

Complex-domain-enhancing neural network for large-scale coherent imaging

Xuyang Chang^{a,b}, Rifa Zhao^{a,b}, Shaowei Jiang^c, Cheng Shen^d, Guoan Zheng^e, Changhui Yang^d, and Liheng Bian^{a,b,f,*}

^aBeijing Institute of Technology, MIIT Key Laboratory of Complex-Field Intelligent Sensing, Beijing, China

^bBeijing Institute of Technology, School of Information and Electronics and Advanced Research Institute of Multidisciplinary Sciences, Beijing, China

^cHangzhou Dianzi University, School of Communication Engineering, Hangzhou, China

^dCalifornia Institute of Technology, Department of Electrical Engineering, Pasadena, California, United States

^eUniversity of Connecticut, Department of Biomedical Engineering, Storrs, Connecticut, United States

^fYangtze Delta Region Academy of Beijing Institute of Technology (Jiaxing), Jiaxing, China

Abstract. Large-scale computational imaging can provide remarkable space-bandwidth product that is beyond the limit of optical systems. In coherent imaging (CI), the joint reconstruction of amplitude and phase further expands the information throughput and sheds light on label-free observation of biological samples at micro- or even nano-levels. The existing large-scale CI techniques usually require scanning/modulation multiple times to guarantee measurement diversity and long exposure time to achieve a high signal-to-noise ratio. Such cumbersome procedures restrict clinical applications for rapid and low-phototoxicity cell imaging. In this work, a complex-domain-enhancing neural network for large-scale CI termed CI-CDNet is proposed for various large-scale CI modalities with satisfactory reconstruction quality and efficiency. CI-CDNet is able to exploit the latent coupling information between amplitude and phase (such as their same features), realizing multidimensional representations of the complex wavefront. The cross-field characterization framework empowers strong generalization and robustness for various coherent modalities, allowing high-quality and efficient imaging under extremely low exposure time and few data volume. We apply CI-CDNet in various large-scale CI modalities including Kramers–Kronig-relations holography, Fourier ptychographic microscopy, and lensless coded ptychography. A series of simulations and experiments validate that CI-CDNet can reduce exposure time and data volume by more than 1 order of magnitude. We further demonstrate that the high-quality reconstruction of CI-CDNet benefits the subsequent high-level semantic analysis.

Keywords: complex-domain neural network; coherent imaging; phase retrieval.

Received Mar. 9, 2023; revised manuscript received May 25, 2023; accepted for publication Jun. 14, 2023; published online Jul. 4, 2023.

© The Authors. Published by SPIE and CLP under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.

[DOI: [10.1117/1.APN.2.4.046006](https://doi.org/10.1117/1.APN.2.4.046006)]

1 Introduction

Large-scale coherent imaging (CI) has brought about a paradigm shift in our understanding of optical imaging, from morphological manifestation to quantitative measurement.^{1–9} The information throughput of the optical system is defined by the space-bandwidth product (SBP), which represents the number of optically resolvable spots within the field of view

(FOV).^{6,10} In CI, the joint reconstruction of amplitude and phase further expands SBP to billions, realizing both wide FOV and high-resolution imaging.^{9,11,12} The remarkable throughput and resolving capacity provide cellular and molecular insights for biomedical research.^{13–15} Large-scale CI techniques generally require certain types of diversity measurements in the spatial domain (e.g., lensless on-chip systems^{16–20}) or the Fourier domain (e.g., Fourier ptychography^{8,9}). Tens or hundreds of intensity-only measurements are often needed to reconstruct the sample's complex wavefront. Such high-volume data make large-scale

*Address all correspondence to Liheng Bian, bian@bit.edu.cn

imaging time-consuming and computationally expensive. Although reducing measurement data volume and exposure time are straightforward strategies, they would sacrifice imaging resolution and signal to noise ratio (SNR).

The image denoising technique has emerged as an effective method for improving imaging quality and confronting insufficient measurement and illumination. However, the conventional model-driven denoising techniques^{21–25} suffer from high computational complexity, making them impractical for high-throughput CI. Recent advanced deep-learning technology introduces the data-driven strategy for image enhancement tasks, providing rapid and flexible solutions for computational imaging.^{26,27} In one instance,²⁷ the convolutional neural network (CNN) is able to learn a mapping from noisy images to noise-free images directly, reducing several orders of magnitude of reconstruction time. Although CNN-based techniques have achieved great success in real-domain image denoising, there are several challenges for large-scale CI. First, the existing CNN architecture, training strategy, and degradation model of data sets are designed for intensity-only images. They do not consider the amplitude-phase correlations of complex-domain signals that have been widely used in neurosciences and speech signal processing fields.^{28–31} Second, conventional real-domain enhancement neural networks typically rely on distinct parameters to adequately capture the characteristics of both amplitude and phase, in order to achieve satisfactory denoising performance. This trade-off between denoising performance and efficiency poses a challenge for such networks.^{32,33} Third, image denoising often smooths edges and sacrifices imaging resolution, which is contradictory to the goal of superresolution coherent reconstruction. In summary, the existing large-scale CI techniques require a trade-off between imaging quality and efficiency, which restricts their clinic applications for rapid and low-phototoxicity imaging.³⁴

Recent advancements in complex-domain neural networks^{28,30,35} have achieved significant success in complex signal processing. For example, Trabelsi et al.³⁰ applied it to speech signal processing, resulting in improved accuracy. Zhang et al.³⁵ combined complex-domain neural networks with deep unfolded techniques to achieve high-quality lensless imaging. In this work, we introduce the complex-domain neural network to enhance large-scale CI, termed as CI-CDNet. We demonstrate its wide applications for various large-scale CI modalities with remarkable quality and efficiency. CI-CDNet effectively utilizes latent coupling information, which involves the feature aliasing between amplitude and phase images, to overcome the reconstruction ambiguity associated with phase information. By doing so, it enables multidimensional representations of the complex wavefront, thereby facilitating the effective suppression of complex multisource measurement noise in computational imaging while preserving fine details and achieving high imaging resolution. In addition, CI-CDNet processes the complex wavefront in a one-step and end-to-end manner, maintaining remarkable performance and efficiency. Specifically, we derived the two-dimensional complex-domain convolution unit, the corresponding activate function, and built the comprehensive multisource noise model for CI, which includes speckle noise, Poisson noise, Gaussian noise, and superresolution reconstruction noise. We then trained CI-CDNet using the derivate multisource noise model and demonstrated it in various large-scale CI modalities, including noniterative Kramers–Kronig-relations (KKR) holography,^{3,36–38} Fourier

ptychographic microscopy (FPM),^{8,9} and lensless coded ptychography (LCP).^{19,39–41} The results indicate that CI-CDNet obtains state-of-the-art performance in accuracy, computational efficiency, and imaging resolution. It is able to reduce exposure time and data volume by more than 1 order of magnitude. Finally, we further demonstrated that the high-quality reconstruction of the proposed technique benefits the subsequent high-level semantic analysis, such as cell segmentation and virtual staining.

