

Learning from hybrid bright-dark field imaging for resolution-enhanced digital pathology

Runjie Huang, Runnan Zhang, Xiyu Chen, Yefeng Shu, Wenjian Chen, Kun Gui, Kehui Wang, Jiasong Sun*, Juergen W. Czarske* and Chao Zuo*

Citation: Huang RJ, Zhang RN, Chen XY et al. Learning from hybrid bright-dark field imaging for resolution-enhanced digital pathology. *Opto-Electron Adv* **9**, 260060 (2026).

<https://doi.org/10.29026/oea.2026.260060>

Received: 17 April 2026; Accepted: 6 July 2026; Published online: 28 August 2026

Related articles

Multi-prior physics-enhanced neural network enables pixel super-resolution and twin-image-free phase retrieval from single-shot hologram

Xuan Tian, Runze Li, Tong Peng, et al

Opto-Electronic Advances 2024, **7**(9): 240060 doi: [10.29026/oea.2024.240060](https://doi.org/10.29026/oea.2024.240060)

Dual-frequency angular-multiplexed fringe projection profilometry with deep learning: breaking hardware limits for ultra-high-speed 3D imaging

Wenwu Chen, Yifan Liu, Shijie Feng, et al

Opto-Electronic Advances 2025, **8**(9): 250021 doi: [10.29026/oea.2025.250021](https://doi.org/10.29026/oea.2025.250021)

Enhanced photoacoustic microscopy with physics-embedded degeneration learning

Haigang Ma, Shili Ren, Xiang Wei, et al

Opto-Electronic Advances 2025, **8**(3): 240189 doi: [10.29026/oea.2025.240189](https://doi.org/10.29026/oea.2025.240189)

Aberration-corrected differential phase contrast microscopy with annular illuminations

Yao Fan, Chenyue Zheng, Yefeng Shu, et al

Opto-Electronic Science 2025, **4**(8): 240037 doi: [10.29026/oes.2025.240037](https://doi.org/10.29026/oes.2025.240037)

More related articles in Opto-Electronic Journals Group website



Learning from hybrid bright-dark field imaging for resolution-enhanced digital pathology

Runjie Huang^{1,2,3,4†}, Runnan Zhang^{1,2,3,4†}, Xiyu Chen^{1,2,3,4†}, Yefeng Shu^{1,2,3,4}, Wenjian Chen³, Kun Gui⁵, Kehui Wang⁵, Jiasong Sun^{1,2,3,4*}, Juergen W. Czarske^{6,7*} and Chao Zuo^{1,2,3,4*}

Abstract: High-throughput whole-slide imaging (WSI) confronts a trade-off between scanning speed and optical resolution, rendering low-numerical-aperture objectives inadequate for resolving critical microstructures. Computational resolution enhancement offers an alternative; however, the ill-posedness of single-image resolution enhancement leads to over-smoothed reconstructions or hallucinations, undermining clinical reliability. We present HBDF-RE, a hybrid bright-field and dark-field resolution enhancement framework that exploits complementary physical contrast to constrain the reconstruction problem. A single dark-field acquisition, enriched with high-frequency scattering information inaccessible to conventional bright-field imaging, serves as a reference that reduces the solution space and guides reconstruction from low-resolution bright-field inputs. HBDF-RE integrates multimodal feature fusion, spatial attention, and joint spatial-frequency domain optimization to preserve global tissue morphology while recovering fine textural details. Experimental results demonstrate a 2.1× enhancement in resolution (from 950 nm to 445 nm), a 3.2 dB improvement in PSNR (Peak Signal-to-Noise Ratio), and an 84% reduction in reconstruction artifacts compared with representative single-image resolution enhancement methods. Importantly, in large-scale AI-assisted cervical cancer screening, HBDF-RE-reconstructed images yield an 11.14% improvement in diagnostic sensitivity, with pronounced gains in clinically ambiguous lesion categories. These results establish HBDF-RE as a physically grounded and clinically reliable resolution enhancement paradigm for digital pathology, offering a pathway toward next-generation WSI and computer-aided diagnosis.

Keywords: resolution enhancement; bright-dark field imaging; deep learning; digital pathology; whole-slide imaging

DOI: [10.29026/oea.2026.260060](https://doi.org/10.29026/oea.2026.260060) | CSTR: [32247.14.oea.2026.260060](https://cstr.net.cn/32247.14.oea.2026.260060)

Citation: Huang RJ, Zhang RN, Chen XY et al. Learning from hybrid bright-dark field imaging for resolution-enhanced digital pathology. *Opto-Electron Adv*, 260060 (2026).

1 Introduction

Whole-slide imaging (WSI), through multi-resolution digitization of entire tissue specimens, has substantially advanced the field of digital pathology¹. With the increasingly widespread adoption of WSI in clinical practice, the demand for high-resolution pathological imaging continues to grow, driven primarily by its critical role in disease diagnosis^{1,2},

tumor detection³⁻⁵, and pathological analysis^{6,7}. However, acquisition of high-resolution images typically relies on expensive technologies and hardware systems, posing challenges in meeting the escalating demands of pathologists in both clinical practice and research applications. Image super-resolution (SR) technology offers a cost-effective solution by employing computational imaging approaches⁷⁻¹⁰

Received: 17 April 2026

Accepted: 6 July 2026

Published online: 28 August 2026

¹Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; ²Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing 210019, China; ³Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense Nanjing University of Science and Technology, Nanjing 210094, China; ⁴State Key Laboratory of Extreme Environment Optoelectronic Dynamic Measurement Technology and Instrument, Taiyuan 030051, China; ⁵Konfoong Biotech International Co., LTD, Yuyao 315400, China; ⁶Competence Center for Biomedical Computational Laser Systems (BIOLAS), TU Dresden, Helmholtzstrasse 18, 01069 Dresden, Germany; ⁷Laboratory of Measurement and Sensor System Technique (MST), TU Dresden, Helmholtzstrasse 18, 01069 Dresden, Germany.

[†]These authors contributed equally to this work.

*Correspondence: JS Sun, E-mail: sunjiasong@njust.edu.cn; JW Czarske, E-mail: juergen.czarske@tu-dresden.de; C Zuo, E-mail: zuochao@njust.edu.cn

resolution beyond the conventional optical constraints¹⁰, reconstructing high-resolution (HR) images from readily accessible low-resolution (LR) images and achieving spatial resolution beyond the constraints imposed by objective numerical aperture¹⁰. In recent years, single-image super-resolution (SISR) technology has attracted considerable attention, particularly deep learning-based methods^{10–16}. Convolutional neural network-based SISR approaches have achieved substantial progress in this field^{17–21}; however, these methods typically produce over-smoothed reconstruction results, leading to blurred outputs. Generative adversarial network-based SISR methods^{22,23} have demonstrated breakthroughs in visual quality; nevertheless, they frequently generate artificial textures and unrealistic artifacts such as hallucinations. Due to insufficient prior knowledge, individual LR images often contain inadequate information, rendering deep learning reconstruction methods susceptible to severe artifacts including over-smoothing and ringing effects, thereby limiting reconstruction efficacy. These issues are particularly critical in pathological imaging applications, directly affecting accurate tissue structure identification and diagnostic reliability, and ultimately undermining the clinical utility of super-resolution technology.

In recent years, reference-based super-resolution (RefSR) methods^{24–27} have been introduced to the medical research domain to address the inherent limitations of traditional single-image super-resolution approaches. The core principle of this technology lies in leveraging additional reference images to provide complementary information for guiding high-quality reconstruction of low-resolution images. In the field of medical imaging, RefSR methods have achieved substantial progress in reconstruction accuracy by exploiting reference image information in combination with deep learning techniques^{28–30}, realizing $2\times$ spatial resolution enhancement. Subsequent advances introduced the attention mechanism³¹ to prevent learning of task-irrelevant interfering features, achieving more precise feature selection and fusion. In the domain of live-cell imaging, RefSR technology has likewise demonstrated considerable potential, with relevant methods integrating multiple bright-field and dark-field images through multi-scale learning strategies to achieve HR reconstruction across continuous scale spaces, alleviating random phase noise and fluctuation artifacts. Further, the introduction of frequency-domain constraint loss functions³² explicitly constrains solution space mapping through ensuring reconstruction results conform to actual measurements. RefSR methods continue to evolve through innovations in network architecture, learning strategies, and loss function design, demonstrating promising applications in medical imaging and live-cell imaging. However, in the field of pathological imaging, RefSR methods remain largely unexplored.

