



## OPEN ACCESS

## EDITED BY

Hemalatha K.L.,  
Sri Krishna Institute of Technology, India

## REVIEWED BY

Ananda Babu Jayachandra,  
Malnad College of Engineering, India  
Yi Ding,  
Second Affiliated Hospital of Jilin University,  
China  
Abdul Lateef Haroon P. S.,  
Ballari Institute of Technology and  
Management, India

## \*CORRESPONDENCE

Lin Zhang

✉ Lin.zhang2@monash.edu

Guoliu Ye

✉ 012005013@bbmu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work

RECEIVED 08 May 2025

ACCEPTED 13 August 2025

PUBLISHED 12 September 2025

## CITATION

Li B, Chen L, Sun C, Wang J, Ma S, Xu H,  
Wang L, Rong T, Hu Q, Wei J, Lu L, Bai G,  
Liu Z, Luo P, Xu A, Liu L, Ye G and Zhang L  
(2025) Leveraging deep learning  
for early detection of cervical cancer  
and dysplasia in China using  
U-NET++ and RepVGG networks.  
*Front. Oncol.* 15:1624111.  
doi: 10.3389/fonc.2025.1624111

## COPYRIGHT

© 2025 Li, Chen, Sun, Wang, Ma, Xu, Wang,  
Rong, Hu, Wei, Lu, Bai, Liu, Luo, Xu, Liu, Ye and  
Zhang. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Leveraging deep learning for early detection of cervical cancer and dysplasia in China using U-NET++ and RepVGG networks

Baoqing Li<sup>1†</sup>, Lulu Chen<sup>2†</sup>, Chudi Sun<sup>3†</sup>, Jian Wang<sup>4,5</sup>,  
Sicong Ma<sup>4,5</sup>, Hang Xu<sup>6</sup>, Luyao Wang<sup>7</sup>, Taotao Rong<sup>7</sup>, Qun Hu<sup>7</sup>,  
Jie Wei<sup>7</sup>, Lijuan Lu<sup>8</sup>, Guannan Bai<sup>9</sup>, Zhangdaihong Liu<sup>10,11</sup>,  
Peng Luo<sup>12</sup>, Aimin Xu<sup>13</sup>, Li Liu<sup>14</sup>, Guoliu Ye<sup>7\*</sup> and Lin Zhang<sup>4,5\*</sup>

<sup>1</sup>Department of Laboratory Medicine, The Second Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, <sup>2</sup>Department of Gynecology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, <sup>3</sup>School of Computer Science and Technology, Soochow University, Suzhou, China, <sup>4</sup>School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia, <sup>5</sup>Suzhou Industrial Park Monash Research Institute of Science and Technology, Suzhou, Jiangsu, China, <sup>6</sup>School of Computer Science, The University of Sydney, Sydney, NSW, Australia, <sup>7</sup>Department of Gynecology, The First Affiliated Hospital of Bengbu Medical University, Bengbu, Anhui, China, <sup>8</sup>Comprehensive Cancer Prevention and Treatment Center, Nantong University Affiliated Jiangyin Hospital, Jiangyin, China, <sup>9</sup>Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, Zhejiang, China, <sup>10</sup>Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom, <sup>11</sup>Oxford-Suzhou Centre for Advanced Research, Suzhou, Jiangsu, China, <sup>12</sup>Department of Oncology, Zhujiang Hospital, Southern Medical University, Guangzhou, China, <sup>13</sup>College of Big Data and Software Engineering, Zhejiang Wanli University, Ningbo, China, <sup>14</sup>Data Center, Affiliated Hospital of Jiangnan University, Wuxi, China

**Background:** Cervical cancer is a significant global public health issue, primarily caused by persistent high-risk human papillomavirus (HPV) infections. The disease burden is disproportionately higher in low- and middle-income regions, such as rural China, where limited access to screening and vaccinations leads to increased incidence and mortality rates. Cervical cancer is preventable and treatable when detected early; this study utilizes deep learning to enhance early detection by improving the diagnostic accuracy of colposcopic image analysis.

**Objective:** The aim of this study is to leverage deep learning techniques to improve the early detection of cervical cancer through the enhancement of colposcopic image diagnostic accuracy.

**Methods:** The study sourced a comprehensive dataset of colposcopic images from The First Affiliated Hospital of Bengbu Medical University, with each image manually annotated by expert clinicians. The U-NET++ architecture was employed for precise image segmentation, converting colposcopic images into binary representations for detailed analysis. The RepVGG framework was then applied for classification, focusing on detecting cervical cancer, HPV infections, and cervical intraepithelial neoplasia (CIN). From a dataset of 848 subjects, 424 high-quality images were selected for training, with the remaining 424 used for validation.

**Results:** The deep learning model effectively identified the disease severity in colposcopic images, achieving a predictive accuracy of 83.01%. Among the 424 validation subjects, cervical pathology was correctly identified in 352, demonstrating high diagnostic precision. The model excelled in detecting early-stage lesions, including CIN I and CIN II, which are crucial for initiating

timely interventions. This capability positions the model as a valuable tool for reducing cervical cancer incidence and improving patient outcomes.