2 Methods

2.1 Complex-Domain Neural Network

The architecture of the proposed CI-CDNet is presented in Fig. 1(a). The input contains a complex wavefront and a noise map. The noise map makes the denoising degree flexible in the iterative reconstruction, which is able to balance the smoothness and fidelity (Note 1 in the [Supplemental Material](#)). The backbone of CI-CDNet is a complex-domain U-Net that contains multiple residual blocks to increase the modeling capacity. Specifically, it contains four downsampling and upsampling scales with 64, 128, 256, and 512 channels, respectively. Each scale has an identity skip connection between 2×2 complex-domain strided convolution (CD-SConv) downsampling and 2×2 complex-domain transposed convolution (CD-TConv) upsampling operations. In addition, we employed successive complex-domain residual blocks, which consist of CD-Conv, CD-Relu, and CD-Conv in the downscaling and upscaling of each scale. The proposed CI-CDNet utilized the complex operation and block as the basic units. The detailed formalism of each block is shown below.

2.1.1 Complex-domain convolution

Figure 1(b) shows the complex convolution operator. Assume that the complex feature map and convolution kernel are represented as $F = F_R + iF_I$ and $K = K_R + iK_I$, respectively, where F_R and K_R are the real parts, F_I and K_I are the imaginary parts, and i is the imaginary unit. Then, the complex-domain convolution can be indicated as

$$\begin{aligned} F * K &= (F_R + iF_I) * (K_R + iK_I) \\ &= (F_R * K_R - F_I * K_I) + i(F_R * K_I + F_I * K_R), \end{aligned} \quad (1)$$

where $*$ denotes the convolution operation. The complex-domain convolution can also be presented in a matrix format as follows:

$$\begin{bmatrix} \text{Re}(F * K) \\ \text{Im}(F * K) \end{bmatrix} = \begin{bmatrix} F_R & -F_I \\ F_I & F_R \end{bmatrix} * \begin{bmatrix} K_R \\ K_I \end{bmatrix}. \quad (2)$$

2.1.2 Complex-domain activation function

The activation function plays a great role in increasing the nonlinear modeling ability of a neural network. We employed the rectifier linear unit (ReLU) as the activation function and implemented it in the real and imaginary parts independently. Thus, the complex-domain activation function (CReLU) can be expressed as

$$\text{CReLU}(F) = \text{ReLU}(F_R) + i\text{ReLU}(F_I), \quad (3)$$

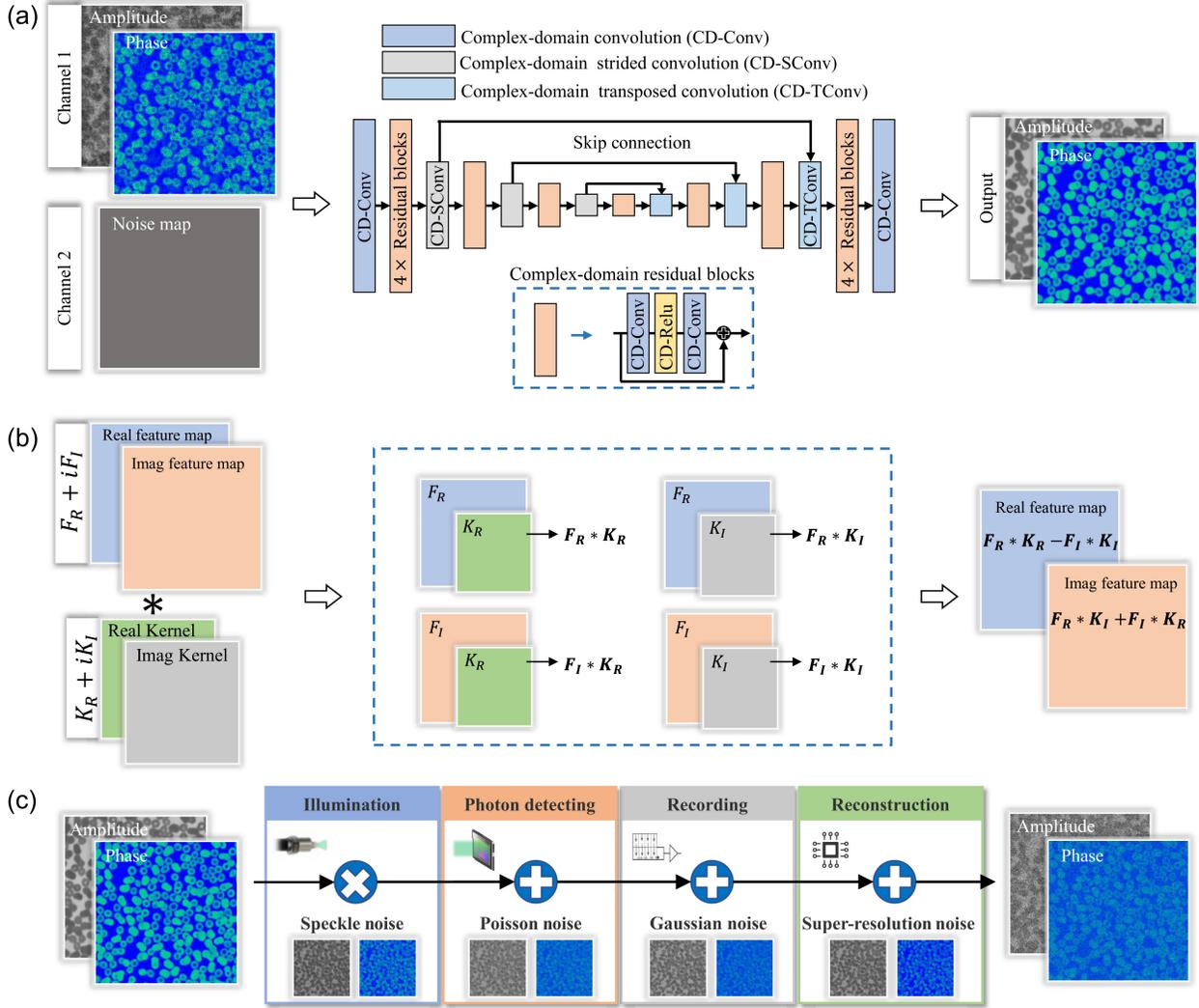


Fig. 1 Architecture of the proposed CI-CDNet. (a) Complex-domain neural network architecture. (b) Complex-domain convolution operation. (c) Multisource noise model for CI.