In this work, we propose a Hybrid Bright-field and Dark-field resolution enhancement framework (HBDF-RE) that effectively exploits complementary information from bright-

field and dark-field images to enhance detail reconstruction in pathological imaging. Here, the term “resolution enhancement” is used in the practical context of digital pathology to denote cross-NA image reconstruction from low-NA multimodal measurements toward a higher-NA bright-field reference. To enable multimodal image acquisition, the halogen lamp illumination module in conventional pathology scanners was replaced with a custom-designed programmable LED light source. We employ a joint optimization strategy integrating a spatial-frequency domain perceptual loss function and attention mechanisms to explicitly constrain the solution space through physical constraints, effectively suppressing reconstruction artifacts and mitigating noise interference. Our approach requires only a single dark-field reference image, while improving the input image resolution from 950 nm to 445 nm. Compared to existing models, our method achieves a 3.2 dB improvement in PSNR (Peak Signal-to-Noise Ratio), reduces reconstruction artifacts by 84%, and enhances computational efficiency by 11%. In downstream gland segmentation tasks, the HBDF-RE method significantly outperforms existing pathological resolution enhancement approaches. More importantly, HBDF-RE-reconstructed images provide higher-quality inputs for downstream artificial intelligence (AI) analysis, improving cervical cancer diagnostic sensitivity by 11.14%. It is worth noting that, while prior works have explored the combination of bright-field and dark-field illumination in the context of quantitative phase imaging—including Fourier ptychographic microscopy (FPM) with hybrid coherent and incoherent illumination³³, Fourier holo-ptychographic microscopy integrating digital holography with Fourier ptychography³⁴, and a transport-of-intensity equation (TIE)-based framework leveraging bright-field and dark-field acquisitions for high-throughput phase retrieval³⁵, these approaches are fundamentally distinct from the present work in objective, methodology, and application context. All three methods target quantitative phase reconstruction through iterative or physics-model-based algorithms and do not address intensity-image resolution enhancement of histopathological sections. Moreover, these methods generally require the acquisition of a large number of raw images and iterative reconstruction procedures to synthesize extended apertures or recover phase information, which may limit imaging throughput and reconstruction speed. In contrast, HBDF-RE constitutes a data-driven deep learning framework that exploits the scattering contrast of dark-field images as a physically grounded reference constraint to directly guide high-resolution intensity image reconstruction, without invoking phase retrieval in any form. HBDF-RE requires only two raw measurements (one bright-field and one dark-field image) and enables rapid non-iterative reconstruction through learned nonlinear mapping of high-frequency structural information, thereby achieving both high imaging efficiency and fast reconstruction performance. To our knowledge, this study is among

the first to investigate the use of dark-field images for reference-guided resolution enhancement in pathological imaging, demonstrating substantial potential for clinical applications.

2 Methods

2.1 Hybrid bright-field and dark-field resolution enhancement framework

We propose utilizing dark-field images as reference inputs to assist in reconstructing resolution enhancement images from low-resolution acquisitions. The experimental system is shown in the Fig. 1(a). By integrating a programmable LED light source into the existing scanner framework, we enable multimodal image acquisition. As shown in Fig. 1(b), the light source employs a dome-shaped geometric configuration with a curvature radius of 35 mm to achieve wide-angle illumination coverage. Each LED element exhibits a divergence half-angle of approximately $\pm 30^\circ$, enabling efficient angular illumination distribution across the imaging field. Detailed LED array configuration and illumination uniformity analysis are provided in Supplementary Section 1. This design facilitates rapid programmable switching between illumination modes (response time < 1 ms), with the industrial camera operating at a single-frame exposure time of $< 1/30$ s. As a result, a complete bright-field/dark-field image pair can be acquired within approximately $1/15$ s³⁶. Within this minimal acquisition window, sample

drift of fixed specimens remains negligible during stage stationary periods, with measured displacements consistently below 0.5 pixels, thereby precluding any significant impact on network training or reconstruction quality. Subsequently, a motorized translation stage drives the sequential scanning of successive positions, thereby achieving full field-of-view multimodal imaging.

Specifically, bright-field imaging is achieved through direct illumination from the central region of the LED array, whereas dark-field imaging is realized via oblique illumination from the peripheral annular region of the array. The integration of these two microscopy modalities is particularly attractive for biomedical applications. In bright-field microscopy, transmitted light measurements primarily reflect the light absorption properties of tissue (amplitude modulation), with image contrast arising from absorption differences. In contrast, dark-field microscopy illuminates the sample at oblique angles and detects diffracted or scattered light, whereby image contrast is predominantly based on light scattering, enabling visualization of refractive index variations (phase modulation) as well as sharp absorptive structures. It should be noted that dark-field intensity images acquired under a given objective numerical aperture (NA_{obj}) remain subject to the same diffraction-limited resolution constraints as their bright-field counterparts, with both channels structurally bounded by a maximum spatial-frequency cutoff of $2NA_{\text{obj}}$ in the intensity spectrum. Yet, this shared diffraction limit does not imply that the

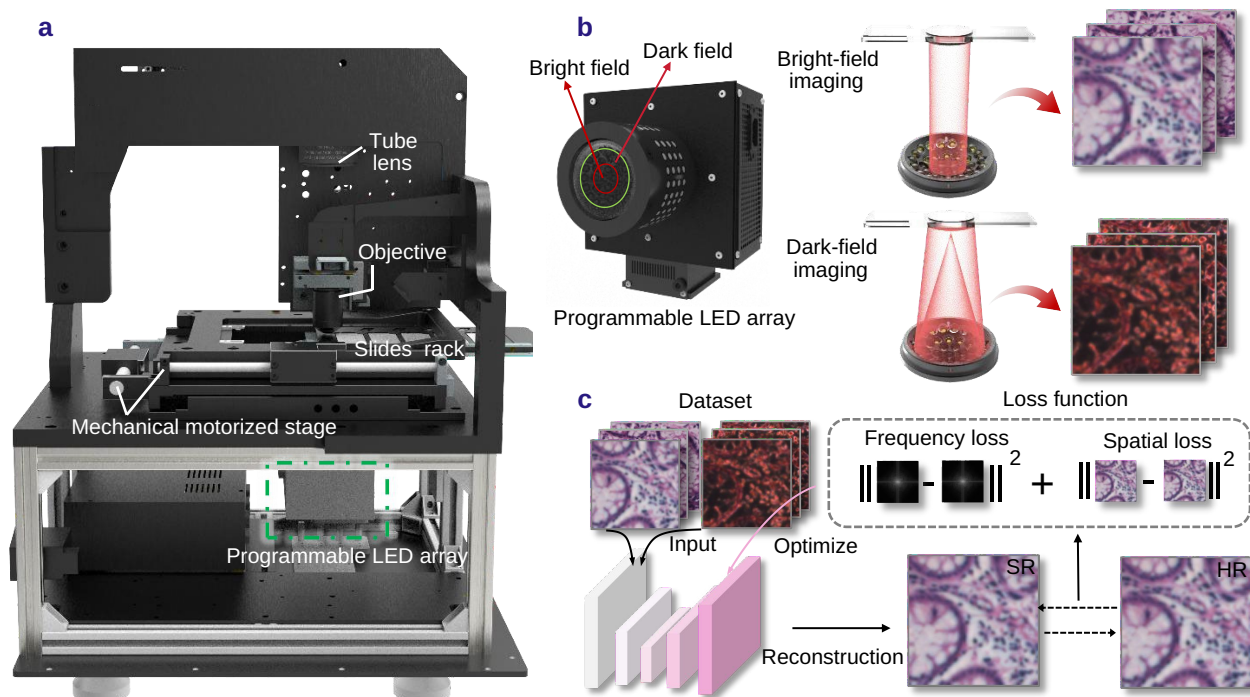


Fig. 1 | Overview of our system and algorithm diagram. (a) Imaging system. (b) Acquisition of bright-field and dark-field images. (c) The whole process of HBDF-RE.

dark-field image encodes identical spatial-frequency information. Under the well-established imaging model akin to FPM, an oblique plane-wave illumination introduces a frequency shift to the object's complex spectrum in Fourier space³⁷, mathematically translating the object's high-frequency components into the low-frequency passband of the detection objective. Critically, while the complex amplitude undergoes linear low-pass filtering by the pupil function, the subsequent square-law detection by the camera sensor records only the optical intensity; consequently, the high-frequency structural priors captured in the dark-field image are not merely translated, but rather embedded as a complex, nonlinear mapping of the object's high-frequency spectrum. Although the recorded dark-field intensity cannot provide unconstrained access to all high-frequency components, it encodes distinct high-frequency information that is inherently absent from the near-axial bright-field channel, thereby providing physically distinct, illumination-shifted structural information with a higher effective frequency support than a mere enhancement in image contrast. The inherent ill-posedness of single-image resolution enhancement implies that a single low-resolution image theoretically corresponds to infinitely many high-resolution solutions. Conventional methods rely on statistical priors learned from training data to constrain the solution space, but such implicit constraints frequently lead to over-smoothing and artifacts. Our method addresses this by introducing dark-field reference images as physically grounded constraints, leveraging their complementary spectral content to decode this nonlinear mapping and guide high-resolution reconstruction beyond the limits of bright-field data alone. This physically grounded complementary contrast, encompassing scattering-sensitive structural details, edge information, and refractive index-related features inaccessible to conventional bright-field imaging, enables the network to more accurately reconstruct the fine structural details present in high-NA bright-field images, thereby reducing the infinite solution set to a finite set consistent with physical observations. This fundamentally improves the well-posedness of the problem. Meanwhile, bright-field images acquired under the $20\times/0.4$ -NA objective serve as the higher-NA reference target for supervised network training, acknowledging that these images remain diffraction-limited and represent a higher-resolution benchmark within the conventional bright-field modality rather than an absolute ground-truth representation of the specimen. The training pipeline is illustrated in Fig. 1(c).

