



Leveraging Deep Learning in Real-Time Intelligent Bladder Tumor Detection During Cystoscopy: A Diagnostic Study

Zixing Ye, MD¹, Yingjie Li, MD¹, Yujiao Sun, RN¹, Chengqing He, MD², Guanglin He, MD³, and Zhigang Ji, MD¹

¹Department of Urology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; ²Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; ³Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou, China

ABSTRACT

Background. Accurate detection of bladder lesions during cystoscopy is crucial for early tumor diagnosis and recurrence monitoring. However, conventional visual inspection methods have low and inconsistent detection rates. This study aimed to evaluate the effectiveness of the HRNetV2 deep learning model for intelligent bladder lesion detection, focusing on its performance at different image resolutions.

Patients and Methods. We recruited 94 patients undergoing cystoscopy or transurethral resection of bladder tumor and collected 102 white-light cystoscopy videos between July 2022 and July 2023. Frames containing suspected bladder lesions were manually annotated. HRNetV2, a semantic segmentation model, was used to identify tumor-related morphological features. Images were categorized as high- or low-resolution, and the dataset was split into training and test sets (4:1 ratio). Diagnostic performance was assessed using sensitivity, precision, and mean Dice (mDice) score.

Results. A total of 33,657 frames were annotated. The overall sensitivity and precision for the test set were 91.6% and 91.3%, respectively, with an mDice score of 80.3%. In the high-resolution group, sensitivity and precision were 94.8% and 94.4%, while in the low-resolution group, they were 75.6% and 74.8%. The mDice scores for high- and low-resolution images were 84.7% and 56.6%, respectively.

Conclusions. HRNetV2 demonstrated excellent performance, particularly with high-resolution images, offering high sensitivity and precision in bladder lesion detection. The results suggest that HRNetV2 has significant potential to improve detection accuracy in clinical practice. Future work should focus on optimizing the model further and testing it with larger, multicenter datasets.

Keywords Bladder lesion · Cystoscopy · Transurethral resection of bladder tumor · Artificial intelligence · Deep learning

Endoscopic procedures (cystoscopy and transurethral resection of bladder tumor [TURBT]) are essential investigation methods for the identification and treatment of bladder lesions, particularly in patients with bladder cancer who require long-term follow-up. However, modern cystoscopy continues to rely on the conventional white light technique developed eight decades ago. This reliance can lead to problems, including subjective interpretation and variability in diagnostic outcomes among examiners with different expertise levels or during repeat cystoscopies by the same examiner; this can potentially result in misdiagnosis, the missing of early stage lesions, and the complication of objective comparison of cystoscopy findings across reports.¹ Previous studies have reported that approximately 10–20% of bladder cancers are missed during cystoscopy,² with reported sensitivity and specificity of 68–100% and 57–97%, respectively.^{3,4} Consequently, improving cystoscopy accuracy is essential for detecting lesions and early tumor recurrence. Although blue light and narrow band imaging have transformed endoscopic imaging and significantly improved bladder tumor detection rates,^{5,6} subjective identification of

Zixing Ye and Yingjie Li are contributed equally to this work.

© The Author(s) 2025

First Received: 27 December 2024

Accepted: 30 January 2025

Published online: 6 March 2025

Z. Ji, MD

e-mail: jizhigang@pumch.cn

cystoscopy images with the naked eye is inevitable, and complex procedures limit their further application.