where the ReLU is

$$\text{ReLU}(F) = \begin{cases} F & \text{if } F \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

2.1.3 Complex-domain weight initialization

The complex value with a mean of zero is employed to implement weight initialization,

$$W = |W|e^{i\theta} = W_R + iW_I, \quad (5)$$

where $|W|$ and θ are the amplitude and phase, respectively. In our implementation, $|W|$ follows Rayleigh distribution and θ follows uniform distribution in the range of $(-\pi, \pi)$. The variance of the complex-domain weight is

$$\text{Var}(W) = \mathbb{E}[WW^*] - (\mathbb{E}[W])^2 = \mathbb{E}[|W|^2] - (\mathbb{E}[W])^2. \quad (6)$$

Because W is symmetrically distributed around 0, thus

$$\text{Var}(W) = \mathbb{E}[|W|^2]. \quad (7)$$

It is hard to compute $\mathbb{E}[|W|^2]$ directly.³⁰ We can introduce an auxiliary variable $\text{Var}(|W|)$, which can be obtained through

$$\text{Var}(|W|) = \mathbb{E}[|W||W|^*] - (\mathbb{E}[|W|])^2 = \mathbb{E}[|W|^2] - (\mathbb{E}[|W|])^2. \quad (8)$$

Putting Eqs. (7) and (8) together, $\text{Var}(|W|)$ can be indicated as $\text{Var}(|W|) = \text{Var}(W) - (\mathbb{E}[|W|])^2$. Thus, the variance of W is expressed as

$$\text{Var}(W) = \text{Var}(|W|) + (\mathbb{E}[|W|])^2. \quad (9)$$

2.2 Multisource Noise Model for Large-Scale CI

In general, the measurement noise is modeled as additive Gaussian noise. Although it has been validated that a CNN trained with synthesis Gaussian noise data has the capacity for removing mixed noise by setting a large noise variance,²⁷ the image details would be sacrificed. To break this limitation, we built a multisource CI noise model to match the real-world noise, as shown in Fig. 1(c). Specifically, we considered the following four noise types.

2.2.1 Gaussian noise

Additive white Gaussian noise models the generalized detector's noise, such as nonuniform illumination noise and thermal noise. We added Gaussian noise in the training data with random noise variance (from 0 to 0.3).

2.2.2 Poisson noise

Poisson noise models the photons' statistical characteristic, which is related to light intensity. It occurs severely in low-light and short-exposure conditions. In order to simulate different

Poisson noise levels, we generated a random multiplicative coefficient 10^α ($\alpha \in [2,3]$) to the complex-domain images. After adding Poisson noise, the images divide back by 10^α .

2.2.3 Speckle noise

Speckle noise usually appears in CI modalities. It is a multiplicative noise that can be modeled by Gaussian distribution. We simulated multiplicative speckle noise with the same variance range as the Gaussian noise.

2.2.4 Superresolution noise

Large-scale CI usually employed superresolution reconstruction techniques, for instance, ptychography imaging synthesizes spatial or Fourier domain to extend the SBP. Although the super-resolution reconstruction does not introduce noise, it would magnify noise and affect its distribution. To model the super-resolution noise, we utilized the bicubic interpolation⁴² to resize the noisy complex-domain wavefront with a superresolution ratio of 2.

We utilized a random shuffle strategy to add the above-mentioned multisource noise in the real and imaginary parts