2.2 Network architecture

Medical image reconstruction tasks require architectures that effectively balance feature extraction with fine-grained spatial recovery³⁸. The U-Net architecture³² addresses this through symmetric encoder-decoder paths and skip connections that mitigate deep-network information loss. Despite

its success in single-modality applications, conventional U-Net lacks effective mechanisms for integrating multimodal inputs, presenting challenges for feature fusion and cross-modal information propagation. Building upon the advantages of the U-Net architecture while addressing its limitations, we propose HBDF-RE, as illustrated in Fig. 2(a), specifically optimized for resolution enhancement reconstruction tasks involving bright-field and dark-field image fusion. Considering the natural alignment characteristics between dark-field image features and LR bright-field image features, we introduce a multimodal feature fusion mechanism into HBDF-RE. The network workflow proceeds as follows: bright-field and dark-field images undergo feature extraction layers to obtain high-dimensional feature representations. Feature maps are progressively down-sampled through pooling operations along the contracting path, followed by channel-wise concatenation at the bottleneck layer to achieve multimodal information fusion. Subsequently, spatial resolution is gradually restored through transposed convolutions along the expanding path, with the final reconstruction layer outputting the SR image. A deep CNN architecture is employed for this resolution enhancement task. The network is trained on paired data triplets $(x_{\text{BF}}, x_{\text{DF}}, y_{\text{BF}})$ by minimizing the loss function in a supervised way:

$$\operatorname{argmin}_{\theta} \left\{ \frac{1}{N} \sum_{n=1}^N \|F([x_{\text{BF}}^n, x_{\text{DF}}^n]; \theta) - y_{\text{BF}}^n\|_2^2 \right\}, \quad (1)$$

where $F(\cdot; \theta)$ denotes the mapping function parameterized by θ , representing the weights and biases across all network layers. N denotes the total number of training triplets. Throughout this process, dense connections³⁹ are employed to improve cross-layer feature flow and enhance training efficiency, while effectively increasing network depth without getting trapped in local minima. In Fig. 2(b), we illustrate the dense block structure, where the l -th layer receives feature maps from all preceding $l-1$ layers. The structure is designed as follows:

$$x_l = D_l([x_0, \dots, x_{l-1}]), \quad (2)$$

where x_l denotes the concatenation of feature maps from layers $0, \dots, l-1$, and $D()$ includes batch-normalization, nonlinear activation using the LeakyReLU function, and 3×3 convolution. The key hyperparameters of the dense block include the Growth Rate is set to 12 and the Number of Convolutions is set to 5.

Furthermore, when evaluating the performance of convolutional neural networks in preserving high-frequency content in reconstructed images, relying solely on traditional spatial-domain loss functions may not be optimal. This is because minimizing pixel-level loss functions (e.g., MSE) tends to neglect high-frequency details, thereby biasing toward smooth solutions even when achieving lower loss values. Taking into account training stability, color fidelity,

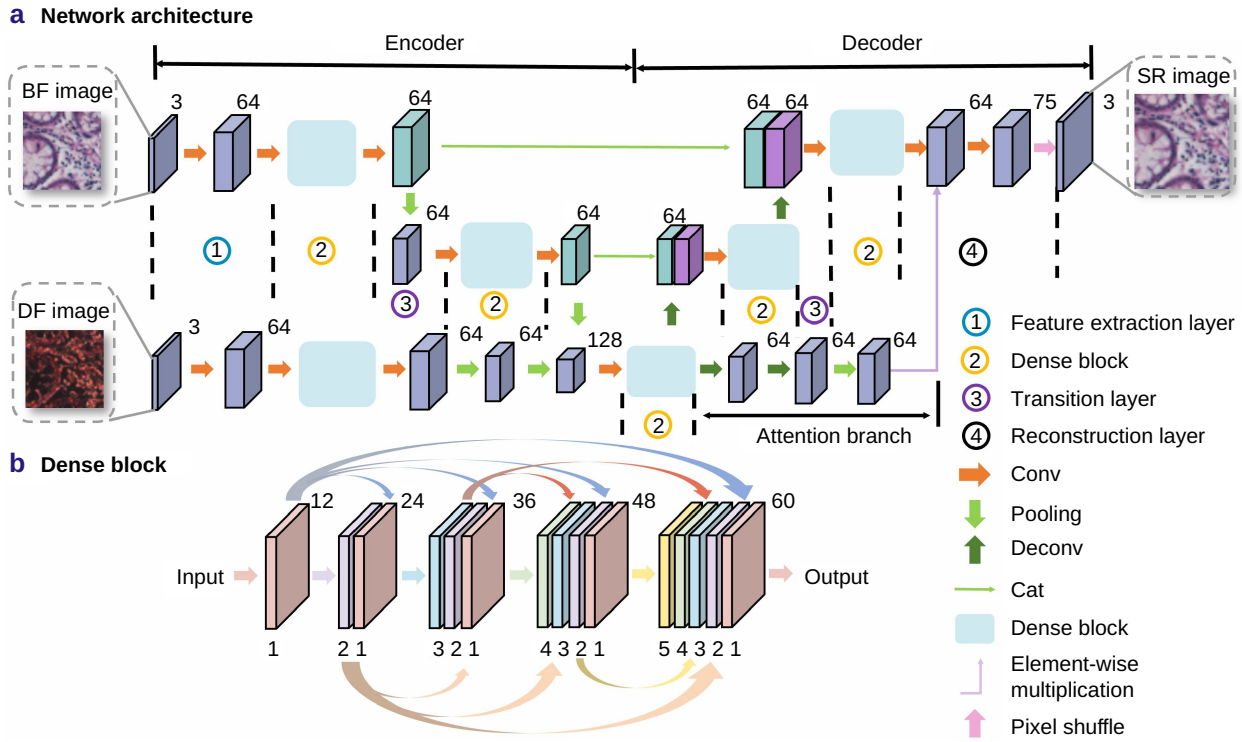


Fig. 2 | The architecture of HBDF-RE follows the form of a Encoder-Decoder framework. (a) Network architecture. (b) The components of Dense Block.

and related considerations, we propose a spatial-frequency domain perceptual loss function, defined as a weighted combination of multiple loss components. This formulation ensures that the model learns features encompassing both low-frequency and high-frequency information within images. This dual-constraint mechanism enables the model to maintain precise reproduction of fine textures while pursuing overall reconstruction quality. Specifically, our loss function comprises two components: spatial-domain mean squared error (MSE) loss and pixel-wise Fourier-domain mean squared error (FMSE) loss. The formulation is as follows:

$$Loss = \alpha Loss_{MSE} + \beta Loss_{FMSE}, \quad (3)$$

where

$$Loss_{MSE} = \frac{1}{r^2 WH} \|f(I_{LR}^n) - I_{HR}^n\|_2^2, \quad (4)$$

$$Loss_{FMSE} = \frac{1}{r^2 WH} \|\mathcal{F}(f(I_{LR}^n)) - \mathcal{F}(I_{HR}^n)\|_2^2, \quad (5)$$

where $\mathcal{F}$ denotes the 2D Fourier transform, $\|\cdot\|$ represents the L2 norm, n denotes the number of images, r represents the enhancement factor, W and H denote the width and height of the image, respectively. α and β are hyperparameters that control the relative weights of each loss component. The weight coefficients are experimentally optimized and determined as $\alpha = 0.96$ and $\beta = 0.04$.

To prevent performance degradation from irrelevant features, we incorporate an attention mechanism⁴⁰ that selectively emphasizes diagnostically relevant multimodal information. The attention module dynamically quantifies feature correlation between bright-field and dark-field images through adaptive weight allocation. The attention mechanism workflow proceeds as follows: Input images undergo hierarchical feature extraction and dense connection-based integration to produce feature tensors $x_{DF}^{encoder}$ and $x_{BF}^{encoder}$. These representations are subsequently channel-wise concatenated at the network bottleneck layer, yielding the fused feature tensor. The concatenated features then diverge into two parallel processing pathways. The decoder branch performs progressive spatial reconstruction through cascaded transposed convolution operations to generate the decoded feature tensor $x_{decoder}$:

$$x_{decoder} = f_{decoder}(\text{concat}(x_{DF}^{encoder}, x_{BF}^{encoder})), \quad (6)$$

where $f_{decoder}$ represents the sequence of transposed convolution layers in the decoder pathway. Concurrently, the attention branch applies transposed convolution followed by Sigmoid activation to produce the spatial attention weight tensor $A \in [0, 1]$:

$$A = \sigma(f_{attention}(\text{concat}(x_{DF}^{encoder}, x_{BF}^{encoder}))), \quad (7)$$

where σ denotes the Sigmoid activation function and $f_{attention}$ represents the attention pathway transformations. To leverage the learned attention weights for adaptive feature

modulation, element-wise multiplication is performed between x_{decoder} and A , producing the attention-weighted feature tensor O :

$$O = x_{\text{decoder}} \otimes A. \quad (8)$$

The attention-modulated tensor O is subsequently processed through the reconstruction module to generate the high-resolution pathology image. This attention mechanism enables selective integration of complementary multimodal information, improving quantitative metrics by over 20% compared to direct concatenation. To systematically validate the contribution of each architectural component, we conducted ablation experiments evaluating the dark-field reference input, spatial attention mechanism, and spatial-frequency domain loss function independently and in combination. Results demonstrate that the dark-field reference input constitutes the dominant contributor, accounting for approximately 60% of the total PSNR gain, while the spatial attention mechanism and frequency-domain loss function each provide further complementary improvements. Detailed ablation results are provided in Supplementary Section 2.