**Conclusion:** The integration of deep learning into colposcopic image analysis marks a significant advancement in early cervical cancer detection. The study suggests that AI-driven diagnostic tools can significantly improve screening accuracy. Reducing reliance on human interpretation minimizes variability and enhances efficiency. In rural and underserved areas, the deployment of AI-based solutions could be transformative, potentially reducing cervical cancer incidence and mortality. With further refinement, these models could be adapted for broader population screening, aiding global efforts to eliminate cervical cancer as a public health threat.

#### KEYWORDS

cervical cancer screening, public health strategy, deep learning, colposcopy, early diagnosis, resource-limited settings

## Introduction

Cervical cancer remains a major global health challenge, with over half a million new cases diagnosed annually and significant mortality, especially in low- and middle-income countries (1). The main methods of cervical cancer screening include human papillomavirus test (HPV test) and liquid based cytology test (LCT). LCT image can accurately and intuitively show the morphological characteristics of cervical cells and is also an accurate method to judge the precancerous lesions of cervical cancer (2). Despite the availability of various preventive measures and screening techniques, such as human papillomavirus (HPV) vaccinations and routine Pap smear tests (3, 4), manual screening processes are not always highly accurate (5), potentially leading to delayed diagnoses of related pathological changes (6). Colposcopy is an important diagnostic tool in women's health care and plays an important role, especially in the early detection and prevention of cervical cancer. A low-power microscope that uses acetic acid and Lugol solution to carefully examine the cervix, vagina, and vulva and provide a zoomed view of these areas, enabling healthcare professionals to identify and assess abnormalities that may not be visible to the naked eye. In a large trial of low-grade abnormalities, the initial colposcopy was only 53% sensitive to detecting high-grade disease within the following 2 years. Studies have shown low agreement between colposcopic impressions of disease and final histology. The use of multiple biopsies can improve the accuracy of colposcopy diagnosis. Although colposcopy is practiced by many clinicians (including senior practitioners, gynecologists, gynecologic oncologists, etc.), the standardization of the procedure, the necessary training, and the ongoing development and maintenance of colposcopy skills are generally poor. There is also ample evidence that colposcopy has significant executive-to-practitioner variability and poor reliability. In addition, a lot of complicated colposcopy and image reading increase the workload of clinicians. Furthermore, in the context of uneven

development, many countries and regions face challenges in training or retaining specialized colposcopists and cervical cytopathologists (7). Consequently, the development of a more accurate and cost-effective method for cervical cancer screening has emerged as a primary challenge for the early detection of the disease.

Artificial intelligence (AI), particularly deep learning (DL), has emerged as a powerful tool to address these challenges. Deep learning models excel at analyzing large datasets of medical imagery and offer high precision in detecting pre-cancerous conditions, potentially reducing the need for human expertise and improving diagnostic consistency (2, 8–10). Recent studies have demonstrated the effectiveness of AI models, such as transfer learning for cervical cell image classification and U-net-based models for cancer screening using medical images (11–16). These advancements hold promise for reducing the diagnostic burden in both high- and low-resource settings.

In the diagnosis and treatment of gynecological malignant tumors, the imaging information of the lesion site can provide an important diagnostic basis for clinicians. With the popularization and continuous development of imaging technology, a large number of medical images need to be managed reasonably and efficiently. The novelty of this study lies in its focus on colposcopic image analysis using deep learning, rather than relying solely on traditional cytological or histological methods (7). While AI has been applied to cervical cancer diagnosis before, most efforts have centered on cytological screening. Our work, by contrast, leverages deep learning algorithms to automatically interpret colposcopic images—a more direct, non-invasive screening tool that can be particularly effective in regions with limited access to trained cytologists or pathologists. This approach introduces an innovative pathway for early detection, offering a viable solution that can be more easily integrated into healthcare systems with constrained resources.

Furthermore, our contributions include the development and validation of a deep learning model that combines advanced image segmentation techniques and transfer learning strategies. These enhancements allow for more accurate classification of cervical abnormalities and can significantly reduce human error in the screening process. The model was validated on multiple independent datasets, ensuring robustness and generalizability across diverse clinical environments. By focusing on improving diagnostic accuracy in low-resource settings, our system offers a practical, cost-effective solution that addresses the global disparity in cervical cancer outcomes.

This research aligns with global health initiatives aimed at reducing cancer-related mortality by expanding access to cutting-edge diagnostic tools, especially in underserved populations. The use of AI in colposcopy represents a promising step forward in the effort to democratize healthcare and deliver more equitable outcomes worldwide.