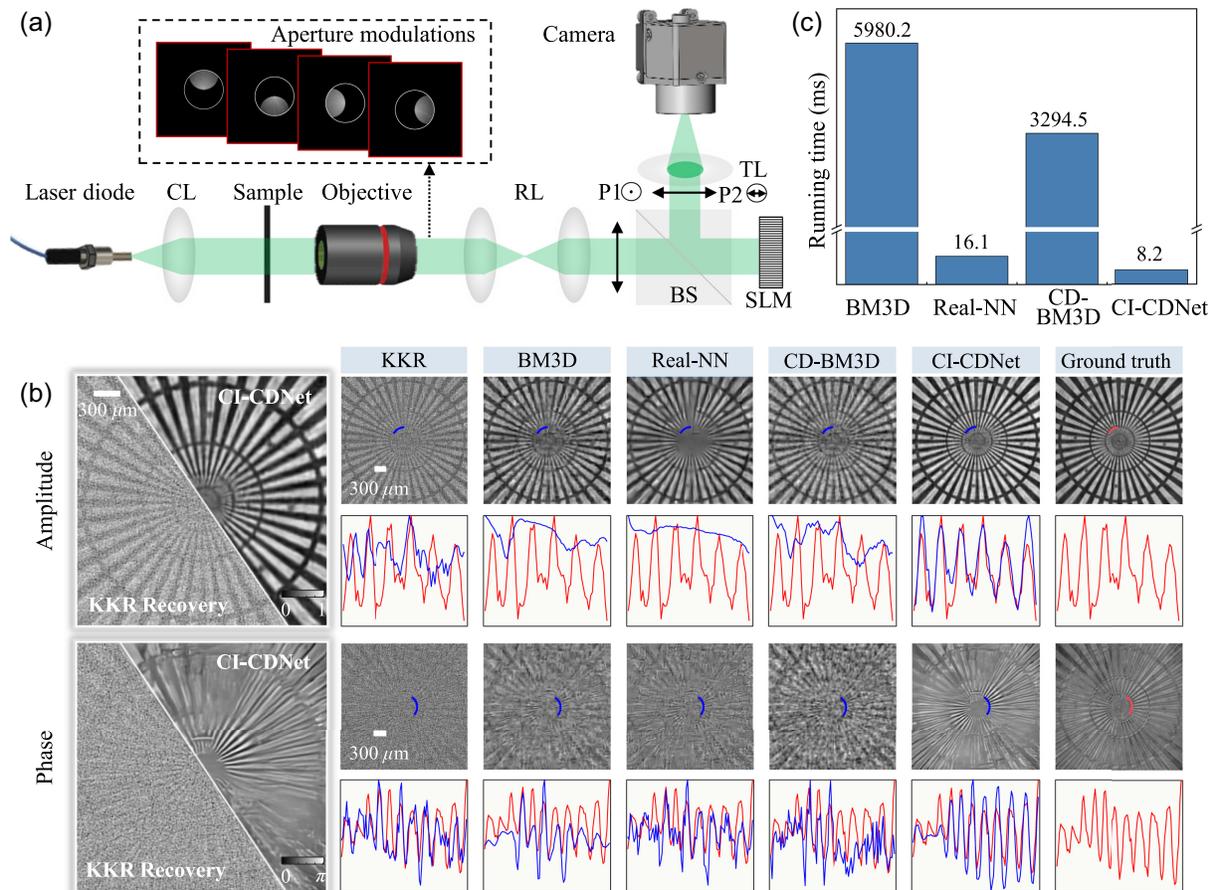


Fig. 2 Experimental setup and resolution test of KKR holography under only 1 ms exposure time. (a) Experimental setup. CL, collimating lens; RL, relay lens; P, polarizer; BS, beam splitter; TL, tube lens; SLM, spatial light modulator. (b) Resolution test results of different enhancing methods using Siemens star under 1 ms exposure time. The blue and red curves are the cross sections of the images, which represent pixel-wise errors. The extremely short exposure time results in a low SNR of KKR direct reconstruction. (c) Running time (ms) of different enhancing methods.

of the complex wavefront. Specifically, the additive Gaussian noise is first added due to its strong generalization for different noise sources. After that, the speckle noise and Poisson noise randomly appeared with the probability of 50%. Finally, we resized the noisy wavefront to simulate the superresolution reconstruction noise.

2.3 Training Details

We employed 10,000 synthetic data sets and added the multi-source noise to train CI-CDNet (Note 1 in the [Supplemental Material](#)). We used L1 loss and Adam optimizer to update parameters with a batch size of 16. The epoch is 400 with an initial learning rate of 1×10^{-5} ; then the learning rate is shrunk by a factor of 0.5 every 150 epochs. The training was implemented in Pytorch 1.8.1 and NVIDIA 2080ti GPU for about 4 days.

3 Results

We applied the proposed CI-CDNet to enhance the reconstructed wavefront and explored its potential for reducing exposure time and data volume. The comparison methods included BM3D,²⁴ complex-domain BM3D (CD-BM3D),²⁵ and conventional real-domain neural network (Real-NN). Real-NN has the same architecture and training process as the CI-CDNet. These comparison methods are state-of-the-art representations of model-driven and data-driven methods. We employed BM3D and Real-NN to denoise the amplitude and phase of the complex-domain wavefront independently. The CD-BM3D and CI-CDNet were used to denoise the complex wavefront directly.

3.1 Kramers–Kronig-Relations Holography

Wavefront reconstruction via KKR is a recent high-SBP and noniterative CI technique,^{3,36–38} which has been used in both two-dimensional holographic imaging^{36,38} and three-dimensional refractive index tomography.^{3,37} KKR combines the real and imaginary parts of a complex function that is analytic in the upper half-plane and requires multiple measurements of different illumination angles³⁷ or aperture modulation³⁸ to satisfy the analyticity. We applied these denoising methods to the KKR reconstructed wavefront, aiming to reduce exposure time and accelerate measurement acquisition.

Our experimental setup is shown in Fig. 2(a). It contained a 532 nm laser diode (Thorlabs DJ532-40) as a light source, an objective (10× Mitutoyo Plan Apo infinity-corrected objective, 0.28 NA), a reflective spatial light modulator (Holoeye LC-R 1080), and a camera (Allied Vision Prosilica GX 6600) with 5.5 μm pixel size. We employed the aperture modulation strategy to satisfy the analyticity of KKR. Specifically, the Fourier plane was relayed outside the objective onto the spatial light modulator (SLM) plane, and the edge of the generated modulation aperture strictly crosses the objective’s pupil center. To obtain the complete Fourier spectrum within the pupil, we implemented four modulations and acquired corresponding intensity measurements under 1 to 1000 ms exposure time (Note 3 in the [Supplemental Material](#)). The reconstruction result of KKR using 1000 ms measurements was used as the ground truth (GT) to quantitatively compare the performance of different techniques.

Figure 2(b) shows the resolution test results of Siemens star under 1 ms exposure time. Due to the short exposure time, the

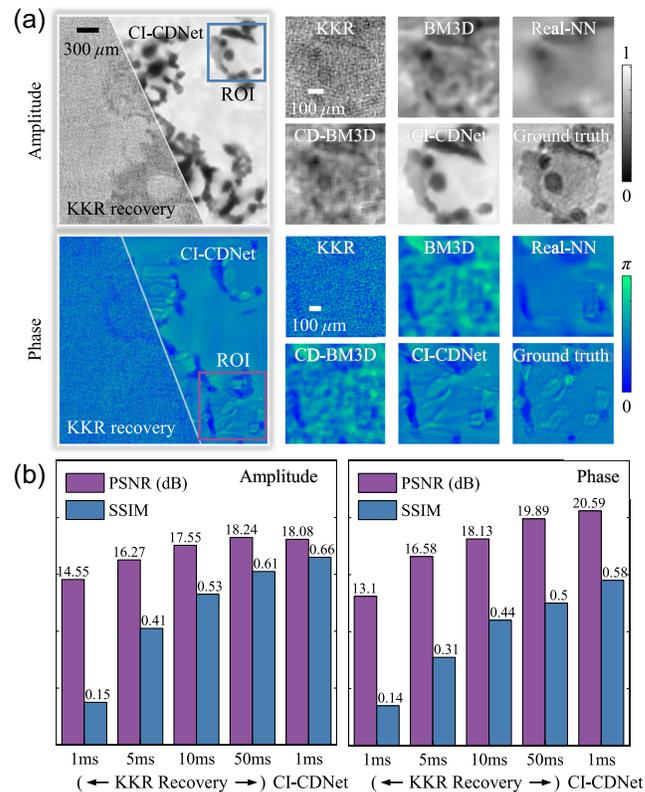


Fig. 3 Quantitative results of KKR holography using a biological sample. (a) Enhancing results of different methods using a papillary thyroid carcinoma slide. (b) Quantitative results of CI-CDNet for reducing exposure time. The result of CI-CDNet under 1 ms exposure time is close to the results of KKR recovery under 50 ms exposure time (more results can be seen in Note 5 in the [Supplemental Material](#)).