Beyond architectural design, training data quality significantly impacts reconstruction performance. Since our dataset comprises images acquired at varying magnifications, substantial color space discrepancies arise from instrumental differences, optical characteristics, and environmental factors^{41–43}. These inconsistencies cause models to expend computational resources learning color transformations rather than spatial detail reconstruction, degrading output quality and generalization. To address this, we implement a preprocessing strategy based on standard color charts with known spectral properties, establishing device-independent mappings that enable the network to focus on genuine spatial reconstruction. Detailed procedures are provided in Supplementary Section 3.

3 Results

3.1 Comparisons with other resolution enhancement methods

To validate the efficacy of the proposed model, we selected the mainstream single-image resolution enhancement method U-Net as the baseline and retrained and evaluated all models on the color-corrected dataset. PSNR and structural similarity index measure (SSIM) were employed as quantitative evaluation metrics. As shown in Fig. 3(a), our method achieves PSNR/SSIM of 28.80 dB/0.88, surpassing U-Net by 3.2 dB and SRGAN by 3.7 dB. While U-Net yields higher PSNR than SRGAN, GAN-based methods produce sharper visual details. Our method combines both advantages. U-Net and SRGAN serve as baselines, representing the typical performance upper bound under identical hardware conditions (i.e., using only low-resolution bright-field

images), whereas HBDF-RE demonstrates the achievable performance when a single dark-field acquisition is added at minimal additional hardware cost. The purpose of this comparison is to quantify the gain contributed by the dark-field reference.

Figure 3(b) presents intensity profiles at the marked line positions in Fig. 3(a), demonstrating that HBDF-RE reconstructed images can clearly resolve microscopic features of approximately 445 nm, approaching the theoretical half-pitch resolution limit predicted by the synthetic aperture principle ($\delta = \lambda / (NA_{\text{obj}} + NA_{\text{ill}}^{\text{DF}}) / 2 \approx 431 \text{ nm}$). Compared to the input low-resolution images, HBDF-RE achieves an approximately $2.1\times$ resolution enhancement ($4 \times / 0.16\text{-NA}$, incoherent half-pitch resolution $\delta = 1.22 \times \lambda / 2NA_{\text{obj}} / 2 \approx 953 \text{ nm}$), substantially exceeding the reconstruction capability of conventional methods. To further validate this conclusion, a quantitative evaluation experiment based on the USAF 1951 resolution target was conducted, with detailed results provided in Supplementary Section 4. The optical resolution of U-Net reconstructed images is approximately 710 nm, while GAN-based methods achieve 690 nm.

Notably, statistical analysis on the test set reveals that HBDF-RE achieves an approximately 84% reduction in artifact incidence relative to conventional deep learning approaches (averaged across U-Net and GAN-based baselines). As illustrated in Fig. 3(c), HBDF-RE reconstruction results exhibit virtually no artifacts such as checkerboard effects and ringing artifacts commonly observed in traditional deep learning models. This is attributed to the high-frequency information provided by dark-field reference images, which effectively constrains the solution space and better preserves tissue structural integrity and texture detail clarity. These introduced artifact deviations not only mislead pathologists' visual judgment but also severely compromise feature extraction and classification performance of downstream AI analysis models, leading to false-positive or false-negative diagnoses. Consequently, downstream task performance serves as a critical metric for assessing model reconstruction quality. To this end, we designed a gland segmentation experiment to evaluate the practical utility of images reconstructed by different resolution enhancement methods (detailed analysis provided in Supplementary Section 5).

Furthermore, it is worth noting that our method achieves approximately 11% improvement in reconstruction speed compared to U-Net models, completing resolution enhancement reconstruction of a single image ($3600 \mu\text{m}^2$) within 60 ms. Visualization S1 presents a comparison of WSI of pathological sections before and after HBDF-RE reconstruction. Meanwhile, to further assess the generalizability of HBDF-RE across diverse clinical scenarios, reconstruction experiments were additionally conducted on tissue specimens prepared with different histological staining protocols, including Hematoxylin and Eosin (H&E) and Immunohistochemistry (IHC) staining, with detailed results provided in Supplementary Section 6. HBDF-RE consistently maintains

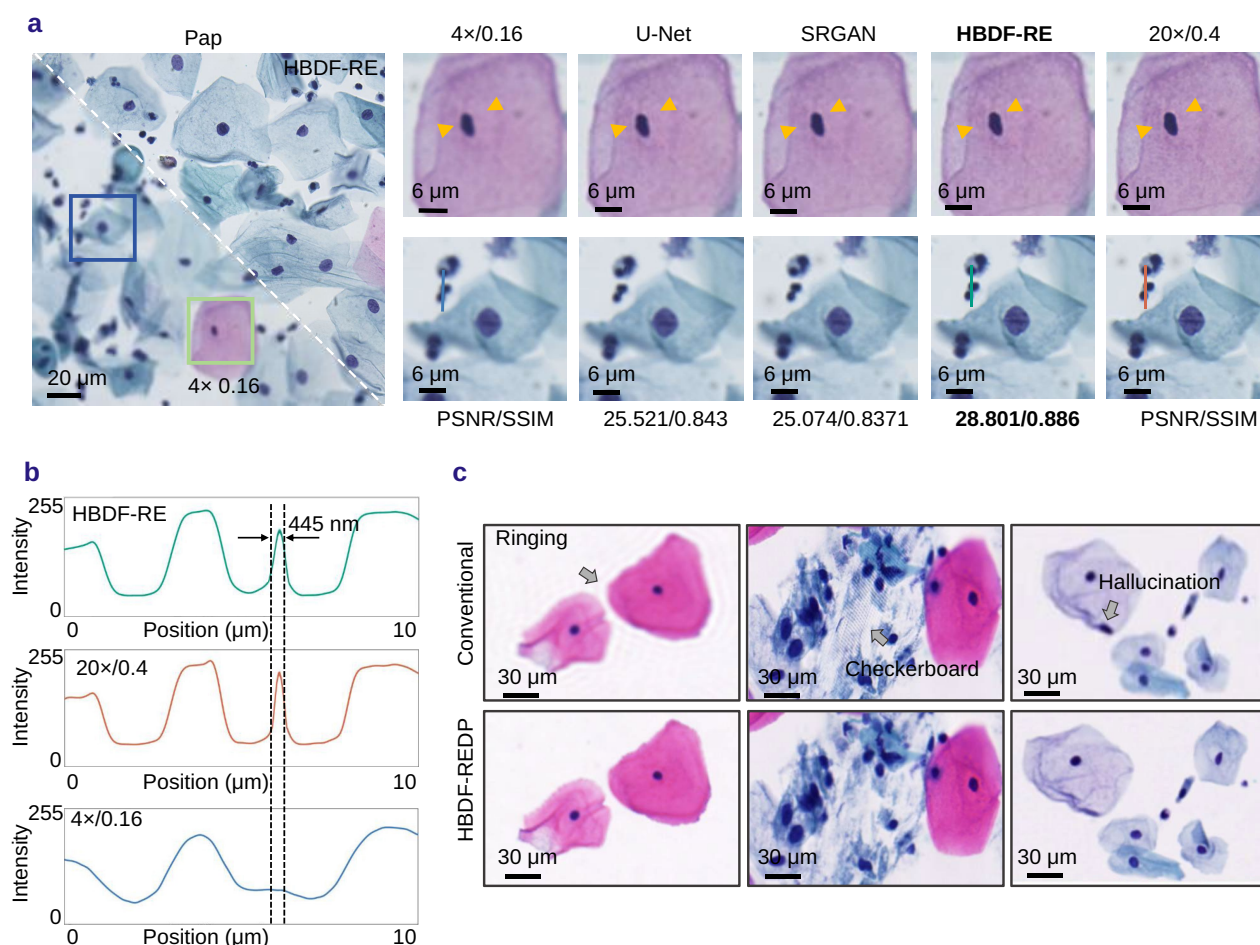


Fig. 3 | Comparisons with other resolution enhancement methods. (a) Reconstruction results display. (b) Reconstruction section line comparison. (c) Comparison of reconstruction artifacts.

high reconstruction fidelity across both staining modalities and heterogeneous tissue architectures, with stable PSNR and SSIM values observed regardless of staining characteristics. These results confirm the robustness of the proposed method and underscore its potential for broad deployment across diverse pathological workflows in clinical practice.