## Methods

### Data preparation

Our dataset, collected from the Department of Gynecology, the First Affiliated Hospital of Bengbu Medical University, Anhui, PR China, under ethical approval (Approval Number: the First Affiliated Hospital of Bengbu Medical University 2024 【427】), consists of 7,612 colposcopic images from 848 distinct female subjects, with each subject contributing multiple images. These images represent 27 different pathological conditions, and it is common for a single colposcopic image to be associated with multiple diseases. To streamline classification, these conditions have been grouped into two primary categories: “mildly normal” and “severe”. The “mildly normal” category includes conditions that are asymptomatic or do not require further medical attention,

while the “severe” category encompasses conditions that necessitate additional treatment. Accurately distinguishing between these two categories can play a crucial role in early screening, helping patients promptly identify the need for further medical care, thereby improving overall health management. An example of this is shown in Figure 1.

In colposcopic image analysis, it's common to find that images may contain limited feature information from the regions of interest, leading to potential bias toward non-critical areas during the model training process. This can result in the model learning less relevant information. To address this, we employed a sampling methodology that leverages closed curves to segment areas of interest based on their morphological characteristics. As depicted in Figure 2, red closed curves are used to highlight potential disease regions.

A total of 51 colposcopic images underwent rigorous screening and pixel-level annotation by two experienced colposcopists. While formal inter-rater statistics were not calculated, annotations were completed in consensus based on clinical diagnostic criteria. These annotated images were used solely for segmentation model training.

Of these, 235 images were classified as indicative of severe conditions, and 5 images were determined to be clearly normal. Additionally, through an image screening process that aligns each image with its respective disease type, 184 colposcopic images were identified as having features that, while not overtly abnormal, were confirmed as normal after examination. This brings the total number of images classified as normal and non-diseased to 189, resulting in a class ratio of approximately 1:1.24 (severe:normal). This stratified sampling technique helps to address data imbalance, improving model training in subsequent phases.

Note: The 51 annotated images refer to a segmentation subset used for U-Net++ training. In contrast, the classification dataset (424 training + 424 validation images) was constructed separately through stratified sampling of diagnostically validated images from the full set of 7,612.



mildly normal

severe

FIGURE 1  
Sample images.



FIGURE 2  
Example of drawing segmentation regions in the image.

## Process and methods

This study establishes a model framework for early cervical cancer screening using convolutional neural networks (CNNs). The operational workflow of the model is illustrated in Figure 3 and comprises the following stages:

1. Pre-training and Segmentation: To address the limited availability of segmentation annotation data in our dataset, we utilized the publicly available AnnoCerv dataset (17) for pre-training on a U-Net++ model (18). This pre-trained network was subsequently fine-tuned with our dataset, enabling the model to achieve faster convergence.
2. Feature-based Classification: For the regions of interest identified by the segmentation model, we employed the RepVGG model (19) to perform classification training. This allowed the model to focus on disease-relevant features.
3. Incorporating Global Information: Since the classification was initially based solely on feature regions, the model could develop biases. To mitigate this, we incorporated global context by including the original, unsegmented images along with the segmented feature-only images in the training process. This ensured that the model retained both local and global information.
4. Final Model Output: The final model uses a voting mechanism that considers predictions from both the

original and segmented images across multiple iterations. The category with the highest voting weight is selected as the final model output, ensuring robust and reliable classification results.

## Methodology

Initially, we employed the U-Net++ architecture to perform segmentation on our dataset. The segmentation quality was quantified using the Intersection over Union (IoU), defined as:  $IoU = \frac{A \cap B}{A \cup B}$ , where  $A \cap B$  represents the number of pixels in the intersection of the predicted segmentation result A and the ground truth B, and  $A \cup B$  denotes the number of pixels in the union of A and B.

When trained directly on our entire dataset, the network's performance yielded an IoU range of 0.05 to 0.60 across 51 segmented images, reflecting suboptimal segmentation quality. To address this, we utilized the publicly available AnnoCerv dataset for pre-training, followed by fine-tuning on our specific dataset. This enhanced the model's performance, with an improved IoU range of 0.27 to 0.71 for the same 51 segmented images, a significant improvement over the direct training approach. As shown in Figure 4, we compare the IOU values of each picture. Our observations more intuitively reveal a phenomenon: in the majority of images, the segmentation performance of models with pre-training is only slightly better than that of models without pre-training. However, in more than ten images, the segmentation performance of the pre-trained models significantly surpasses that of the non-pre-trained methods. This observation, to a certain extent, validates the effectiveness of the strategies we have employed.

The segmented images primarily captured regions associated with potential pathological features. These segmented images were then employed for classification training using the RepVGG network. The classification training on our dataset, consisting of 424 images, achieved an accuracy of 84.19%. In the remaining set of validation images, there are 419 instances classified as 'severe'. In contrast, only a mere five images are categorized as 'mildly normal.' The likely reason for this disparity within the dataset is that the majority of individuals undergoing colposcopic examinations are presumed to be symptomatic, suggesting a potential diagnosis. Consequently, the scarcity of 'mildly normal' images may be

