analysis. After these optimizations are implemented, the software's efficiency can be aligned with the high volume of clinical demands, thereby enhancing pathologists' workflow efficiency.

Furthermore, we aim to incorporate biomarker analysis into the existing tumor, node, metastasis (TNM) staging classification system, upgrading it to a TNM-biomarkers framework. This enhancement will provide greater convenience for precision diagnosis research and contribute to more comprehensive clinical assessments. These advancements will further improve the method's clinical value and significance, contributing to more precise treatment plans and better patient outcomes.

6. Conclusions

The IHC quantification method developed in this study, which is based on DL algorithms and image processing techniques, offers a highly effective and reliable solution for automated quantitative analysis of IHC staining. It overcomes the limitations of traditional manual interpretation methods, providing more accurate, consistent, and objective results. The method has broad application prospects in the clinical diagnosis and management of various diseases, contributing to more precise treatment plans and improved patient outcomes. Future research may focus on further optimizing the method, expanding its application range, and addressing remaining challenges in standardization, antibody quality, and interpretation variability.

CRediT authorship contribution statement

Yongjian Deng: Writing – review & editing, Validation, Supervision, Resources, Investigation, Data curation. **Xiaomei Wang:** Supervision, Project administration, Conceptualization. **Bojun Cai:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation.

Ethical statement

The study adheres to all relevant ethical guidelines and regulations. The research involves the analysis of de-identified IHC-stained tissue samples obtained from surgical resection, which were provided by Guangzhou Huitong Yidian Pathological Diagnosis Center Co., Ltd. The use of these samples for research purposes was approved by the institutional review board, and all necessary consents were obtained from the patients.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships, which may be considered as potential competing interests: Xiaomei Wang and Yongjian Deng are currently employed by Hangzhou MediGenPro Biotechnology Co., Ltd.; Xiaomei Wang and Bojun Cai are currently employed by Hangzhou PathoAI Technology Co., Ltd.

References

1. Tang L, Huang ZP, Mei H, Hu Y. Insights gained from single-cell analysis of chimeric antigen receptor T-cell immunotherapy in cancer. *Mil Med Res.* 2023;10(1):52.
2. Nambirajan A, Jain D. Immunohistochemical detection of p40 expression in lung cancer clinical samples. *Methods Mol Biol.* 2021;2279:13–21.
3. Hamilton PW, Wang Y, Boyd C, et al. Automated tumor analysis for molecular profiling in lung cancer. *Oncotarget.* 2015;6(29):27938–27952.
4. Wu S, Li X, Miao J, et al. Artificial intelligence for assisted HER2 immunohistochemistry evaluation of breast cancer: A systematic review and meta-analysis. *Pathol Res Pract.* 2024;260:155472.
5. Yue M, Zhang J, Wang X, et al. Can AI-assisted microscope facilitate breast HER2 interpretation? A multi-institutional ring study. *Virchows Arch.* 2021;479(3):443–449.
6. Malebary SJ. Intelligent machine learning based brain tumor segmentation through multi-layer hybrid U-net with CNN feature integration. *Comp Mater Continua.* 2024;79(1).
7. Lerousseau M, Vakalopoulou M, Classe M, et al. Weakly supervised multiple instance learning histopathological tumor segmentation. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23*. Springer; 2020:470-479.
8. Farshid G, Armes J, Dessauvagie B, et al. Independent validation of a HER2-low focused immunohistochemistry scoring system for enhanced pathologist precision and consistency. *Mod Pathol.* 2025;38(4):100693.