Image recognition and segmentation algorithms based on deep learning have emerged as powerful artificial intelligence technologies widely used in medical practice, including tumor pathological classification and lesion detection using endoscopy.⁷ The convolutional neural network (CNN), a subset of deep learning algorithms, is notable for its robust performance in image classification.⁸ Bladder lesion detection using CNN algorithms is accurate, objective, reproducible, and automatic.⁹ Shkolyar et al. introduced CystoNet, a CNN-based automated image-processing platform for bladder cancer diagnosis, which demonstrated excellent diagnostic performance¹⁰ and has been used for prospective tests.¹¹ Furthermore, Takuya Iwaki's deep learning system was more accurate in diagnosing Hunner-type interstitial cystitis than urologists.¹² Hence, deep learning-based cystoscopy augmentation is expected to significantly promote the clinical practice of comprehensive bladder lesion detection and individualized management.¹¹

Image semantic segmentation has evolved from traditional methods based on shallow features to deep learning techniques, including fully convolutional network (FCN) and its extensions.¹³ A high-resolution network (HRNet) was developed in 2019¹⁴ and subsequently modified to HRNetV2, a superior version of HRNet.¹⁵ HRNetV2 significantly enhances contextual feature extraction and improves semantic segmentation accuracy by resolving downsampling effects and closing semantic gaps in information integration.¹⁶ Moreover, HRNetV2 maintains high-resolution representations throughout the process, without the conventional bottleneck design.¹⁷ Kinshuk et al. reported that HRNetV2 exhibited superior mask detection performance in social distancing calculation tasks.¹⁸ However, no study has been conducted on the HRNetV2 model application in medical image semantic segmentation.

In this study, we utilized the HRNetV2 to detect real-time intelligent bladder lesions during cystoscopy. This study aimed to investigate the potential of deep learning technology in assisting examiners in rapidly and accurately identifying bladder lesions that the naked eye may miss and laying the foundation for further pathological and diagnostic applications.

PATIENTS AND METHODS

Data Collection

This study was approved by the Ethics Committee of Peking Union Medical College Hospital, Chinese Academy of Medical Sciences (I-23PJ1996). The study adhered to the ethical standards of the Helsinki Declaration. We recruited 94 patients who underwent cystoscopy or TURBT at Peking

Union Medical College Hospital and collected 102 white light videos between 1 July 2022 and 1 July 2023. All patients provided written informed consent and complete baseline data. Two urologists confirmed that all the collected videos contained bladder lesions and were randomly divided into training and test sets in a 4:1 ratio. All images containing bladder lesions were extracted from video segments at fixed intervals, and manual annotation of bladder lesion outlines was performed frame by frame. The final labeled results were reviewed and verified independently by two different urologists.

Semantic Segmentation Method

The semantic segmentation method was utilized for bladder lesion detection. We utilized HRNetV2, a modified version of HRNet,¹⁷ as the encoder. FCN¹⁹ served as the decoder in the FCN_HRNetV2 segmentation network formation (Fig. 1). In this configuration, the encoder maintains high resolution, allowing for the extraction of detailed positional features. The HRNetV2 produces feature maps at four scales, which were subsequently unified to a consistent resolution using deconvolution. The decoder restores the feature maps to their original image size.

Data Preprocessing and Training

The original input data resolutions for bladder tumor recognition were 544×672 , 480×720 , and 1080×1920 . Before inputting the data into the network, images of these sizes were standardized using red, green, and blue (RGB) channels, with the mean and variance values of (123.675, 116.28, and 103.53) and (57.12, 57.12, and 57.12), respectively. After normalization, the images were fed into the network, which produced output sizes that matched the original input dimensions. After segmentation, a simple post-processing step was implemented using a 29×29 kernel. This involved an initial morphological opening operation to remove edge spikes in the predicted regions and a closing operation to remove small fragmented areas.

We utilized a batch size 16 on a single NVIDIA Tesla V100 Tensor Core graphics processing unit (GPU) during training. The backbone employed an ImageNet pretrained model optimized using the stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.01. We used a polynomial learning rate decay strategy, ensuring a minimum learning rate of 1×10^{-4} . The training involved cropping the images to 512×512 dimensions. To ensure stability in bladder lesion detection despite differences in size and angles during cystoscopy, we employed image augmentation techniques, including flipping and contrast and brightness enhancement. In addition, we utilized a feature pyramid structure to integrate feature maps of different scales,