3.2 Analysis of resolution enhancement performance under different dark-field illumination angles

The aforementioned experiments demonstrate that introducing dark-field reference images effectively enhances reconstruction resolution and suppresses artifact generation. To systematically investigate the influence of illumination angle on reconstruction performance, we designed three dark-field imaging modes with different illumination angles: P1 (corresponding to numerical aperture $NA_{\text{III}}^{\text{DF}} \approx 0.21$), P2 ($NA_{\text{III}}^{\text{DF}} \approx 0.42$), and P3 ($NA_{\text{III}}^{\text{DF}} \approx 0.63$), and trained corresponding models and conducted testing under identical experimental conditions. As shown in Fig. 4(b), the frequency distributions of each illumination mode in

Fourier space exhibit significant differences, with yellow circles demarcating the numerical aperture range of the objective. Under incoherent illumination, the optical transfer function (OTF) support domain of the bright-field channel is bounded by a cutoff frequency of $2NA_{\text{obj}}/\lambda$, corresponding to approximately 0.32/500 nm for the 4x /0.16-NA objective. Dark-field oblique illumination, by contrast, causes the scattered light to carry spatial frequency components beyond this bright-field OTF cutoff, thereby providing the deep learning network with high-frequency structural priors that are entirely inaccessible from bright-field images alone. The frequency coverage range of these modes directly determines the quantity and quality of such high-frequency prior information available for reconstruction.

Quantitative evaluation results (Fig. 4(a)) indicate that the P2 mode achieves optimal performance, with PSNR reaching 28.59 dB and resolution of 445 nm, representing significant improvements over P1 (26.93 dB, 680 nm) and P3 (27.46 dB, 650 nm), respectively. Figure 4(c) presents the Fourier spectra corresponding to Fig. 4(a). The spectrum of the bright-field (LR) image is predominantly concentrated

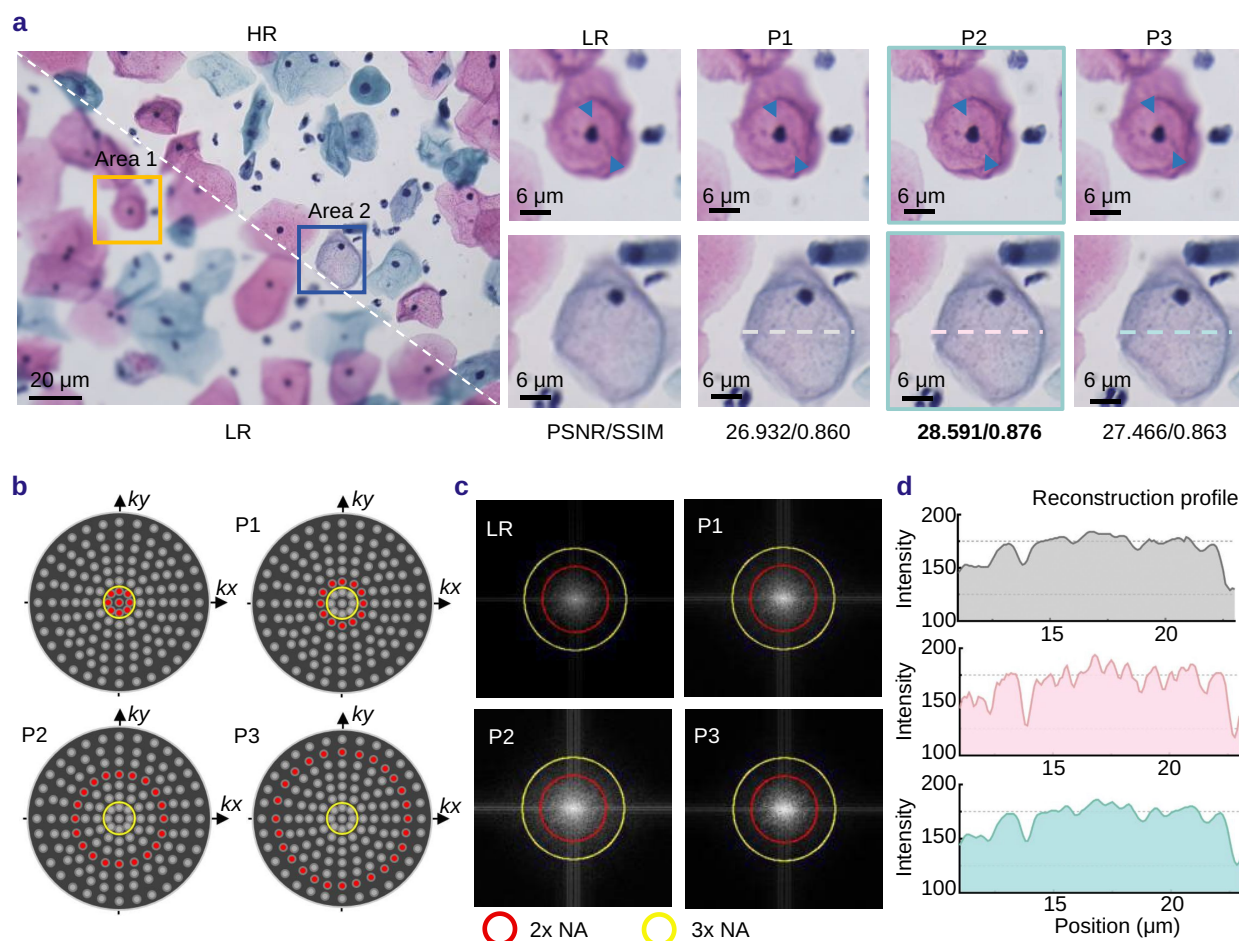


Fig. 4 | Systematic investigation of dark-field illumination angle effects on resolution enhancement reconstruction quality. (a) Reconstruction results of the model under different illumination patterns. (b) Different illumination patterns. (c) Fourier analysis of the reconstructed images. (d) Profile lines based on details in region.

within an aperture radius of $1 \times \text{NA}$, extending to the boundary of the OTF support domain ($2 \times \text{NA}$). The P1 mode effectively covers the OTF support domain, the P2 mode extends frequency coverage to approximately $3 \times \text{NA}$ —supplying the network with the richest high-frequency constraint information and most effectively reducing the ill-posed solution space of resolution enhancement, whereas the effective frequency coverage of reconstructed images under the P3 mode exhibits a significant decline relative to P2, accompanied by a markedly reduced signal-to-noise ratio in the high-frequency spectral region, as evidenced by the Fourier spectral analysis in Fig. 4(c). It should be emphasized that the resolution enhancement achieved in this work does not rely on direct frequency-domain synthesis in the manner of FPM; rather, the deep learning network exploits the high-frequency scattering information embedded in the dark-field image as a physical constraint to guide the learning of a mapping from low-resolution bright-field inputs to their high-resolution counterparts. The physical basis of this resolution gain therefore lies in the dark-field image serving

as a source of high-frequency structural priors that effectively narrows the solution space of the inverse problem, enabling more reliable recovery of fine structural detail. The grayscale intensity profiles in Fig. 4(d) further confirm the superiority of the P2 mode in edge sharpness and high-frequency information recovery. More importantly, compared to the P2 mode, the artifact incidence in reconstructed images under P1 and P3 illumination conditions increased by approximately 20%–30%. This phenomenon indicates that in dark-field-assisted resolution enhancement reconstruction, the selection of dark-field reference images is particularly critical. Under the P1 mode, the frequency coverage provided by dark-field images is limited, resulting in insufficient physical constraints on the solution space; whereas the P3 mode, despite possessing broader frequency coverage, suffers from two compounding degradation mechanisms. First, the larger dark-field illumination numerical aperture ($\text{NA}_{\text{DF}}^{\text{DF}} = 0.63$) introduces a greater oblique illumination angle, leading to further attenuation of the scattered signal energy. Under an unchanged camera noise floor, the

weakened dark-field signal becomes increasingly susceptible to imaging noise, thereby compromising the reliability of the high-frequency structural priors. Second, the incoherent bright-field channel is limited by a cutoff frequency of $2NA_{\text{obj}}/\lambda$, corresponding to spectral support up to approximately $NA = 0.32$. Under the P2 configuration, the dark-field annular passband ($NA \approx 0.26\text{--}0.58$) partially overlaps with the bright-field spectrum, enabling continuous spectral coverage from low to high spatial frequencies. In contrast, the P3 passband shifts outward ($NA \approx 0.47\text{--}0.79$), introducing an unsampled spectral gap between $NA \approx 0.32\text{--}0.47$. This discontinuity limits the network's ability to learn a physically continuous mapping between low- and high-frequency information, thereby reducing the reliability of high-frequency constraints for robust resolution enhancement. Meanwhile, the three dark-field modes were evaluated across multiple histological staining protocols, including Hematoxylin and Eosin (H&E), Papanicolaou (Pap), and Immunohistochemistry (IHC) staining. The P2 mode consistently achieves superior PSNR and SSIM performance across all staining conditions, confirming that the optimal illumination selection criterion identified above exhibits robust generalizability independent of staining modality, with detailed results provided in Supplementary Section 6. Therefore, selecting dark-field images with optimal illumination parameters and conducting rigorous quantitative characterization of illumination uniformity across different dark-field modes are both crucial for achieving robust, low-artifact reconstruction.