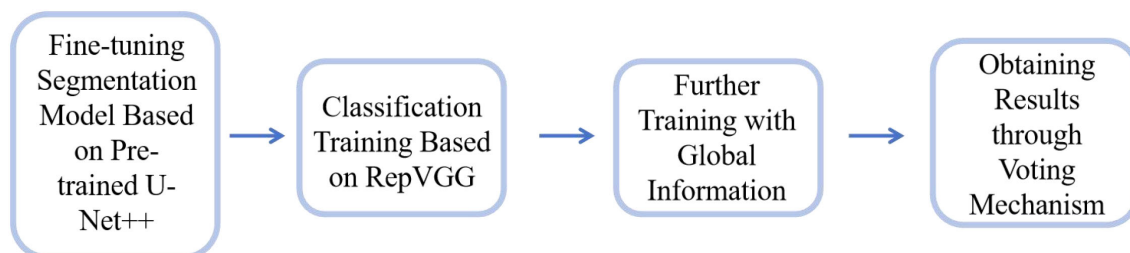


FIGURE 3  
Model framework workflow.



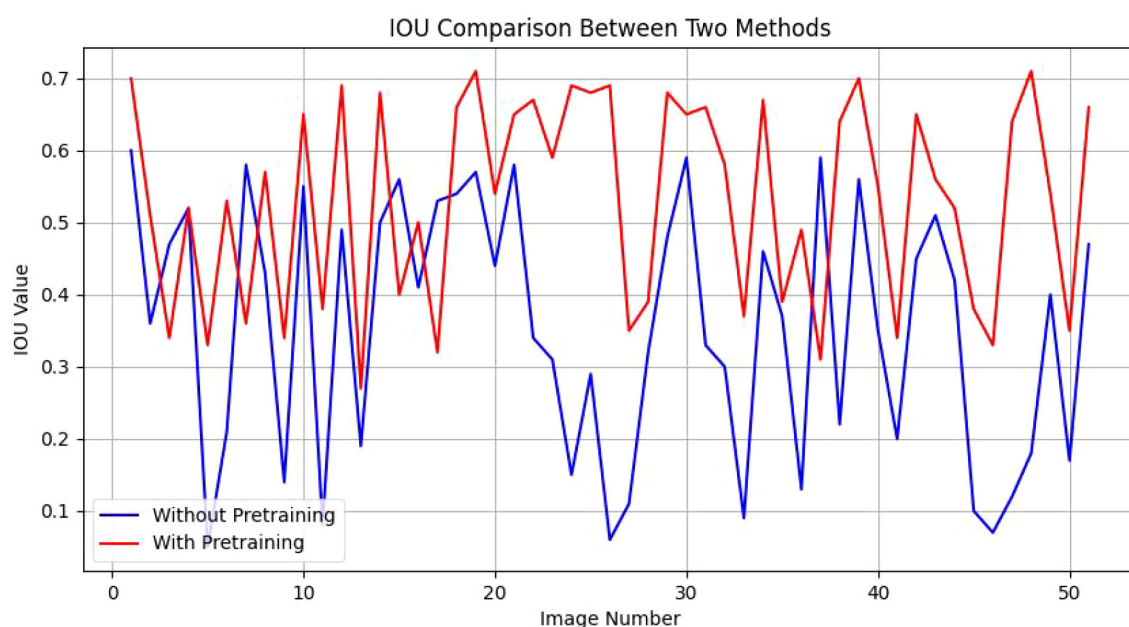


FIGURE 4  
IOU comparison between two methods.

attributed to this bias toward symptomatic patients. When applied to disease assessment of 424 women in validation subset, the model's accuracy was 74.05%, which fell short of the threshold required for clinical applicability. This suggests that models trained solely on regions of interest may miss critical global information in unseen test data.

To address this limitation, we integrated the original images with their corresponding segmented regions for joint training, improving the model's training accuracy to 88.44%. This approach also enhanced the disease classification accuracy for the 424 women in validation subset to 81.83%.

Additionally, we introduced a voting mechanism to further refine the model's output. For each patient, an original image (Image 1) is passed through the segmented model to isolate the region of interest (Image 2). Both Image 1 and Image 2 are then flipped horizontally, producing Images 3 and 4, respectively. To account for varying real-world lighting conditions, the luminance of these images is adjusted, resulting in Images 5 and 6. The voting process, based on the weights assigned to each image (as detailed in Table 1), determines the final classification. An image is classified as belonging to a particular category if the voting score for that category exceeds 0.5. For patients with multiple images, each original image is assigned equal weight, and the final disease status is determined by the collective results of all images.

This voting mechanism increased the overall accuracy in validating the disease status of all women to 83.01%, demonstrating a significant improvement in the model's diagnostic performance.

The voting weights assigned to each image variant (e.g., flipped, brightness-adjusted) were determined empirically through iterative tuning. While not obtained via formal grid search, these weights

reflect their relative contributions to improved ensemble classification accuracy on the validation set.