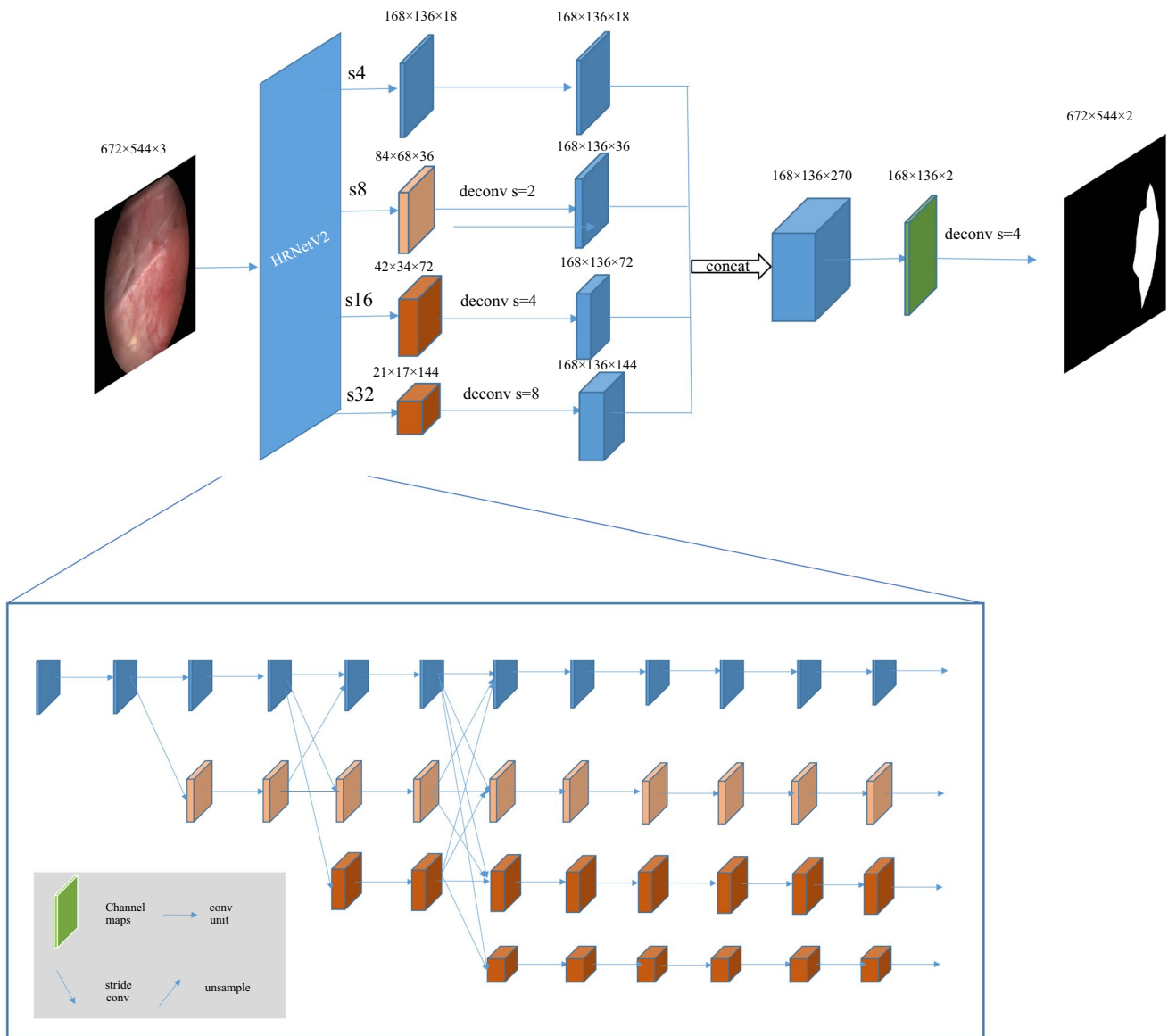


FIG. 1 Architecture of HRNetV2 network

enhancing the performance of the model across diverse scenarios. The final model was achieved after 80,000 iterations. We employed a single V100 tensor core GPU for inference.

