




High-resolution lensless holographic microscopy using a physics-aware deep network

Ashwini S. Galande , Vikas Thapa, Aswathy Vijay , and Renu John *

Indian Institute of Technology Hyderabad, Department of Biomedical Engineering,
Medical Optics and Sensors Laboratory, Hyderabad, Telangana, India

ABSTRACT. **Significance:** Lensless digital inline holographic microscopy (LDIHM) is an emerging quantitative phase imaging modality that uses advanced computational methods for phase retrieval from the interference pattern. The existing end-to-end deep networks require a large training dataset with sufficient diversity to achieve high-fidelity hologram reconstruction. To mitigate this data requirement problem, physics-aware deep networks integrate the physics of holography in the loss function to reconstruct complex objects without needing prior training. However, the data fidelity term measures the data consistency with a single low-resolution hologram without any external regularization, which results in a low performance on complex biological data.

Aim: We aim to mitigate the challenges with trained and physics-aware untrained deep networks separately and combine the benefits of both methods for high-resolution phase recovery from a single low-resolution hologram in LDIHM.

Approach: We propose a hybrid deep framework (HDPHysNet) using a plug-and-play method that blends the benefits of trained and untrained deep models for phase recovery in LDIHM. The high-resolution phase is generated by a pre-trained high-definition generative adversarial network (HDGAN) from a single low-resolution hologram. The generated phase is then plugged into the loss function of a physics-aware untrained deep network to regulate the complex object reconstruction process.

Results: Simulation results show that the SSIM of the proposed method is increased by 0.07 over the trained and 0.04 over the untrained deep networks. The average phase-SNR is elevated by 8.2 dB over trained deep models and 9.8 dB over untrained deep networks on the experimental biological cells (cervical cells and red blood cells).

Conclusions: We showed improved performance of the HDPHysNet against the unknown perturbation in the imaging parameters such as the propagation distance, the wavelength of the illuminating source, and the imaging sample compared with the trained network (HDGAN). LDIHM, combined with HDPHysNet, is a portable and technology-driven microscopy best suited for point-of-care cytology applications.

© The Authors. Published by SPIE 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.JBO.29.10.106502](https://doi.org/10.1117/1.JBO.29.10.106502)]

Keywords: lensless holography; physics-aware neural networks; cervical cells; high-resolution

Paper 240081GRR received Mar. 22, 2024; revised Sep. 4, 2024; accepted Sep. 6, 2024; published Oct. 8, 2024.

*Address all correspondence to Renu John, renujohn@bme.iith.ac.in

1 Introduction

Quantitative phase imaging (QPI) has enabled label-free quantitative investigations of biological cells and tissues by measuring their optical, chemical, and mechanical properties.¹ Digital holographic microscopy (DHM), a QPI modality, is an interferometric technique for measuring morphological characteristics along with dynamics.^{2,3} The core idea of DHM is that a coherent light wave, on passing through an object, experiences diffraction and change of phase induced by the object. This information is captured in the interference pattern formed by the superposition of the wave diffracted/scattered by the object and the un-scattered wave at some angle. The goal is to generate a 3D shape of the object by processing the captured hologram. However, the stringent requirements of a coherent source, environmental stability, high numerical aperture (NA) microscope objectives, and device complexity limit its use in low resource settings and point-of-care applications. The development of a low-cost, portable, and technology-driven imaging modality is highly desirable for point-of-care applications. The real-time analysis allows for regular clinical check-ups and early diagnosis of severe diseases such as cancer, anemia, and bacterial infections, which can save many lives in rural areas.