results of KKR recovery contained serious background noise and detail loss. Although the conventional denoising algorithms can suppress noise, the resolution was sacrificed (as presented in the cross-sectional curve). In comparison, the proposed CI-CDNet outperformed other methods in both noise suppression and resolution maintenance. Figure 2(c) shows the running time (ms) of different methods. The proposed CI-CDNet had the best running efficiency. Quantitatively, CI-CDNet reduced running time by 2 orders of magnitude compared with the conventional model-based techniques (BM3D and CD-BM3D).

Then, we employed a biological sample to quantitatively explore the performance of CI-CDNet for reducing exposure time. Figure 3(a) shows the results of papillary thyroid carcinoma slide under 1 ms exposure time. Figure 3(b) shows the quantitative results under different exposure time. The evaluation indexes included peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). We can see that the result of CI-CDNet under 1 ms exposure time is close to the result of KKR under 50 ms exposure time (more results can be seen in Note 5 in the [Supplemental Material](#)). Thus, CI-CDNet can reduce more than 1 order of magnitude in exposure time.

3.2 Fourier Ptychographic Microscopy

FPM is a novel technique for wide-field and high-resolution imaging.^{8,9} It extends microscopy’s bandwidth to a billion pixels

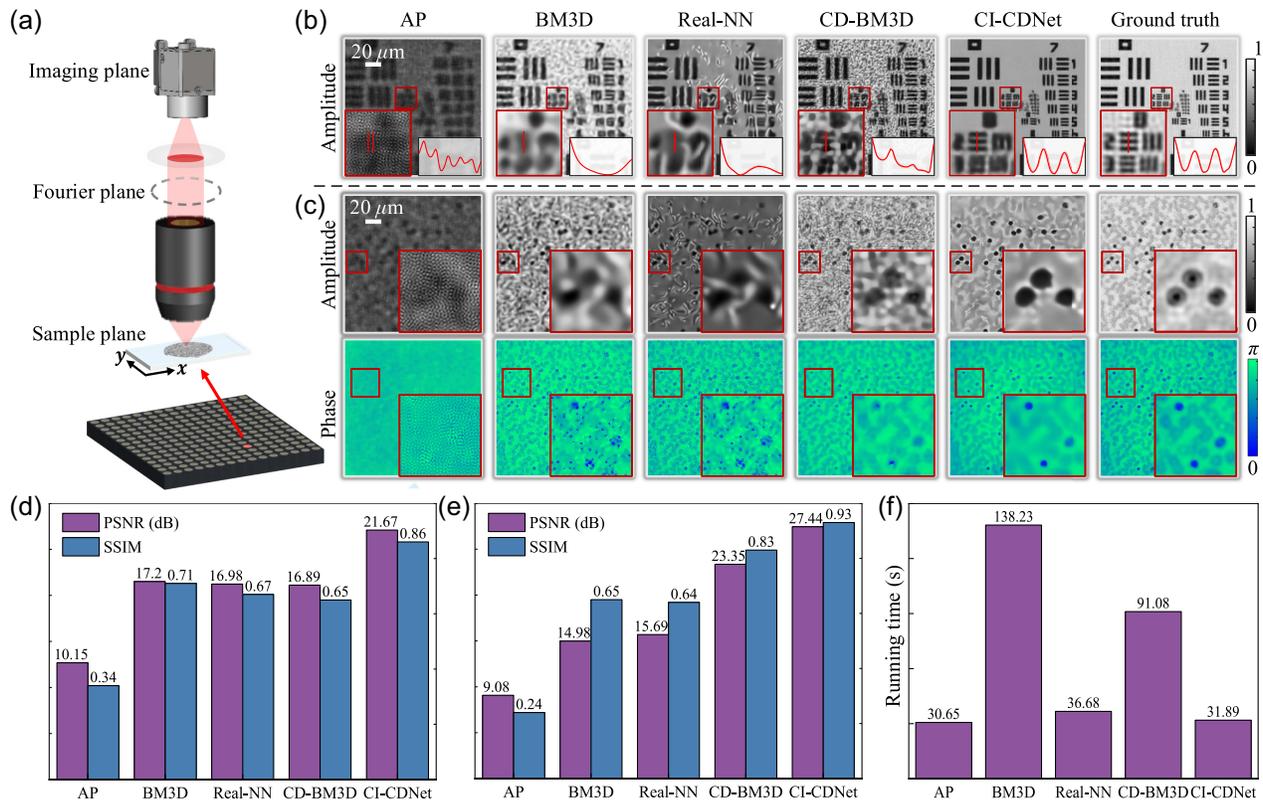


Fig. 4 Results of FPM under only 0.15 ms exposure time. (a) Experimental setup. (b) Resolution test of different enhancing methods using the USAF resolution test chart under 0.15 ms exposure time. (c) Enhancing results of the unstained blood smear under 0.15 ms exposure time. (d), (e) Quantitative results (blood smear) of amplitude and phase, respectively. The proposed CI-CDNet obtains more than 11 dB (amplitude) and 18 dB (phase) improvement on the PSNR index compared with the conventional AP method. (f) Running time (s) of different enhancing methods. AP is the baseline algorithm. Other enhancing methods are regularizers in FPM reconstruction; thus their running time includes the iteration time of data-fidelity term (Note 2 in the [Supplemental Material](#)).

through multiple illumination angles that correspond to different subregions in the Fourier domain. Different from the direct wavefront enhancement in KKR holography, we applied CI-CDNet as a regularizer during the iterative FPM reconstruction^{20,33} (Note 2 in the [Supplemental Material](#)).