Model Evaluation

The performance of the segmentation model was evaluated using intersection over union (IOU), with a 0.1 IOU threshold chosen to prioritize bladder lesion detection over precise localization. The diagnostic performance was further evaluated on the basis of sensitivity and precision in the test set at the specified IOU threshold. In addition, the mean Dice (mDice) score was used as the evaluation metric for the detection performance of the model, characterizing its

ability to detect bladder tumors across different resolution levels.

RESULTS

In this study, 33,657 frames were manually annotated, outlining 37,947 targets. Among these, 75 patients contributed 30,103 targets (26,654 frames) to the training set, while 19 patients contributed 7,844 targets (7003 frames) to the test set. Overall, the diagnostic sensitivity of the neural network model in the test set was 91.6%, with a precision of 91.3% (Table 1). The video increment experiment demonstrated a steady improvement in model performance with increasing data (Table 2).

TABLE 1 Diagnostic performance of HRNetV2 at different resolution levels

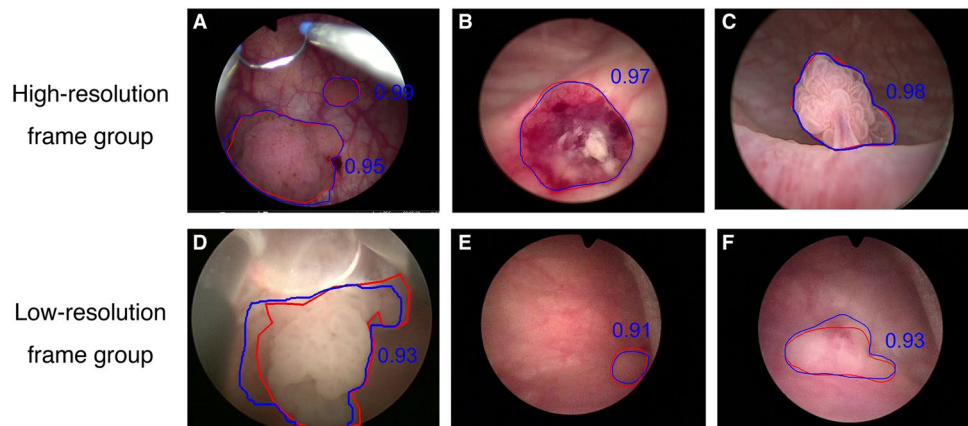
	Training set	Test set		
		High-resolution group	Low-resolution group	Total
Patients	75	–	–	19
Videos	82	–	–	20
Frames	26,654	5897	1106	7003
Ground truth	30,103	6522	1322	7844
Sensitivity	–	94.8%	75.6%	91.6%
Precision	–	94.4%	74.8%	91.3%
mDice score	–	84.7%	56.6%	80.3%

TABLE 2 Results of video increment experiment

Videos trained	Frames trained	Sensitivity (%)	Precision (%)
15	5421	76.7	39.2
68	22,270	91.1	90.1
82	29,528	91.6	91.3

Frames in the test set were artificially categorized into high-resolution and low-resolution frame groups on the basis of lesion clarity and contrast with surrounding tissues to better align with clinical practice and facilitate subsequent work. Frames in the high-resolution group typically exhibited clear morphology and a relatively intact structure, resembling aquatic grass or cauliflower. The low-resolution frames were characterized by an incomplete appearance due to excision, an unclear feature from the camera distance, or turbid fluid interference (Fig. 2). The sensitivity and precision in the high-resolution group were 94.8% and 94.4%, respectively, compared with 75.6% and 74.8% in the low-resolution group. The mDice scores for these groups were 84.7% and 56.6%, respectively, with an overall mDice of 80.3% (Table 1). Supplementary files 1 and 2 show representative demos demonstrating real-time bladder lesion detection at different resolutions.