3.3 HBDF-RE for AI-assisted cervical cancer diagnosis

Typical cervical cytology WSIs contain tens of thousands of cells, rendering manual evaluation not only time-consuming and labor-intensive but also subject to significant inter-observer variability. Despite the widespread implementation of cervical cytology screening programs, diagnostic precision remains critically influenced by pathologist proficiency and clinical experience. Recent developments in artificial intelligence have revealed considerable promise for automating cervical cytology evaluation; nevertheless, diagnostic efficacy continues to be fundamentally limited by the quality of input images. While employing low-magnification objective lenses facilitates rapid specimen throughput, the resultant WSIs often exhibit regions of cellular blurring, which substantially impairs the extraction of critical morphological features and compromises diagnostic certainty. This tension reflects a fundamental dilemma inherent to current clinical pathological workflows: low-magnification scanning (e.g., $4\times /0.16\text{-NA}$) offers high throughput and low cost conducive to large-scale screening, yet provides insufficient resolution for reliable identification of diagnostically critical features such as nuclear atypia and mitotic figures; conversely, high-magnification scan-

ning (e.g., $20\times /0.4\text{-NA}$ or $40\times /0.65\text{-NA}$) delivers adequate diagnostic resolution at the expense of scanning times approximately an order of magnitude longer and data storage requirements nearly two orders of magnitude greater, severely constraining the feasibility of population-scale screening programs. HBDF-RE addresses this inherent throughput-resolution trade-off by supplementing low-magnification bright-field acquisition with a single dark-field capture per field of view, incurring an additional temporal overhead of less than 1/30 second, thereby achieving image quality approaching that of high-magnification scanning at minimal hardware cost. This capability carries particular practical significance for resource-limited primary healthcare settings and large-scale population screening initiatives. Accordingly, optimizing the resolution and image clarity of digitized slides represents an essential prerequisite for enhancing AI-supported diagnostic accuracy. Based on these considerations, liquid-based cervical cytology specimens were selected to evaluate the clinical applicability of the HBDF-RE methodology within AI-enhanced diagnostic pipelines.

To develop the cell-level classification model, we employed a ResNet-50-based architecture trained on liquid-based cytology specimens obtained from the repository at Ningbo Pathology Diagnostic Center. A total of 980 WSIs were acquired from the center, from which 103,258 liquid-based cytology cell specimens were extracted. The 980 WSIs were subsequently partitioned at the slide level into training, validation, and test subsets following a 7 : 1 : 2 stratified allocation, yielding 686, 98, and 196 WSIs, respectively. All specimens were digitally scanned and annotated by experienced pathologists who classified them into five categories based on clinical management protocols: The high-grade lesion and carcinoma category includes high-grade squamous intraepithelial lesions (HSIL) and squamous cell carcinoma (SCC), all indicating significant precancerous progression or confirmed malignancy necessitating prompt colposcopic examination, tissue biopsy, and oncological treatment pathways. The ASC-US category consists of atypical squamous cells of undetermined significance, representing minimal malignant transformation potential and requiring human papillomavirus (HPV) testing to determine subsequent management. The LSIL category encompasses low-grade squamous intraepithelial lesions, indicating mild cellular changes that warrant routine colposcopic observation. The glandular abnormality category includes atypical glandular cells (AGC) and adenocarcinoma in situ (AIS), requiring targeted diagnostic approaches including endocervical curettage or endometrial tissue collection. Lastly, the normal findings category comprises cases showing Negative for Intraepithelial Lesion or Malignancy (NILM), requiring no additional clinical action. The class distribution across the five categories is as follows: NILM accounts for approximately 68.3% of specimens, followed by ASC-US (12.7%), LSIL (9.8%), HSIL/SCC (6.4%), and AGC/AIS (2.8%),

reflecting the inherent class imbalance characteristic of clinical cytological screening populations. To mitigate the adverse effects of this imbalance on model performance, targeted data augmentation strategies, including random rotation, horizontal and vertical flipping, and color jittering, were applied exclusively to the minority classes (AGC/AIS and HSIL/SCC), thereby enhancing the model's discriminative capacity for these diagnostically critical yet underrepresented categories.

The AI-assisted diagnostic workflow is depicted in Fig. 5(a). Following digitization of liquid-based cytology slides via scanning, a pre-trained deep learning classification model autonomously identifies the majority of cells, enabling pathologists to concentrate exclusively on AI-designated regions of diagnostic interest. This paradigm obviates the necessity for exhaustive manual examination of hundreds to thousands of individual cells, thereby markedly enhancing diagnostic efficiency. Nevertheless, inherent resolution constraints of conventional scanning systems frequently compromise AI model performance when confronted with suboptimal image quality. As demonstrated in Fig. 5(b), morphologically ambiguous cells often result in systematic misclassification, with atypical and malignant lesions being erroneously categorized as cellular debris. Conversely, application of HBDF-RE resolution enhancement reconstruction substantially ameliorates diagnostic accuracy, enabling the AI model to correctly identify pathological entities that would otherwise be misclassified. To rigorously quantify the clinical utility of HBDF-RE enhancement, we implemented a comprehensive comparative framework comprising six

distinct diagnostic paradigms: AI-based autonomous interpretation, expert pathologist interpretation, and AI-augmented collaborative interpretation, with each modality systematically evaluated on both native scanner-acquired images and HBDF-RE reconstructed counterparts. The test set comprised 34,638 cellular specimens extracted from 196 WSIs. Twelve board-certified pathologists from Zhejiang Provincial People's Hospital were stratified into two independent annotation groups of six members each, with discordant classifications adjudicated by senior expert consensus to establish definitive reference standards.

The results demonstrate that HBDF-RE substantially enhances the entire AI-assisted diagnostic process, from data quality to interpretation reliability. Systematic performance evaluation across all diagnostic categories revealed consistent improvements in accuracy, sensitivity, and specificity (Fig. 5(c)), with sensitivity showing a relative enhancement of 11.14%. Comprehensive metrics across all diagnostic categories are summarized in Extended Data Table 1 and Fig. 5(d). Notably, the HBDF-RE-enhanced AI system achieved superior diagnostic performance compared to all other evaluated paradigms, with particularly pronounced improvements in diagnostically challenging categories such as AGC/AIS. Compared to the original images, U-Net reconstruction improved diagnostic performance by 8.48% and SRGAN by 9.76% in AGC/AIS cases, while HBDF-RE achieved a substantially higher improvement of 17.32%. These findings collectively demonstrate that HBDF-RE reconstruction amplifies AI diagnostic capability throughout the clinical workflow. Moreover, the substantial artifact

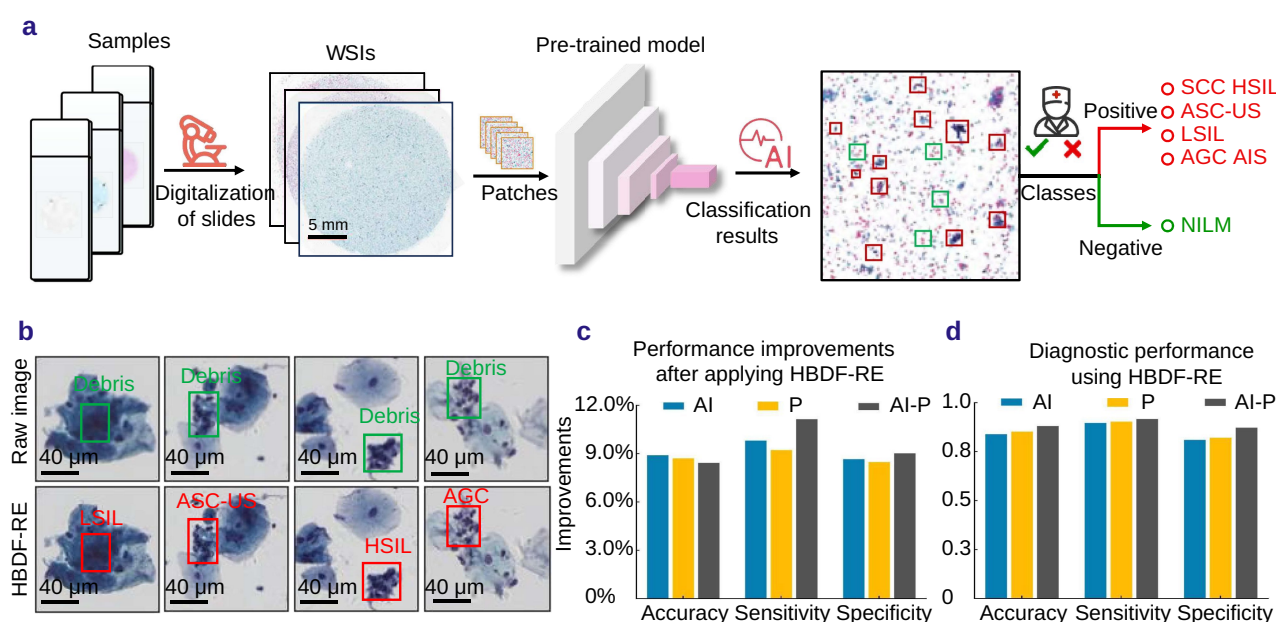


Fig. 5 | Comparative analysis under different diagnostic conditions. (a) AI-assisted cervical cancer diagnosis process. (b) Comparisons of AI-assisted results between Raw WSI and HBDF-RE WSI. (c) Relative improvements in accuracy, sensitivity, and specificity after applying HBDF-RE (without NILM cases). (d) Accuracy, sensitivity, and specificity of full cervical cytological grading after applying HBDF-RE.