Figure 4(a) presents the experimental setup. It contained a 15×15 LED illumination array with 632 nm central wavelength, an objective with 2×0.1 NA, and a camera with $1.85 \mu\text{m}$ pixel size. We captured 225 low-resolution (LR) images from different illumination angles under 0.15, 0.25, and 4 ms exposure time, respectively. We used the alternating projection (AP)⁴³ technique to reconstruct the measurements under 4 ms exposure time as GTs.

Figures 4(b) and 4(c) show the reconstruction results of the USAF resolution test chart (amplitude sample) and unstained blood smear under 0.15 ms exposure time. Figures 4(d) and 4(e) show the quantitative results (blood smear) of amplitude and phase, respectively. We can see that the conventional AP algorithm (baseline) failed due to serious noise and distortion. The regularization methods with real-domain denoising techniques (BM3D and Real-NN) are able to enhance the imaging resolution, which can resolve group 7, element 5 of the USAF target, but the spatial distortion seriously affected the

reconstruction quality (especially the phase images of blood smear). The CD-BM3D outperformed the real-domain denoising techniques, with higher resolution to resolve group 7, element 6 of the USAF target and better phase image quality of the blood smear. However, the high computational complexity and long running time make it unsuitable for rapid large-scale imaging. In comparison, the proposed CI-CDNet obtained the best performance. It can resolve group 8, element 2 of the USAF target and recover clear cell structures for both amplitude and phase images of blood smear. The PSNR and SSIM indexes also validated the advantage of the proposed CI-CDNet. Figure 4(f) presents the running time (s) of different methods. We should note that the running time of these enhancing methods included the iteration time of data-fidelity terms based on AP. Benefiting from the one-step and end-to-end strategies, CI-CDNet is efficient in the iterative reconstruction, which only consumed about a quarter of running time compared with the conventional BM3D method.

3.3 Lensless Coded Ptychography

LCP with a random diffuser has emerged as a low-cost high-SBP technique that can bypass the throughput limit of optical

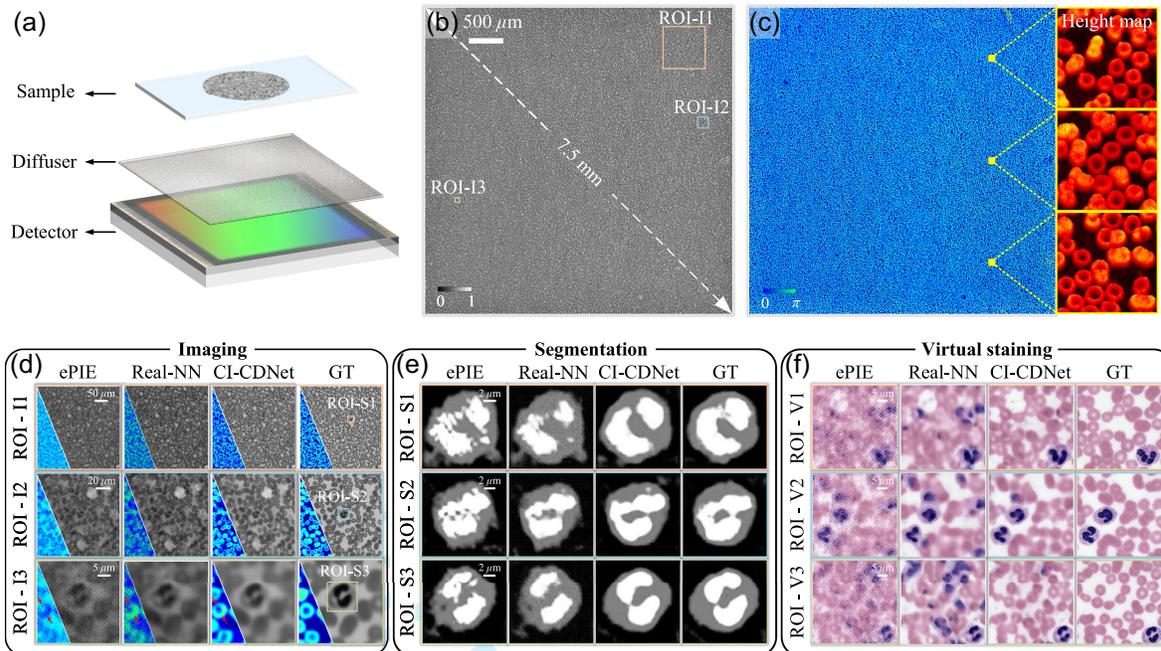


Fig. 5 Results of LCP. (a) Schematic diagram of LCP system. (b), (c) Reconstructed amplitude and phase of the unstained blood smear using CI-CDNet with only 50 captured images. (d) Close-ups of three ROIs. The pseudo-color part is the phase and the gray part shows the amplitude. The GT is the result of ePIE using 900 images. (e) Results of white blood cell segmentation. (f) Results of virtual staining.

systems.^{19,39–41} In LCP, a diffuser is placed between the sample and detector to modulate the wavefront and encode the high-frequency information (Note 3 in the [Supplemental Material](#)). In general, LCP requires nearly thousands of LR measurements to iteratively recover the high-resolution sample and unknown diffuser's profile simultaneously, which makes the data acquisition time-consuming and cumbersome. Thus, we aim at reducing data volume requirements and acquisition time.

Figure 5(a) shows the experiment setup. We applied the glass etching chemicals to a coverslip and coated carbon nanoparticles to produce a random diffuser. It realized micrometer-level phase scattering and subwavelength intensity absorbing. The light source was a fiber-coupled diode with 532 nm wavelength. We used an unstained blood smear as a sample and continuously moved it to $900x - y$ positions. The shift step size is 1 to 3 μm to balance the motion blur and similarity. A detector (Sony IMX226, 1.85 μm pixel size) was used to capture the corresponding intensity diffraction images at a fixed frame rate (30 FPS), and the data collection consumed ~ 30 s. We compared the ePIE algorithm,⁴⁴ Real-NN and CI-CDNet regularization algorithm (Note 2 in the [Supplemental Material](#)) to superresolution reconstruct (4 \times) the sample and the diffuser's profile using only 50 captured images. The BM3D and CD-BM3D methods failed due to their excessive computational complexity and unacceptable long running times. The results of ePIE using 900 images were regarded as the GT. The recovered complex-domain diffuser's profile and sample shift positions are shown in Fig. S4 of Note 3 in the [Supplemental Material](#).