We investigated the impact of ground truth (GT) area proportion on diagnostic performance. Table 3 shows that the model performed poorly when the GT area proportion was small (0–0.02), with a recall rate of 0.564. Suboptimal performance was observed when the GT area approached

FIG. 2 Representative frames of high-resolution frame group and low-resolution frame group; the numbers represent IOU values**TABLE 3** Diagnostic performance under different GT area proportions

Proportion	Recall	GT	Correct predictions	Total predictions	Sensitivity (%)	Precision (%)
0–0.02	564	1000	745	1230	56.4	60.6
0.02–0.05	1384	1484	1209	1276	93.3	94.7
0.05–0.1	1208	1256	1090	1124	96.2	97.0
0.1–0.2	2307	2330	2144	2195	99.0	97.7
0.2–0.3	1000	1007	1045	1067	99.3	97.9
0.3–0.4	411	414	507	518	99.3	97.9
0.4–0.5	187	187	333	349	100	95.4
0.5–0.6	93	114	154	159	81.6	96.9
0.6–0.7	27	49	24	26	55.1	92.3
0.7–1.0	3	3	4	4	100	100

half of the image (0.5–0.7). However, this conclusion may lack reliability owing to the limited number of targets in this category.

In addition to assessing diagnostic and detection performance, we analyzed the impact of image quality and potential factors contributing to false diagnoses. False positives were primarily due to abnormal mucosal texture features and the misidentification of small targets as lesions. False negatives were attributed mainly to atypical target features, relatively small target sizes, and targets that were too close to the observation point (Fig. 3).

DISCUSSION

Bladder tumors are typically numerous, immobile, and frequently recurring. Papillary tumors are easy to identify. However, they are difficult to classify, whereas flat tumors, including carcinoma in situ (CIS), are difficult to detect and can grow rapidly without intervention.²⁰ Therefore, conducting a thorough and systematic examination of all potential lesions is essential when planning future follow-ups or determining the need for adjuvant treatments.²¹ However, experienced medical personnel may overlook bladder tumors because of their inherent subjectivity and the possibility of interference from operational or observational factors. We employed the recently proposed HRNetV2 neural network model for real-time bladder lesion detection and achieved excellent diagnostic and detection performance. This was evidenced by high sensitivity, precision, and mDice scores.

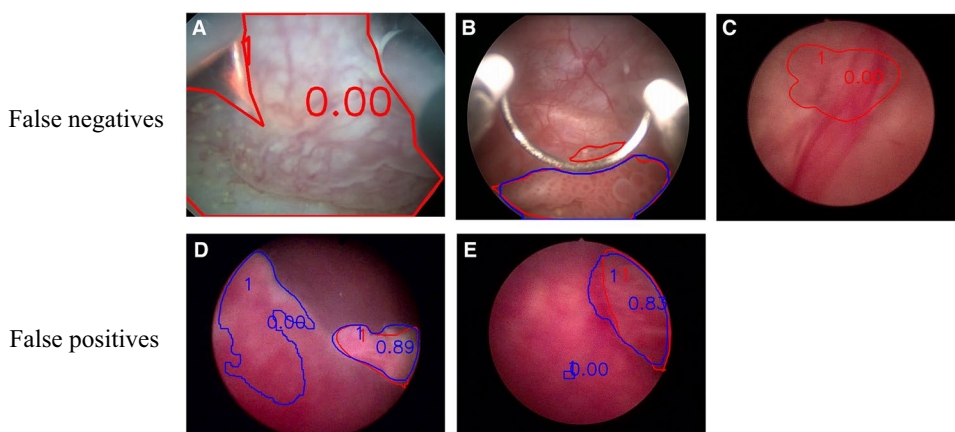
Deep learning techniques for computer vision include semantic segmentation and object detection. Unlike object detection, which provides object location, semantic segmentation precisely delineates the spatial boundaries of targets, providing more information for classification and subgroup analysis.²² Our study aimed to achieve real-time and comprehensive detection of all suspicious lesions within the bladder cavity using deep learning methods. This will serve as the foundation for subsequent lesion evaluation and

therapeutic guidance by testing the diagnostic and detection performance of the deep learning model. Therefore, employing the semantic segmentation model HRNetV2 and selecting patients with suspected bladder lesions rather than those with bladder tumors is consistent with the aim of this study.