Table 1 | Sensitivity, specificity and accuracy for AI, P, and AI-P across different categories under raw image and HBDF-RE method.

Performance metrics	AI Raw image	P Raw image	AI-P Raw image	AI HBDF-RE	AI HBDF-RE	AI-P HBDF-RE	AI HBDF-RE vs. AI Raw image	P HBDF-RE vs. P Raw image	AI-P HBDF-RE vs. AI-P Raw image
SCC HSIL									
Sensitivity (95% CI)	0.907 [0.883 - 0.958]	0.919 [0.891 - 0.964]	0.924 [0.878 - 0.956]	0.937 [0.908 - 0.968]	0.974 [0.931 - 0.982]	0.989 [0.907 - 1.000]	3.31%	5.98%	7.53%
Specificity (95% CI)	0.920 [0.889 - 0.954]	0.939 [0.901 - 0.956]	0.957 [0.904 - 0.971]	0.977 [0.931 - 0.997]	0.983 [0.946 - 0.989]	0.989 [0.956 - 0.998]	6.19%	4.58%	3.57%
Accuracy (95% CI)	0.909 [0.882 - 0.949]	0.927 [0.906 - 0.947]	0.939 [0.896 - 0.969]	0.956 [0.905 - 0.972]	0.965 [0.929 - 0.981]	0.987 [0.955 - 0.996]	5.10%	4.09%	5.12%
ASC-US									
Sensitivity (95% CI)	0.853 [0.810 - 0.878]	0.867 [0.801 - 0.880]	0.889 [0.825 - 0.903]	0.941 [0.908 - 0.968]	0.959 [0.931 - 0.982]	0.979 [0.967 - 0.990]	10.66%	10.69%	10.51%
Specificity (95% CI)	0.847 [0.802 - 0.854]	0.853 [0.824 - 0.879]	0.879 [0.830 - 0.886]	0.950 [0.901 - 0.987]	0.963 [0.946 - 0.989]	0.978 [0.953 - 0.998]	12.28%	12.97%	11.26%
Accuracy (95% CI)	0.838 [0.818 - 0.849]	0.840 [0.805 - 0.857]	0.882 [0.855 - 0.898]	0.936 [0.919 - 0.972]	0.956 [0.929 - 0.981]	0.973 [0.955 - 0.996]	11.69%	13.81%	10.32%
LSIL									
Sensitivity (95% CI)	0.875 [0.842 - 0.887]	0.907 [0.891 - 0.930]	0.910 [0.902 - 0.955]	0.923 [0.897 - 0.998]	0.957 [0.906 - 0.996]	0.984 [0.903 - 0.998]	5.46%	5.80%	8.21%
Specificity (95% CI)	0.851 [0.809 - 0.864]	0.886 [0.823 - 0.895]	0.917 [0.833 - 0.928]	0.926 [0.899 - 0.967]	0.938 [0.902 - 0.989]	0.985 [0.929 - 0.998]	8.73%	6.13%	7.29%
Accuracy (95% CI)	0.857 [0.809 - 0.877]	0.896 [0.849 - 0.928]	0.933 [0.861 - 0.958]	0.927 [0.896 - 0.980]	0.945 [0.907 - 0.992]	0.988 [0.918 - 0.997]	8.17%	5.67%	5.89%
AGC AIS									
Sensitivity (95% CI)	0.745 [0.713 - 0.772]	0.749 [0.726 - 0.783]	0.768 [0.738 - 0.794]	0.868 [0.838 - 0.898]	0.878 [0.841 - 0.902]	0.901 [0.877 - 0.930]	16.32%	17.22%	17.32%
Specificity (95% CI)	0.798 [0.746 - 0.810]	0.803 [0.732 - 0.837]	0.811 [0.783 - 0.846]	0.901 [0.861 - 0.937]	0.907 [0.856 - 0.929]	0.922 [0.889 - 0.928]	12.20%	12.67%	13.69%
Accuracy (95% CI)	0.766 [0.727 - 0.794]	0.778 [0.727 - 0.802]	0.783 [0.739 - 0.817]	0.885 [0.855 - 0.912]	0.891 [0.879 - 0.911]	0.908 [0.885 - 0.926]	15.54%	14.52%	15.96%
NILM									
Sensitivity (95% CI)	0.928 [0.900 - 0.940]	0.940 [0.938 - 0.962]	0.959 [0.948 - 0.962]	0.963 [0.950 - 0.982]	0.975 [0.965 - 0.985]	0.980 [0.913 - 0.998]	3.77%	3.70%	2.19%
Specificity (95% CI)	0.893 [0.860 - 0.912]	0.897 [0.883 - 0.919]	0.905 [0.882 - 0.908]	0.963 [0.899 - 0.986]	0.974 [0.951 - 0.987]	0.985 [0.950 - 0.999]	7.81%	8.58%	8.84%
Accuracy (95% CI)	0.905 [0.875 - 0.935]	0.919 [0.901 - 0.925]	0.938 [0.921 - 0.951]	0.963 [0.915 - 0.985]	0.979 [0.885 - 0.986]	0.986 [0.911 - 0.998]	6.41%	6.53%	5.42%

reduction not only enhances algorithmic performance but also provides pathologists with clearer, more reliable images for manual interpretation. To complement the objective 84% reduction in artifact incidence, a perceptual evaluation was conducted using a Mean Opinion Score (MOS) protocol, in which six board-certified pathologists independently performed blinded assessments of 1,000 reconstruction images randomly sampled from the test set. Evaluators rated each image on a five-point scale ranging from 1 (severe artifacts present) to 5 (no visible artifacts), with assessment criteria encompassing the severity of checkerboard effects,

ringing artifacts, and textural distortion. Images receiving MOS scores below 3 were subsequently re-evaluated following HBDF-RE processing; 84% demonstrated score improvements to the 4–5 range, corroborating the quantitative findings and confirming that HBDF-RE effectively suppresses the artifact categories most prevalent in deep learning-based resolution enhancement, with strong implications for clinical deployment. HBDF-RE reconstructed images thereby offer artifact-free tissue representation and significantly reduced scanning time, providing key clinical advantages.

4 Conclusions

In this study, we propose a deep learning approach (HBDF-RE) that integrates bright-field and dark-field dual-modal images to systematically address the prevalent issues of artifact generation and detail loss in traditional single-image resolution enhancement reconstruction. HBDF-RE exploits complementary modal characteristics through an encoder-decoder network integrated with spatial attention mechanisms, fusing low-resolution bright-field and dark-field images to enable end-to-end learning and enhanced feature representation. The spatial-frequency domain joint loss function ensures simultaneous optimization of low-frequency structural integrity and high-frequency details. Experimental results show that HBDF-RE reconstructs images approaching the quality of the higher-NA bright-field reference, achieving an approximately $2.1\times$ effective structural resolvability enhancement relative to the low-NA input, and compared with typical single-image resolution enhancement methods, the PSNR is improved by 3.2 dB, while the occurrence of artifacts is reduced by 84%. More importantly, in AI-assisted cervical cancer diagnosis, the diagnostic sensitivity of HBDF-RE reconstructed images is improved by 11.14%. These results demonstrate that HBDF-RE provides a practical and scalable solution for robust resolution enhancement in digital pathology. By enabling high-performance imaging through affordable, low-NA microscopy rather than expensive optics system, this method represents a paradigm shift from hardware-dependent to deep learning-enabled enhancement, paving the way toward accessible whole-slide imaging in clinical settings and integration into next-generation diagnostic platforms. Future work will focus on rigorous quantitative validation through OTF analysis of effective contrast transfer under different illumination modes, which will further consolidate the theoretical foundation of the proposed framework and strengthen its translational potential in clinical deployment.

5 Experimental setup and dataset preparation

Given the absence of publicly available cytopathological section datasets for model training, we utilized 90 tissue samples provided by KonFoong Biotech International Co., Ltd. (Yuyao, Zhejiang, China), encompassing multiple staining types. By integrating a custom-designed programmable illumination source into the company's KF-PRO-005-EX scanner system with control software, we digitized the 90 tissue samples into WSI across different modalities (bright-field and dark-field) and resolutions ($4\times$ /0.16-NA and $20\times$ /0.4-NA). The system is equipped with the color camera (The Imaging Source, DFK 33UX226, $1.85\ \mu\text{m}$ pixel size) operating at a frame rate of 30 fps. Bright-field and dark-field images acquired under the $4\times$ objective served as model inputs, while bright-field images acquired under the

$20\times$ objective served as ground-truth training labels.