## Results

To demonstrate the overall efficacy of our algorithmic framework, we conducted comparative experiments with several state-of-the-art methods and performed ablation studies on our approach. Since our ultimate goal is to achieve a classification result, the neural network methods we compared include the conventional models such as VGG-16, ResNet-50, and Densenet-121. All models were trained with consistent hyperparameters using PyTorch 2.1.0 and torchvision v0.15.0 on an NVIDIA RTX 4070 Ti SUPER GPU (16 GB VRAM), with 80 training epochs (no early stopping), batch size 8, Adam optimizer (learning rate = 0.001), and categorical cross-entropy loss.

Our ablation experiments investigated several configurations: models that do not employ pre-trained segmentation procedures, models that entirely omit the segmentation process, and models that utilize segmentation but forgo the voting mechanism. Notably, the model excluding both the segmentation process, and the voting mechanism corresponds to the original RepVGG model. The results of these experiments are summarized in Table 2 below.

To more clearly present our data, we have plotted a bar chart as shown in Figure 5 below:

An analysis of the comparative results reveals that the impact of various classical classification network models on the final image classification outcome is generally minimal. However, this effect becomes more pronounced when using a limited amount of

TABLE 1 Corresponding weight information of images.

| Figure | 1   | 2    | 3    | 4   | 5    | 6    |
|--------|-----|------|------|-----|------|------|
| Weight | 0.2 | 0.25 | 0.15 | 0.2 | 0.11 | 0.09 |

segmented image data, particularly if no additional data is used for pre-training. Interestingly, the model that excludes additional data for segmentation sometimes yields results that are slightly inferior to those of the original classification network. In contrast, the extraction of key information areas through the segmentation operation prior to classification significantly enhances the model’s performance. Consequently, our experimental design methodology demonstrates a substantial degree of effectiveness and robustness.

In fact, in the classification of images, the most intuitive way to reflect the effectiveness of the model is to verify the accuracy of the set, and at the same time, we draw its confusion matrix, as shown in Figure 6. Although the number of images classified as “mildly normal” is relatively small, making it difficult to accurately assess the model’s recognition rate for this category from the experiments, we can observe from the figures that our model’s misclassification of “severe” images is significantly lower than that of other models. In the context of medical research, the tolerance for misclassifying disease-free images as diseased is considerably higher than that for misclassifying diseased images as disease-free, as the consequences of overlooking a disease can be far more severe than the inconvenience caused by false positives. This is particularly crucial in diagnostic settings where early and accurate detection of diseases can significantly impact patient outcomes. From this standpoint, our model’s ability to minimize the misclassification of diseased images as disease-free demonstrates its vastly superior performance compared to other models, which is a critical advantage in clinical applications.

Discussion

Cervical cancer remains the leading cause of cancer-related death among women in developing countries. However, the disease is highly curable when detected early through effective screening. In this study, we developed a novel, automated cervical cancer

screening system based on deep learning, aiming to assist clinicians in identifying early signs of precancerous lesions and cancer via colposcopy image analysis. By integrating advanced image processing and machine learning techniques, the system demonstrates the potential to improve diagnostic accuracy, reduce clinician workload, and minimize misdiagnosis, particularly in resource-limited settings where trained specialists may be scarce.

The proposed system incorporates segmentation-guided classification, enabling it to extract and analyze key regions of interest in colposcopic images. Our experimental results show that incorporating global information, using pre-trained models for segmentation, and applying a voting mechanism significantly enhance the model’s performance. Specifically, the final model achieved an overall accuracy of 83.01% in a test cohort of 848 women, supporting the robustness and effectiveness of the design. Moreover, we found that accurate localization and segmentation of diagnostic regions had a direct and substantial impact on downstream classification performance, emphasizing the necessity of precise image preprocessing in automated screening workflows.

Despite the encouraging results, we acknowledge several limitations that warrant further investigation. First, although the study involved multi-institutional collaboration, the dataset was collected from a single medical center. This introduces potential bias related to device-specific settings and clinical protocols, which may affect the model’s generalizability. To address this, we are actively working with additional clinical sites to build a multi-center dataset for comprehensive external validation.

Second, the validation cohort had a limited number of “mildly normal” cases, reflecting the symptomatic referral nature of the data source. We recognize that this may influence specificity and are currently enriching the dataset to better reflect the full spectrum of screening populations.

Third, while the segmentation module plays a critical role in model performance, only 51 images were pixel-wise annotated for quantitative evaluation. Although this subset enabled proof-of-concept validation, we plan to expand the number of annotated samples to allow more statistically robust assessments.

Additionally, the voting weights used in the ensemble process were selected empirically. Future work will explore formal optimization strategies, such as grid search or Bayesian optimization, to determine optimal fusion parameters.