Numerous studies have investigated the use of CNNs in the intelligent diagnosis of intravesical lesions. For example, Ikeda et al. conducted a pilot study with 1671 normal and 431 tumor images to diagnose bladder cancer. The deep learning model GoogLeNet achieved a sensitivity of 89.7% and a specificity of 94.0%.²³ Refining algorithms led to superior performance. Zhang et al. investigated attention mechanisms in bladder tumor segmentation using the classical U-Net model, and the new network achieved an improved mDice score of 82.7%.²⁴ Strategies, including expanding sample datasets, contribute to more stable diagnosis performance. Wu et al. developed the cystoscopy artificial intelligence diagnostic system (CAIDS) framework using a pyramid scene parsing network. They collected > 69,000 cystoscopy images from > 10,000 patients to train the model for tumor diagnosis, and the accuracies of different external validation sets all surpassed 97%. Our study demonstrated that the stability of the model improved as the sample size increased. Additionally, CAIDS was found to diagnose carcinoma CIS more accurately than urologists.²⁵ A previous study compared eight deep neural networks for the semantic segmentation of cystoscopy images and found that the Pyramid Attention Network (PAN) model performed better than the others.²² Algorithm enhancement and dataset expansion are crucial for the further advancement of deep learning for bladder lesion detection based on the methodologies employed in various studies.

Cystoscopy uses water for observations, unlike most medical endoscopes, including gastrointestinal endoscopy, which uses air to conduct light. Therefore, cloudy urine, blood mist, and light scattering, commonly encountered in clinical practice, can affect the imaging quality. Low-visibility medical images obscure essential disease details

FIG. 3 Analysis of false negatives and false positives; the main reasons for false negatives are (A) close targets, (B) small size, and (C) atypical features; the main reasons for false positives are abnormal (D) texture and (E) small targets mistaken for lesions; red lines indicate manual annotations, and blue lines represent algorithm predictions; the numbers represent IOU value



and decrease diagnostic accuracy, which are major causes of missed and false detections of deep learning models. We performed subgroup analysis to address the influence of image quality on detection performance. We demonstrated that HRNetV2 exhibits superior sensitivity and precision for frames with higher resolution than other models, reaching 94.8% and 94.4%, respectively, with a mDice score of 84.7%. Considering that the overall sensitivity and precision exceeded 90% and the overall mDice score exceeded 80%, it is evident that, despite cystoscopy being conducted in a liquid environment, the HRNetV2 model can achieve satisfactory detection performance similar to that of endoscopy and even surpass it in cases where visibility is high. However, there is still a considerable gap in detecting low-resolution frames compared with other published models. The diagnostic sensitivity and precision for low-resolution bladder lesions are 75.6% and 74.8%, respectively, with a mDice score of 56.6%. Therefore, maintaining good transparency of the bladder cavity fluid is essential to ensuring the effectiveness of deep learning.

We investigated the observational and operational factors influencing poor visibility, as it is the most common cause of abnormal detections. We prioritized missed detections above false positives, since missed detections can lead to more severe adverse consequences. Reduced visibility can occur when the lens is too close or far from the target, resulting in an out-of-focus image. Moving the lens can partially improve this. Another cause of undiagnosed lesions is the indistinct differentiation between bladder tumor features and normal mucosal tissue, especially for flat lesions. In these cases, the lack of multiangle observations makes it difficult for model identification. Further improvement requires additional training data with multiangle and dynamic observation views to improve performance. Morphological similarity to trained features is the leading cause of false positive detection. This highlights the need for more data to improve the segmentation effect of bladder lesions. A low GT ratio can lead to false negative and false positive detections, as indicated by our GT area ratio test. Real-time panoramic image stitching for cystoscopy may be a useful solution to this problem. Further investigation into image enhancement techniques, including noise filtering and image dehazing,²¹ could improve performance.