With the recent advances in computational capability, reconstruction algorithms, and faster imaging devices with low cost, the concept of lensless holographic microscopy^{4,5} has emerged as an attractive alternative to the traditional DHM.⁶ Using a partially coherent source in lensless digital inline holographic microscopy (LDIHM) enables portable, low-cost, and speckle-free imaging devices for point-of-care applications.^{7,8} However, twin image artifacts, which are caused by the propagation of the conjugate wavefront and the resolution limit, make the reconstruction challenging with single-shot hologram recording. Hologram reconstruction is formulated as an inverse problem (IP) approach that searches for the reconstructed object, in the solution space, that is most consistent with the captured hologram given some prior knowledge of the object under reconstruction (as shown in Fig. 1).⁹ This is an ill-posed problem due to the availability of multiple solutions and requires regularization [$\rho(u_o)$]. Methods such as Hand-crafted priors,⁹ sparsity,¹⁰ and learned priors^{11,12} are imposed as a constraint along with the data fidelity term in the minimization process. Several optimization methods have been proposed to solve this optimization problem.^{13–16} However, the reconstruction quality is highly dependent on the choice of the suitable handcrafted prior, the optimal tuning parameter, and the heuristic search that may get stuck at local optima. Recent studies have shown interest in combining the alternating projections strategy (phase retrieval methods) with the inverse problem approach for single-shot phase reconstruction.¹⁷ However, convergence on a single low-resolution hologram with a large field-of-view (FOV) is challenging.

Deep learning (DL) methods can effectively solve inverse hologram reconstruction problems with superior performance over conventional methods.^{18–21} Based on the requirement of training data, we provide a taxonomy (as shown in Fig. 1) for the classification of DL methods in hologram reconstruction. In Fig. 1, H represents the forward imaging model, u_o is the object under reconstruction, R_θ represents the deep network with θ as the learnable parameters, and I is the

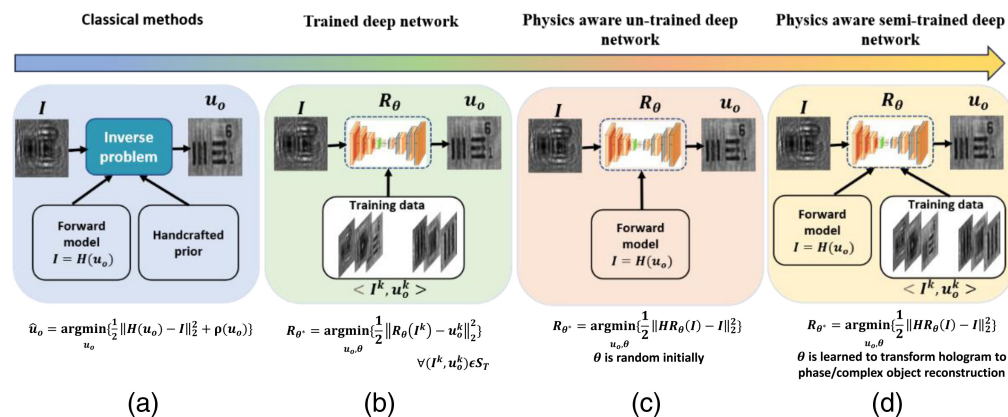


Fig. 1 (a)–(d) Hologram reconstruction based on inverse problem approach: classic versus trained deep models versus un-trained deep models.

captured hologram. $\rho(u_o)$ is the handcrafted regularization/constraint on the object under reconstruction. In the classic inverse problem approach [Fig. 1(a)], the reconstruction is formulated as the optimization problem in which the error between the predicted and captured hologram is minimized while following the constraint on the reconstructed object. However, the handcrafted priors often fail to capture the rich structure of many natural signals due to a lack of discriminative power.