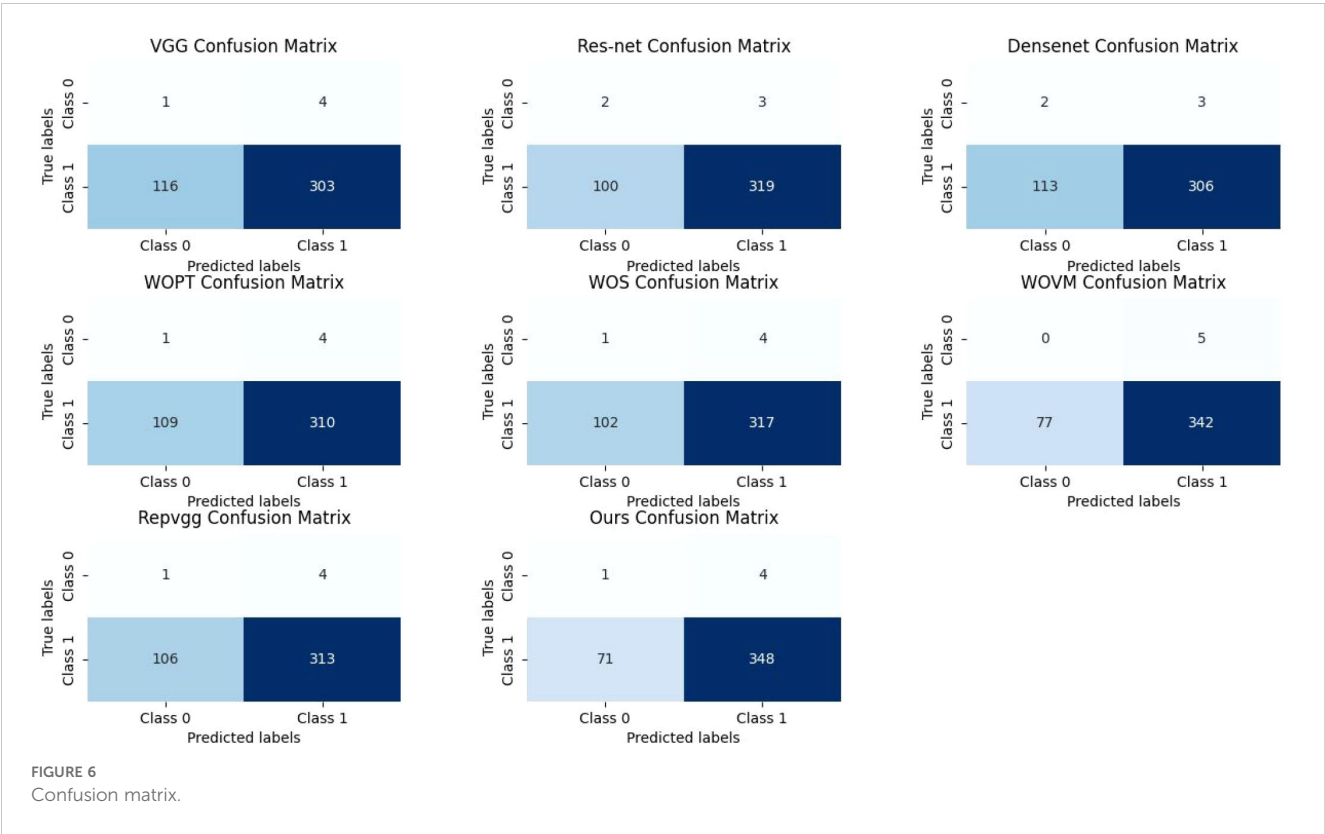
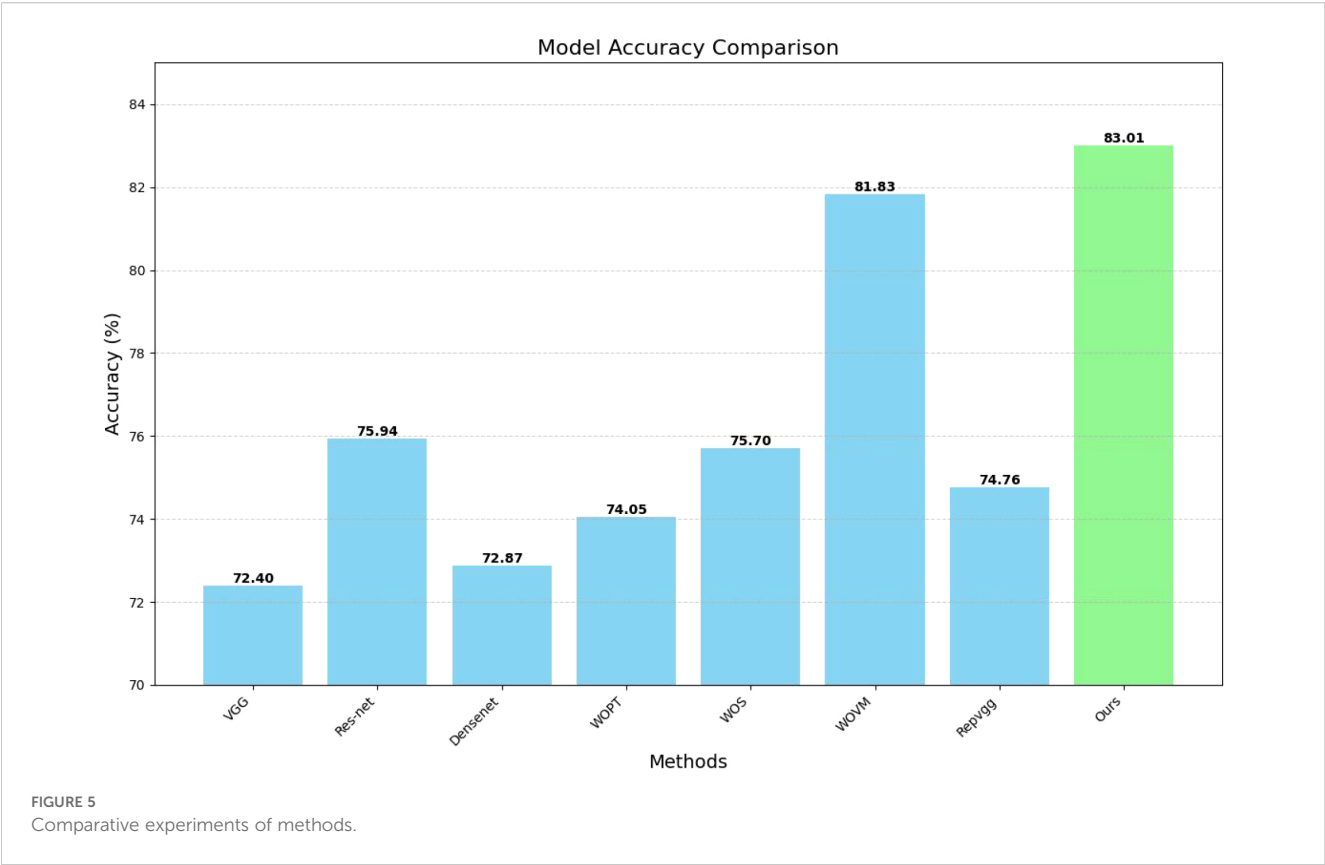
Regarding evaluation metrics, although we primarily report classification accuracy—a widely accepted benchmark in deep learning classification tasks—we acknowledge that additional diagnostic metrics such as ROC curves, AUC, sensitivity, specificity, and precision-recall curves provide complementary insights. These will be reported in follow-up studies to provide a more comprehensive assessment of diagnostic performance.