This study has some limitations. First, it is a single-center study with limited datasets, which affects the detection efficiency of the model. To address this, we plan to collect more data and conduct multicenter studies. Second, the imbalanced distribution of various pathological types within the datasets, particularly the lack of samples of CIS, prevented tumor classification. This highlights the need to augment the dataset with various pathology types. Third, although HRNetV2 shows satisfactory results, its performance compared with other models may not be optimal. Future research

should focus on improving algorithms based on HRNetV2 for better bladder tumor detection.

CONCLUSIONS

Our study highlights the potential of integrating the HRNetV2 algorithm for real-time intraoperative detection of bladder lesions. This method shows promise in improving the accuracy and efficiency of bladder lesion diagnosis in clinical practice. Continued refinement and optimization of our method will provide valuable insights for future research in the field of medical image processing and lay the foundation for subsequent analyses, including pathology classification.

SUPPLEMENTARY INFORMATION The online version contains supplementary material available at <https://doi.org/10.1245/s10434-025-17015-3>.

ACKNOWLEDGMENT Not applicable.

AUTHOR CONTRIBUTIONS Z.J. and Z.Y. contributed to conception and design. Y.L. performed the data analysis and contributed to the drafting of the manuscript. Y.S. contributed to visualization. C.H. and G.H. contributed to data interpretation. Z.Y. and Z.J. contributed to data processing. All authors contributed to the article and approved the submitted version.

FUNDING This work was supported by the National High Level Hospital Clinical Research Funding (2022-PUMCH-B-008), CAMS Innovation Fund for Medical Sciences (CIFMS, 2024-I2M-C&T-B-020) and China Primary Health Care Foundation (012).

DATA AVAILABILITY The data that support the findings of this study are available from the corresponding author, Z.J., upon reasonable request.

DISCLOSURES The authors declare that they have no competing interests

ETHICS APPROVAL AND CONSENT TO PARTICIPATE This study was approved by the Ethics Committee of Peking Union Medical Hospital, Chinese Academy of Medical Sciences (I-23PJ1996).

CONSENT FOR PUBLICATION Written informed consent was obtained from the patient for publication and accompanying images.