Trained/supervised DL methods to solve inverse problems have demonstrated impressive performance in computational imaging.²² Here, the network R_θ learns the mapping between the input hologram (I^k) and the reconstructed object (u_o) from the training dataset (S_T). Hologram reconstruction based on supervised DL has shown significant improvement in the reconstructed phase from a single hologram.^{18,23} Despite the superior performance of supervised DL methods in phase recovery, they suffer from the drawback of the requirement of a large, object-specific training data pair (I^k, u_o^k), which is very difficult to achieve. Also, these models learn the morphological characteristics of the samples for pixel-to-pixel transformation, which may lead to failure for the patterns that were not provided in the training dataset. These data-driven end-to-end networks do not consider the physics of the hologram formation process (imaging physics).^{18–21}

To alleviate the data requirement problems, recently proposed untrained deep models incorporate the physics of hologram formation in the hologram reconstruction process.^{24–30} The network could learn to reconstruct 3D objects from holographic data by directly capturing patterns and relationships from the input. As shown in Fig. 1(c), the error between the captured and measured holograms is used to update the network parameters (θ). The interplay between the network (R_θ) and the forward model ($H(I)$) learns the network parameters. These models aim to leverage the expressive power of deep networks while ensuring that the reconstructed object adheres to the fundamental laws of wave propagation and interference. However, the process of fitting reconstruction to the single measured hologram results in noisy reconstruction. Explicit regularization methods have been proposed to improve the reconstruction by eliminating noise.^{30,31} However, we observed in our previous work³¹ that the untrained networks perform better for low-complexity signals as they lack the learned knowledge and predictive capabilities that come with training. Hence, the synergy of supervised deep models with the physics of holography is highly desirable for the hologram reconstruction process (we refer to them as semi-trained models).

The recently proposed semi-trained methods can refine results and improve interpretability by leveraging the ability of deep networks to extract meaningful representations from the limited data and constraining the fundamental laws of wave propagation and interference on the reconstruction.^{32,33} Here, θ is learned from the limited dataset and further improved with the physics consistency. Based on the model architecture and type of dataset required for pre-training, we roughly categorized these methods into three groups: unpaired training approach, model-based deep unfolding approach, and plug-and-play (PnP) approach. The unpaired training approach uses the cycle consistency loss and a generative adversarial network (CycleGAN) in which the real distribution of the object is not necessary and hologram and ground truth are irrelevant. The CycleGAN-based reconstruction of holograms is robust against the aberrations present in imaging and can be used for real-time reconstruction.^{34,35} Model-based deep unfolding networks are designed to imitate the behavior of the conventional alternate projection optimization methods. Each stage consists of a trainable deep network that implements one iteration of the alternating projection method. Furthermore, the physics-based constraints enforced in each stage/iteration benefit from the representational capabilities of deep networks as well as the interpretability of traditional iterative algorithms.³⁶ The third approach, the PnP method, makes use of the alternating directions of multipliers method (ADMM) to combine DL and denoiser of a choice into a robust single-shot phase recovery process.^{37,38} ADMM allows for alternating between a data fidelity term and the denoiser, which promotes parallelism to reduce the computational time.

In this paper, we address two different problems of LDIHM reconstruction: the twin-image free phase recovery from a single-shot hologram and improving the resolution, which is limited by the pixel size of the camera. We propose a hybrid approach using the PnP model that combines the phase recovered by the high-definition GANs (HDGANs)³⁹ as *a priori* to the

physics-aware deep networks (UNet).³¹ We refer to our proposed method as HDPhysNet. The flexibility of the PnP models, which use the ADMM for optimization, makes it possible to plug the learned regularization in the loss function along with the data fidelity term without explicit differentiation. The learned prior inherently denoises the twin image and upsamples the reconstructed phase image by $2\times$ or $4\times$ while leveraging the ability of deep networks to extract meaningful representation. The data fidelity term promotes data consistency using the hologram formation model.

Overall, our main contributions are as follows:

- We propose a completely new semi-trained approach (HDPhysNet) that combines learned knowledge and physics consistency using the PnP model to solve the challenges of limited resolution and twin image-free phase reconstruction in single-shot LDIHM.
- We use HDGAN and train it on synthetic low-resolution holograms to extract phase information and upsample the reconstruction by $2\times$.
- We use ADMM to integrate the benefits of HDGAN into the physics-aware untrained deep network (UNet) to refine and improve the results. Furthermore, it maintains more interpretability on unknown samples than the supervised deep methods alone.
- We demonstrate the improved performance of HDPhysNet against the perturbation of imaging parameters such as propagation distance, imaging sample type, and wavelength. We comparatively evaluate the accuracy, efficiency, and generalization capabilities of HDPhysNet over conventional methods, trained DL methods, and physics-aware untrained deep networks in hologram reconstruction.