Furthermore, although we explored visual interpretability tools (e.g., Grad-CAM) internally to assess model decision-making transparency, these results were excluded due to space limitations and will be detailed in subsequent publications.

Finally, while recent AI-based cervical screening tools provide important contributions to the field, they differ in screening modality (e.g., cytology, HPV biomarkers) and diagnostic focus. Our work,

TABLE 2 Comparative experiments.

| Methods                         | Accuracy (%) |
|---------------------------------|--------------|
| VGG                             | 72.40        |
| Res-net                         | 75.94        |
| Densenet                        | 72.87        |
| Without pre-trained (WOPT)      | 74.05        |
| Without segmentation (WOS)      | 75.70        |
| Without Voting Mechanism (WOVM) | 81.83        |
| Repvgg                          | 74.76        |
| Ours                            | 83.01        |



which centers on colposcopic image analysis, offers a complementary approach that can be integrated into broader screening pathways.

In summary, this study presents a practical, scalable, and interpretable AI-assisted cervical cancer screening system that demonstrates strong diagnostic performance and promising real-world applicability. With continued refinement—particularly in terms of data diversity, interpretability, and evaluation metrics—this framework may help expand access to high-quality screening, especially in under-resourced regions, and ultimately contribute to the global effort to reduce cervical cancer incidence and mortality.

## Conclusion

This study presents a novel and effective automated screening system for cervical cancer based on deep learning, achieving an accuracy of 83.01% in testing. By leveraging advanced image processing and machine learning techniques, the system enhances screening specificity and accuracy, reduces clinician workload, and mitigates subjective biases. It holds promise for resource-limited settings, significantly improving early diagnosis and reducing cervical cancer incidence. Despite its success, challenges such as handling blurry images and ensuring system robustness remain. Future work should focus on larger datasets and more advanced model architectures to further enhance performance and accessibility globally.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## Ethics statement

The studies involving humans were approved by the First Affiliated Hospital of Bengbu Medical University, Anhui, China. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

BL: Writing – review & editing, Investigation. LC: Investigation, Writing – review & editing. CS: Writing – original draft, Methodology, Investigation. JianW: Writing – review & editing. SM: Writing – review

& editing. HX: Writing – review & editing. LW: Writing – review & editing. TR: Writing – review & editing. QH: Writing – review & editing. JieW: Writing – review & editing. LJL: Writing – review & editing. GB: Writing – review & editing. ZL: Writing – review & editing. PL: Writing – review & editing. AX: Writing – review & editing. LL: Writing – review & editing. GY: Writing – review & editing. LZ: Writing – review & editing, Investigation, Methodology, Writing – original draft.

## Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study is supported by the Zhejiang Provincial Natural Science Foundation of China under Grant No. LTGY23H240002, Zhejiang Provincial Natural Science Foundation of China under Grant No. LTGY23H240002, Wuxi Municipal Health Commission Research Project Plan under Grant No. Z202309, Taihu Light Technology Research Project (Medical and Health) Y20232001, Natural Science Research Project of Anhui Educational Committee 2022AH051428.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



## References

1. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol*. (2006) 24:2137–50. doi: 10.1200/JCO.2005.05.2308
2. Wang J, Yu Y, Tan X, Wan H, Zheng N, He Z, et al. Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. *Nat Commun*. (2024) 15:4369. doi: 10.1038/s41467-024-48705-3
3. Sawaya GF, Smith-McCune K, Kuppermann M. Cervical cancer screening: more choices in 2019. *JAMA*. (2019) 321:2018–9. doi: 10.1001/jama.2019.4595
4. Wentzensen N, Lahrmann B, Clarke MA, Kinney W, Tokugawa D, Poitras N, et al. Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical cancer screening. *J Natl Cancer Inst*. (2021) 113:72–9. doi: 10.1093/jnci/djaa066
5. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. *Lancet*. (2019) 393:169–82. doi: 10.1016/S0140-6736(18)32470-X
6. Pollack AE, Tsu VD. Preventing cervical cancer in low-resource settings: building a case for the possible. *Int J Gynaecol Obstet*. (2005) 89 Suppl 2:S1–3. doi: 10.1016/j.ijgo.2005.01.014
7. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Matchar DB, et al. Accuracy of the papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: A systematic review. *Ann Intern Med*. (2000) 132:810–9. doi: 10.7326/0003-4819-132-10-200005160-00009
8. Wentzensen N, Massad LS, Mayeaux EJ Jr, Khan MJ, Waxman AG, Einstein MH, et al. Evidence-based consensus recommendations for colposcopy practice for cervical cancer prevention in the United States. *J Low Genit Tract Dis*. (2017) 21:216–22. doi: 10.1097/LGT.0000000000000322
9. Iannantuono GM, Bracken-Clarke D, Floudas CS, Roselli M, Gulley JL, Karzai F. Applications of large language models in cancer care: current evidence and future perspectives. *Front Oncol*. (2023) 13:1268915. doi: 10.3389/fonc.2023.1268915
10. Choi HS, Song JY, Shin KH, Chang JH, Jang BS. Developing prompts from large language model for extracting clinical information from pathology and ultrasound reports in breast cancer. *Radiat Oncol J*. (2023) 41:209–16. doi: 10.3857/roj.2023.00633
11. Majeed T, Aalam SW, Ahanger AB, Masoodi TA, Macha MA, Bhat AA, et al. Transfer Learning approach for Classification of Cervical Cancer based on Histopathological Images. In: *2023 3rd International Conference on Artificial Intelligence and Signal Processing (AISP)*. VIJAYAWADA, India: IEEE (2023). p. 1–5. doi: 10.1109/AISP57993.2023.10135048
12. Meenakshisundaram N, Ramkumar G. An Automated Identification of Cervical Cancer disease using Convolutional Neural Network Model. In: *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*. Chennai, India: IEEE (2023). p. 1–7. doi: 10.1109/ACCAI58221.2023.10200640
13. Mustafa WA, Noor MI, Alquran H, Ghani MM, Hanafi HF, Lah NHC, et al. Stages classification on cervical cell images: A comparative study. In: *2023 3rd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*. Tumkur, India: IEEE (2023). p. 1–5. doi: 10.1109/ICMNWC60182.2023.10435963
14. Safitri PH, Fatichah C, Zulfa N. Two-stage classification of pap-smear images based on deep learning. In: *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*. Yogyakarta, Indonesia: IEEE (2022). p. 1–6. doi: 10.1109/ICITISEE57756.2022.10057846
15. Jin J, Zhu H, Teng Y, Ai Y, Xie C, Jin X. The accuracy and radiomics feature effects of multiple U-net-based automatic segmentation models for transvaginal ultrasound images of cervical cancer. *J Digit Imaging*. (2022) 35:983–92. doi: 10.1007/s10278-022-00620-z
16. Gautam S, Osman AFI, Richeson D, Gholami S, Manandhar B, Alam S, et al. Attention 3D U-NET for dose distribution prediction of high-dose-rate brachytherapy of cervical cancer: Direction modulated brachytherapy tandem applicator. *Med Phys*. (2024) 51:5593–603. doi: 10.1002/mp.17238
17. Minciună D, Socolov DG, Szocs A, Ivanov D, Giscu T, Nechifor V, et al. AnnoCerv: A new dataset for feature-driven and image-based automated colposcopy analysis. *Acta Universitatis Sapientiae Informatica*. (2023) 15:306–29. doi: 10.2478/ausi-2023-0019
18. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A nested U-net architecture for medical Image segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (2018). Cham, Switzerland: Springer. doi: 10.1007/978-3-030-00889-5\_1
19. Ding X, Zhang X, Ma N, Han J, Ding G, Sun J, et al. RepVGG: Making VGG-style ConvNets great again. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. CVPR: New York. (2021). p. 13728–37.