All data were subjected to color correction using the proposed color calibration strategy. During data processing, $20\times$ images were divided into 640×640 pixel patches with 50% overlap to ensure boundary continuity, while corresponding $4\times$ bright-field and dark-field images were cropped to 128×128 pixels for network training. Information entropy filtering (threshold < 5) was employed to remove blank background regions. Ultimately, 146,776 strictly registered image patch pairs were obtained, with 102,744 pairs allocated for model training and the remaining 44,032 pairs for testing. During testing, weighted average stitching was performed using the ImageJ library in FIJI software to handle overlapping regions. All models were trained using the Adam optimizer with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The initial learning rate was set to 1×10^{-4} , with a scheduling strategy that halved the learning rate every 300 epochs. Training was conducted on a platform equipped with an Intel Core™ i9-14900K processor and NVIDIA RTX 4090 GPU, running Ubuntu 20.04.

References

1. Esteva A, Kuprel B, Novoa RA et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
2. Campanella G, Hanna MG, Geneslaw L et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* **25**, 1301–1309 (2019).
3. Kather JN, Krisam J, Charoentong P et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med* **16**, e1002730 (2019).
4. Schmauch B, Romagnoni A, Pronier E et al. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nat Commun* **11**, 3877 (2020).
5. Wang TJ, Dremel J, Richter S et al. Resolution-enhanced multi-core fiber imaging learned on a digital twin for cancer diagnosis. *Neurophotonics* **11**, S11505 (2024).
6. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J* **16**, 34–42 (2018).
7. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inf* **7**, 29 (2016).
8. Sun JW, Wu JC, Wu S et al. Quantitative phase imaging through an ultra-thin lensless fiber endoscope. *Light Sci Appl* **11**, 204 (2022).
9. Philipp K, Smolarski A, Koukourakis N et al. Volumetric HiLo microscopy employing an electrically tunable lens. *Opt Express* **24**, 15029–15041 (2016).
10. Dong C, Loy CC, He KM et al. Learning a deep convolutional network for image super-resolution. In *Proceedings of the 13th European Conference on Computer Vision* 184–199 (Springer, 2014). http://doi.org/10.1007/978-3-319-10593-2_13.
11. Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 1646–1654 (IEEE, 2016). <http://doi.org/10.1109/CVPR.2016.182>.
12. Lim B, Son S, Kim H et al. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* 136–144 (IEEE, 2017). <http://doi.org/10.1109/CVPRW.2017.151>.

13. Ledig C, Theis L, Huszár F et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 105–114 (IEEE, 2017). <http://doi.org/10.1109/CVPR.2017.19>.
14. Wang XT, Xie LB, Dong C et al. Real-ESRGAN: training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* 1905–1914 (IEEE, 2021). <http://doi.org/10.1109/ICCVW54120.2021.00217>.
15. Chen ZQ, Sun JW, Yang XB et al. Diffusion-driven lensless fiber endomicroscopic quantitative phase imaging toward digital pathology. *Adv Imaging* **2**, 041003 (2025).
16. Helgadottir S, Midtvedt B, Pineda J et al. Extracting quantitative biological information from bright-field cell images using deep learning. *Biophys Rev* **2**, 031401 (2021).
17. Cao JZ, Liang JY, Zhang K et al. Reference-based image super-resolution with deformable attention transformer. In *Proceedings of the 17th European Conference on Computer Vision* 325–342 (Springer, 2022). http://doi.org/10.1007/978-3-031-19797-0_19.
18. Shim G, Park J, Kweon IS. Robust reference-based super-resolution with similarity-aware deformable convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 8422–8431 (IEEE, 2020). <http://doi.org/10.1109/CVPR42600.2020.00845>.
19. Yang FZ, Yang H, Fu JL et al. Learning texture transformer network for image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 5790–5799 (IEEE, 2020). <http://doi.org/10.1109/CVPR42600.2020.00583>.
20. Shevkunov I, Kandhavelu M, Egiazarian K. A deep learning-based concept for quantitative phase imaging upgrade of bright-field microscope. *Appl Phys Lett* **124**, 043702 (2024).
21. Meng Z, Ding LQ, Feng ST et al. Numerical dark-field imaging using deep-learning. *Opt Express* **28**, 34266–34278 (2020).
22. Zheng HT, Ji MQ, Wang HQ et al. CrossNet: an end-to-end reference-based super resolution network using cross-scale warping. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)* 87–104 (Springer, 2018). http://doi.org/10.1007/978-3-030-01231-1_6.
23. Lyu Q, Shan HM, Steber C et al. Multi-contrast super-resolution MRI through a progressive network. *IEEE Trans Med Imaging* **39**, 2738–2749 (2020).
24. Zeng K, Zheng H, Cai CB et al. Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. *Comput Biol Med* **99**, 133–141 (2018).
25. Zheng H, Qu XB, Bai ZJ et al. Multi-contrast brain magnetic resonance image super-resolution using the local weight similarity. *BMC Med Imaging* **17**, 6 (2017).
26. Zheng H, Zeng K, Guo D et al. Multi-contrast brain MRI image super-resolution with gradient-guided edge enhancement. *IEEE Access* **6**, 57856–57867 (2018).
27. Feng CM, Fu HZ, Yuan SH et al. Multi-contrast MRI super-resolution via a multi-stage integration network. In *Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention* 140–149 (Springer, 2021). http://doi.org/10.1007/978-3-030-87231-1_14.
28. Georgescu MI, Ionescu RT, Miron AI et al. Multimodal multi-head convolutional attention with various kernel sizes for medical image super-resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* 2194–2204 (IEEE, 2023). <http://doi.org/10.1109/WACV56688.2023.00223>.
29. Nguyen T, Xue YJ, Li YZ et al. Deep learning approach for Fourier ptychography microscopy. *Opt Express* **26**, 26470–26484 (2018).
30. Xue YJ, Cheng SY, Li YZ et al. Reliable deep-learning-based phase imaging with uncertainty quantification. *Optica* **6**, 618–629 (2019).
31. Wang H, Zhu JB, Li YZ et al. Local conditional neural fields for versatile and generalizable large-scale reconstructions in computational imaging. arXiv (2023-7-12) <https://doi.org/10.48550/arXiv.2307.06207>
32. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention* 234–241 (Springer, 2015). http://doi.org/10.1007/978-3-319-24574-4_28.
33. Fan Y, Sun JS, Chen Q et al. Efficient Fourier ptychographic microscopy with hybrid coherent and incoherent illumination. *Proc SPIE* **12316**, 123160D (2022).
34. Li ZS, Shu YF, Lu LP et al. Fourier holo-ptychographic microscopy: a hybrid digital holography-Fourier ptychography approach to synthetic-aperture quantitative phase microscopy. *ACS Photonics* **12**, 2965–2971 (2025).
35. Lu LP, Li JJ, Shu YF et al. Hybrid brightfield and darkfield transport of intensity approach for high-throughput quantitative phase microscopy. *Adv Photonics* **4**, 056002 (2022).
36. Liu K, Wu J, Cao J et al. High-quality simultaneous bright-field and dark-field imaging based on the light-field matrix model. *Appl Phys Lett* **126**, 173702 (2025).
37. Zheng GA, Horstmeyer R, Yang C. Wide-field, high-resolution Fourier ptychographic microscopy. *Nat Photonics* **7**, 739–745 (2013).
38. Gao ZY, Ge XL, Zhu LC et al. Flow field estimation with distortion correction based on multiple input deep convolutional neural networks and Hartmann-Shack wavefront sensing. *Photonics* **11**, 452 (2024).
39. Huang G, Liu Z, Van Der Maaten L et al. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2261–2269 (IEEE, 2017). <http://doi.org/10.1109/CVPR.2017.243>.
40. Woo S, Park J, Lee JY et al. CBAM: convolutional block attention module. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)* 3–19 (Springer, 2018). http://doi.org/10.1007/978-3-030-01234-2_1.
41. Cheung V, Westland S, Connah D et al. A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms. *Color Technol* **120**, 19–25 (2004).
42. Chen LC, Zhu YK, Papandreou G et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)* 833–851 (Springer, 2018). http://doi.org/10.1007/978-3-030-01234-2_49.
43. Sirinukunwattana K, Pluim JPW, Chen H et al. Gland segmentation in colon histology images: the glas challenge contest. *Med Image Anal* **35**, 489–502 (2017).

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2024YFE0101300, 2022YFA1205002, 2024YFF0505603, 2024YFF0505600), National Natural Science Foundation of China (62227818, 62361136588, 62305162, U21B2033, 62505136, 62575139), Fundamental Research Funds for the Central Universities (30923010206), Fundamental Research Funds for the Central Universities (2024202002), National Key Laboratory of Plasma Physics (JCKYS2024212804), and Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense (JSGPCXZNGZ202401).

Competing interests

The authors declare no conflicts of interest.

Data availability

All data are available in the main text or the supplementary information from the corresponding author upon reasonable request.

Supplementary information

Supplementary information for this paper is available at
<https://doi.org/10.29026/oea.2026.260060>



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 license, and indicate if changes were made. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

©The Author(s) 2026.

Published by Editorial Office of *Opto-Electronic Advance*, Institute of Optics and Electronics, Chinese Academy of Sciences.