OPEN ACCESS This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

1. Lobo N, Afferi L, Moschini M, Mostafid H, Porten S, Psutka SP, et al. Epidemiology, screening, and prevention of bladder cancer. *Eur Urol Oncol*. 2022;5(6):628–39.
2. Mulawkar PM, Sharma G, Tamhankar A, Shah U, Raheem R. Role of macroscopic image enhancement in diagnosis of non-muscle-invasive bladder cancer: an analytical review. *Front Surg*. 2022;9:762027.
3. Kamat AM, Hahn NM, Efstathiou JA, Lerner SP, Malmström PU, Choi W, et al. Bladder cancer. *Lancet*. 2016;388(10061):2796–810.
4. Zhu CZ, Ting HN, Ng KH, Ong TA. A review on the accuracy of bladder cancer detection methods. *J Cancer*. 2019;10(17):4038–44.
5. Daneshmand S, Patel S, Lotan Y, Pohar K, Trabulsi E, Woods M, et al. Efficacy and safety of blue light flexible cystoscopy with hexaminolevulinate in the surveillance of bladder cancer: a phase iii, comparative, multicenter study. *J Urol*. 2018;199(5):1158–65.
6. Houwen B, Vleugels JLA, Pellisé M, Rivero-Sánchez L, Balaguer F, Bisschops R, et al. Real-time diagnostic accuracy of blue light imaging, linked color imaging and white-light endoscopy for colorectal polyp characterization. *Endosc Int Open*. 2022;10(1):E9–e18.
7. Ali N, Bolenz C, Todenhöfer T, Stenzel A, Deetmar P, Kriegmair M, et al. Deep learning-based classification of blue light cystoscopy imaging during transurethral resection of bladder tumors. *Sci Rep*. 2021;11(1):11629.
8. Winkler JK, Fink C, Toberer F, Enk A, Deinlein T, Hofmann-Wellenhof R, et al. Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA Dermatol*. 2019;155(10):1135–41.
9. Guo Y, Li C, Zhang S, Zhu G, Sun L, Jin T, et al. U-net-based assistive identification of bladder cancer: a promising approach for improved diagnosis. *Urol Int*. 2024;108(2):100–7.
10. Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQ, et al. Augmented bladder tumor detection using deep learning. *Eur Urol*. 2019;76(6):714–8.
11. Ikeda A, Nosato H. Overview of current applications and trends in artificial intelligence for cystoscopy and transurethral resection of bladder tumours. *Curr Opin Urol*. 2024;34(1):27–31.
12. Iwaki T, Akiyama Y, Nosato H, Kinjo M, Niimi A, Taguchi S, et al. Deep learning models for cystoscopic recognition of hunner lesion in interstitial cystitis. *Eur Urol Open Sci*. 2023;49:44–50.
13. Hui X, Yuhua Z, Tong Z, Zhihui L. Survey of image semantic segmentation methods based on deep neural network. *J Front Comp Sci Tech*. 2021;15(1):47.
14. Wang J, Sun K, Cheng T, Jiang B, Deng C, Zhao Y, et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans Pattern Anal Mach Intell*. 2021;43(10):3349–64.
15. Ma K, Hao M, Shang W, Liu J, Meng J, Hu Q, et al. Study on the influence of label image accuracy on the performance of concrete crack segmentation network models. *Sensors (Basel)*. 2024;24(4):1068.
16. Abdulwahab S, Rashwan HA, Sharaf N, Khalid S, Puig D. Deep monocular depth estimation based on content and contextual features. *Sensors*. 2023;23(6):2919.
17. Sun K, Zhao Y, Jiang B, Cheng T, Xiao B, Liu D, et al. High-resolution representations for labeling pixels and regions. arXiv preprint <https://arxiv.org/abs/1904.04514>. 2019.
18. Sengupta K, Srivastava PR. HRNET: AI-on-edge for mask detection and social distancing calculation. *SN Comput Sci*. 2022;3(2):157.
19. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(4):640–51.
20. Freitas NR, Vieira PM, Cordeiro A, Tinoco C, Morais N, Torres J, et al. Detection of bladder cancer with feature fusion, transfer learning and CapsNets. *Artif Intell Med*. 2022;126:102275.
21. Jia X, Shkolyar E, Laurie MA, Eminaga O, Liao JC, Xing L. Tumor detection under cystoscopy with transformer-augmented deep learning algorithm. *Phys Med Biol*. 2023;68(16):165013.
22. Varnyú D, Szirmay-Kalos L. A comparative study of deep neural networks for real-time semantic segmentation during the transurethral resection of bladder tumors. *Diagnostics*. 2022;12(11):2849.
23. Ikeda A, Nosato H, Kochi Y, Kojima T, Kawai K, Sakanashi H, et al. Support system of cystoscopic diagnosis for bladder cancer based on artificial intelligence. *J Endourol*. 2020;34(3):352–8.
24. Zhang Q, Liang Y, Zhang Y, Tao Z, Li R, Bi H. A comparative study of attention mechanism based deep learning methods for bladder tumor segmentation. *Int J Med Inform*. 2023;171:104984.
25. Wu S, Chen X, Pan J, Dong W, Diao X, Zhang R, et al. An artificial intelligence system for the detection of bladder cancer via cystoscopy: a multicenter diagnostic study. *J Natl Cancer Inst*. 2022;114(2):220–7.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.