Figures 5(b) and 5(c) show the results of amplitude and phase using CI-CDNet. The reconstructed complex-domain images

have $6144 \text{ pixels} \times 6144 \text{ pixels}$ and a 7.5 mm FOV. Figure 5(d) shows the close-ups of three regions of interest (ROIs). The pseudo-color part is the phase and the gray part is the amplitude. The proposed CI-CDNet can suppress background noise efficiently, providing high-fidelity results for label-free cell observation. Moreover, CI-CDNet can reconstruct the discoid mature erythrocyte, as indicated by the red arrow in the ROI-I3. The quantitative results of Table 1 and visual results in Note 5 in the [Supplemental Material](#) show that the result of CI-CDNet using 50 images is close to the results of ePIE using 500 images. Thus, CI-CDNet can reduce data volume by 1 order of magnitude.

The satisfactory performance of CI-CDNet benefits the subsequent high-level semantic analysis. We demonstrated the high-accuracy white blood cell segmentation^{45,46} and virtual

Table 1 Quantitative results of LCP using different data volumes. The result of CI-CDNet using 50 captured images is close to the results of ePIE using 500 images (more results can be seen in Note 5 in the [Supplemental Material](#)).

Algorithm	Data Volume	Amplitude		Phase	
		PSNR	SSIM	PSNR	SSIM
ePIE	50 images	16.38	0.46	15.14	0.38
	100 images	24.48	0.73	21.67	0.50
	500 images	30.01	0.93	26.84	0.79
CI-CDNet	50 images	30.59	0.89	27.08	0.75

Note: Bold values indicate the best index in each respective column.

staining⁴⁷ (Note 4 in the [Supplemental Material](#)). Figures 5(e) and 5(f) present the segmentation and staining results, respectively. We can see that the results of ePIE contain discontinuous segmentation profiles and incorrect staining. In contrast, the proposed CI-CDNet improved the segmentation and staining accuracy significantly.

4 Conclusion and Discussion

We proposed a novel large-scale CI technique with a complex-domain-enhancing neural network, termed CI-CDNet. CI-CDNet introduced complex-domain operations to the CNN, which can exploit the latent correlations between amplitude and phase. In this way, the proposed technique broke the inherent restriction of the conventional real-domain neural network, realizing cross-field and joint representation of complex wavefront. Furthermore, a multisource noise model of large-scale CI was built to train CI-CDNet. The high-accuracy noise model benefits the network's domain-adaptation ability from synthetic data to real data, improving its performance in various degraded scenes. The data-driven and end-to-end manners brought low computational complexity of CI-CDNet for large-scale CI. We compared CI-CDNet with model-driven methods (BM3D and CD-BM3D) and data-driven methods (Real-NN and dual-channel neural network; see Note 8 in the [Supplemental Material](#)) in a series of large-scale CI modalities, including KKR holography, FPM, and LCP. The results validated its state-of-the-art performance for extremely few data volumes and low exposure time. Specifically, in KKR, CI-CDNet can reduce the exposure time by more than 1 order of magnitude. In FPM, CI-CDNet improved by more than 11 and 18 dB of amplitude and phase respectively on the PSNR index. In LCP, it reduced the data volume by nearly 1 order of magnitude. To conclude, the proposed technique breaks the trade-off among computational complexity, generalization, and reconstruction accuracy. It can be extended for more generalized frameworks and applications in future work.

The noise map of CI-CDNet is an essential parameter for its performance. In our implementation, it relied on heuristic estimation and manual adjustment, which was difficult to estimate accurately for real noise. The recent advanced blind denoising technique⁴⁸ and reinforcement learning technique⁴⁹ are expected to solve the problem and realize noise map estimation and all parameters adjustment automatically during iterations.

The current CI-CDNet requires a preconstructed intermediate step. The two-step processing is unsuitable for computation resource-limited platforms that require real-time imaging. In addition, the performance of CI-CDNet is inseparable from the preconstructed accuracy. End-to-end learning for different modalities using the proposed technique is an effective way to avoid intermediate step. But the generalization would be sacrificed, and the neural network requires to be retrained for different imaging modalities. An alternative solution is combining physics-informed frameworks, such as deep image prior⁵⁰ and deep unfolding⁵¹ techniques. They incorporate the physics model and the built-in smoothness prior of the neural network to optimize the imaging tasks. Nevertheless, their large memory requirement for graphics processing units is a bottleneck for ultralarge-scale imaging.

We believe that the complex-domain neural network is potentially even more broadly transformative for optimizing the whole imaging workflow. Specifically, it can be introduced to the joint optimization of imaging setup and reconstruction,⁵²

for instance, the illumination angle, modulation pattern, imaging distance, or even more generalized physical parameters. In addition, the application scenarios can be also extended, such as multiple dimensions voxel reconstruction and holographic image segmentation and recognition. This may open new insight into complex wavefront representation in various optoelectronics fields.

Data Availability

The neural network and the pretrained model of CI-CDNet are available at github.com/bianlab/CI-CDNet.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61827901, 61991451, 62131003) and the BIT Research and Innovation Promoting Project (Grant No. 2022YCXZ006).