9. Weigert M, Schmidt U, Haase R, Sugawara K, Myers G. Star-convex polyhedra for 3D object detection and segmentation in microscopy. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*; 2020:3666-3673.
10. Trullo R, Bui QA, Tang Q, Olfati-Saber R. Image translation based nuclei segmentation for immunohistochemistry images. *MICCAI Workshop on Deep Generative Models*. Springer; 2022:87–96.
11. Lahiani A, Gildenblat J, Klaman I, Navab N, Klaiman E. Generalizing multistain immunohistochemistry tissue segmentation using one-shot color deconvolution deep neural networks. *arXiv preprint arXiv:180506958*. Published online 2018.
12. Segerer FJ, Nekolla K, Rognoni L, et al. Novel Deep Learning Approach to Derive Cytokeratin Expression and Epithelium Segmentation from DAPI. *arXiv preprint arXiv:220808284*. Published online 2022.
13. Sekhar A, Goel V, Jain G, et al. HER2 and FISH Status Prediction in Breast Biopsy H&E-Stained Images Using Deep Learning. In: *2024 IEEE 24th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE;2024:1-8.
14. Kataria T, Knudsen B, Elhabian SY. StainDiffuser: multitask dual diffusion model for virtual staining. *arXiv preprint arXiv:240311340*. Published online 2024.
15. Priego-Torres BM, Lobato-Delgado B, Atienza-Cuevas L, Sanchez-Morillo D. Deep learning-based instance segmentation for the precise automated quantification of digital breast cancer immunohistochemistry images. *Expert Syst Appl.* 2022;193:116471.
16. Overbeek JC, Partin A, Brettin TS, et al. Assessing Reusability of Deep Learning-Based Monotherapy Drug Response Prediction Models Trained with Omics Data. *arXiv preprint arXiv:240912215*. Published online 2024.
17. Habchi Y, Kheddar H, Himeur Y, et al. Deep transfer learning for kidney cancer diagnosis. *arXiv preprint arXiv:240804318*. Published online 2024.
18. Qaiser T, Rajpoot NM. Learning where to see: A novel attention model for automated immunohistochemical scoring. *IEEE Trans Med Imaging.* 2019;38(11):2620–2631.
19. Gallagher-Syed A, Khan A, Rivellese F, et al. Automated segmentation of rheumatoid arthritis immunohistochemistry stained synovial tissue. *arXiv preprint arXiv:230907255*. Published online 2023.
20. Jabar F, Rasmussen Busund LT, Ricciuti B, et al. Fully Automatic Content-Aware Tiling Pipeline for Pathology Whole Slide Images. *arXiv preprint arXiv: 241116885*. Published online 2024.
21. Ganz J, Marzahl C, Ammeling J, et al. Information mismatch in PHH3-assisted mitosis annotation leads to interpretation shifts in H&E slide analysis. *Sci Rep.* 2024;14(1):26273.
22. Fanous MJ, Seybold CM, Chen H, Pillar N, Ozcan A. BlurryScope: a cost-effective and compact scanning microscope for automated HER2 scoring using deep learning on blurry image data. *arXiv preprint arXiv:241017557*. Published online 2024.
23. Walter FC, Damrich S, Hamprecht FA. Multistar: Instance segmentation of overlapping objects with star-convex polygons. In: *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2021:295-298.
24. Tommasino C, Russo C, Rinaldi AM. NuLite–Lightweight and Fast Model for Nuclei Instance Segmentation and Classification. *arXiv preprint arXiv:240801797*. Published online 2024.
25. Gamper J, Alemi Koohbanani N, Benet K, Khuram A, Rajpoot N. Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In: *Digital Pathology: 15th European Congress, ECDP 2019, Warwick, UK, April 10–13, 2019, Proceedings 15*. Springer; 2019:11-19.
26. Koyuncu D, Tavolara T, Gatti DM, et al. B cells in perivascular and peribronchiolar granuloma-associated lymphoid tissue and B-cell signatures identify asymptomatic *Mycobacterium tuberculosis* lung infection in Diversity Outbred mice. *Infect Immun.* 2024;92(7):e0026323.
27. Hörst F, Rempe M, Heine L, et al. CellViT: Vision transformers for precise cell segmentation and classification. *Med Image Anal.* 2024;94:103143.
28. Wei L, Hua S, Zhang S, Zhang X. DeReStainer: H&E to IHC Pathological Image Translation via Decoupled Staining Channels. In: *MICCAI Workshop on Deep Generative Models*. Springer; 2024:1-10.
29. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol.* 2001;23(4):291–299.
30. Landini G, Martinelli G, Piccinini F. Colour deconvolution: stain unmixing in histological imaging. *Bioinformatics.* 2021;37(10):1485–1487.
31. Landini G, Perryer G. Digital enhancement of haematoxylin- and eosin-stained histological images for red-green colour-blind observers. *J Microsc.* 2009;234(3):293–301.
32. Anglada-Rotger D, Marques F, Pardas M. Color Deconvolution applied to Domain Adaptation in HER2 histopathological images. In: *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*. IEEE; 2023:1-5.
33. Gao H, Zhang X, Fang Y, Yuan J. A line segment extraction algorithm using laser data based on seeded region growing. *Int J Adv Robot Syst.* 2018;15(1): 1729881418755245.
34. Lu Z, Liu Y, Xu J, et al. Immunohistochemical quantification of expression of a tight junction protein, claudin-7, in human lung cancer samples using digital image analysis method. *Comput Methods Programs Biomed.* 2018;155:179–187.
35. Zizzi A, Montironi MA, Mazzucchelli R, et al. Immunohistochemical analysis of chromatin remodeler DAXX in high grade urothelial carcinoma. *Diagn Pathol.* 2013;8:111.
36. 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.