References

1. D. J. Brady et al., "Multiscale gigapixel photography," *Nature* **486**(7403), 386–389 (2012).
2. X. Lin et al., "All-optical machine learning using diffractive deep neural networks," *Science* **361**(6406), 1004–1008 (2018).
3. J. Li et al., "Transport of intensity diffraction tomography with non-interferometric synthetic aperture for three-dimensional label-free microscopy," *Light-Sci. Appl.* **11**(1), 154 (2022).
4. Y. Xue et al., "Single-shot 3D wide-field fluorescence imaging with a computational miniature mesoscope," *Sci. Adv.* **6**(43), eabb7508 (2020).
5. H. Pinkard et al., "Learned adaptive multiphoton illumination microscopy for large-scale immune response imaging," *Nat. Commun.* **12**(1), 1916 (2021).
6. J. Park et al., "Review of bio-optical imaging systems with a high space-bandwidth product," *Adv. Photonics* **3**(4), 044001 (2021).
7. J. Fan et al., "Video-rate imaging of biological dynamics at centimetre scale and micrometre resolution," *Nat. Photonics* **13**(11), 809–816 (2019).
8. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**(9), 739–745 (2013).
9. G. Zheng et al., "Concept, implementations and applications of Fourier ptychography," *Nat. Rev. Phys.* **3**(3), 207–223 (2021).
10. O. Kulce et al., "All-optical information-processing capacity of diffractive surfaces," *Light-Sci. Appl.* **10**(1), 25 (2021).
11. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nat. Photonics* **12**(10), 578–589 (2018).
12. Y. Rivenson et al., "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light-Sci. Appl.* **7**(2), 17141–17141 (2018).
13. Y. Rivenson et al., "Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning," *Nat. Biomed. Eng.* **3**(6), 466–477 (2019).
14. S. Cheng et al., "Single-cell cytometry via multiplexed fluorescence prediction by label-free reflectance microscopy," *Sci. Adv.* **7**(3), eabe0431 (2021).
15. C. Zuo et al., "Deep learning in optical metrology: a review," *Light-Sci. Appl.* **11**(1), 39 (2022).
16. W. Luo et al., "Synthetic aperture-based on-chip microscopy," *Light-Sci. Appl.* **4**(3), e261–e261 (2015).
17. W. Luo et al., "Pixel super-resolution using wavelength scanning," *Light-Sci. Appl.* **5**(4), e16060–e16060 (2016).
18. Y. Gao and L. Cao, "Generalized optimization framework for pixel super-resolution imaging in digital holography," *Opt. Express* **29**(18), 28805–28823 (2021).

19. S. Jiang et al., "Resolution-enhanced parallel coded ptychography for high-throughput optical imaging," *ACS Photonics* **8**(11), 3261–3271 (2021).
20. X. Chang et al., "Plug-and-play pixel super-resolution phase retrieval for digital holography," *Opt. Lett.* **47**(11), 2658–2661 (2022).
21. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006).
22. X. Lan et al., "Efficient belief propagation with learned higher-order Markov random fields," *Lect. Notes Comput. Sci.* **3952**, 269–282 (2006).
23. Y. Weiss and W. T. Freeman, "What makes a good model of natural images?" in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, IEEE, pp. 1–8 (2007).
24. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
25. V. Katkovnik and K. Egiazarian, "Sparse phase imaging based on complex domain nonlocal BM3D techniques," *Digital Signal Process.* **63**, 72–85 (2017).
26. K. Zhang et al., "Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.* **26**(7), 3142–3155 (2017).
27. K. Zhang, W. Zuo, and L. Zhang, "FFDNet: toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.* **27**(9), 4608–4622 (2018).
28. D. P. Reichert and T. Serre, "Neuronal synchrony in complex-valued deep networks," <https://arxiv.org/abs/1312.6115> (2013).
29. G. Shi, M. M. Shanechi, and P. Aarabi, "On the importance of phase in human speech recognition," *IEEE-ACM Trans. Audio Speech Lang. Process.* **14**(5), 1867–1874 (2006).
30. C. Trabelsi et al., "Deep complex networks," <https://arxiv.org/abs/1705.09792> (2017).
31. Y. Gao and L. Cao, "A complex constrained total variation image denoising algorithm with application to phase retrieval," <https://arxiv.org/abs/2109.05496> (2021).
32. S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: fixed-point convergence and applications," *IEEE Trans. Comput. Imaging* **3**(1), 84–98 (2016).
33. X. Chang, L. Bian, and J. Zhang, "Large-scale phase retrieval," *eLight* **1**, 4 (2021).
34. S. Skylaki, O. Hilsenbeck, and T. Schroeder, "Challenges in long-term imaging and quantification of single-cell dynamics," *Nat. Biotechnol.* **34**(11), 1137–1144 (2016).
35. F. Zhang et al., "Physics-based iterative projection complex neural network for phase retrieval in lensless microscopy imaging," in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 10523–10531 (2021).
36. Y. Baek et al., "Kramers–Kronig holographic imaging for high-space-bandwidth product," *Optica* **6**(1), 45–51 (2019).
37. Y. Baek and Y. Park, "Intensity-based holographic imaging via space-domain Kramers–Kronig relations," *Nat. Photonics* **15**(5), 354–360 (2021).
38. C. Shen et al., "Non-iterative complex wave-field reconstruction based on Kramers–Kronig relations," *Photonics Res.* **9**(6), 1003–1012 (2021).
39. S. Jiang et al., "Wide-field, high-resolution lensless on-chip microscopy via near-field blind ptychographic modulation," *Lab Chip* **20**(6), 1058–1065 (2020).
40. S. Jiang et al., "High-throughput digital pathology via a handheld, multiplexed, and AI-powered ptychographic whole slide scanner," *Lab Chip* **22**, 2657–2670 (2022).
41. S. Jiang et al., "Blood-coated sensor for high-throughput ptychographic cytometry on a Blu-ray disc," *ACS Sens.* **7**(4), 1058–1067 (2022).
42. R. E. Carlson and F. N. Fritsch, "Monotone piecewise bicubic interpolation," *SIAM J. Numer. Anal.* **22**(2), 386–400 (1985).
43. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**(15), 2758–2769 (1982).
44. A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy* **109**(10), 1256–1262 (2009).
45. T. Falk et al., "U-Net: deep learning for cell counting, detection, and morphometry," *Nat. Methods* **16**(1), 67–70 (2019).
46. X. Zheng et al., "Fast and robust segmentation of white blood cell images by self-supervised learning," *Micron* **107**, 55–71 (2018).
47. K. de Haan et al., "Deep learning-based transformation of H&E stained tissues into special stains," *Nat. Commun.* **12**(1), 4884 (2021).
48. S. Guo et al., "Toward convolutional blind denoising of real photographs," in *IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)*, pp. 1712–1722 (2019).
49. K. Wei et al., "TFPnP: tuning-free plug-and-play proximal algorithms with applications to inverse imaging problems," *J. Mach. Learn. Res.* **23**(16), 1–48 (2022).
50. F. Wang et al., "Phase imaging with an untrained neural network," *Light-Sci. Appl.* **9**(1), 77 (2020).
51. J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: model-based inspiration of novel deep architectures," <https://arxiv.org/abs/1409.2574> (2014).
52. B. Zhang et al., "End-to-end snapshot compressed super-resolution imaging with deep optics," *Optica* **9**(4), 451–454 (2022).

Biographies of the authors are not available.