2 Principles of LDIHM Imaging

The sample under imaging is illuminated by a partially coherent light emitting diode (LED) light source of wavelength 627 nm, butt-coupled to an optical fiber (Model M15L01, Thorlabs, Newton, New Jersey, United States) with a coupling efficiency of 16.66%.⁸ The wave diffracted by the weakly absorbing object is represented as the object wave (u_o), and the un-scattered wave acts as the reference wave (u_r) in the object plane. A large source-to-sample distance ($z_1 \sim 3 - 6$ cm) and a small sample-to-detector distance (usually in the range of mm) result in an FOV of ~ 29 mm². The schematic of the setup is shown in Fig. 2.

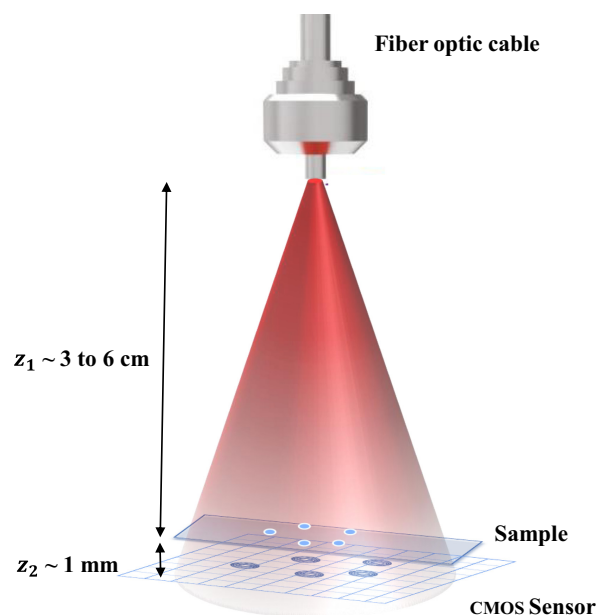


Fig. 2 Schematic of the experimental setup of LDIHM.

The diffracted wavefront in the far field (sensor plane) is given by the Huygens–Fresnel integral as

$$U_o(x, y) = \iint \mathbf{u}_o(\xi, \eta) \exp(ikz)h(x - \xi, y - \eta)d\xi d\eta, \quad (1)$$

where $h(\cdot)$ is the function of free-space propagation of the diffracted object wave $\mathbf{u}_o(\xi, \eta)$ along the optical axis with the wave number $k = 2\pi/\lambda$ (with λ being the wavelength of the propagating wave). The propagation path length, z , is calculated using an autofocusing algorithm.⁴⁰ The reference wave is the recording without a sample, and it is used to normalize the hologram. The interference of the object and reference wave is captured by the complementary metal oxide semiconductor (CMOS) camera with a pixel size of $1.67 \mu\text{m}$ and a resolution of 3840×2784 pixels. The distribution of these two waves at the detector in the plane (x, y) is described as U_o and U_R , respectively. The hologram intensity is given as

$$\mathbf{I} = |U_o + U_R|^2 = |U_o|^2 + |U_R|^2 + U_o^*U_R + U_oU_R^*. \quad (2)$$

After normalization, the effect of $|U_R|^2$ is removed, and the non-linearity due to $|U_o|^2$ is modeled as an error e ; then, the mapping between the object field and hologram is represented as

$$\mathbf{I} = U_o^*U_R + U_oU_R^* + e = H(\mathbf{u}_o), \quad (3)$$

where $H(\cdot)$ represents a mapping function that relates the object field \mathbf{u}_o to the hologram \mathbf{I} .

The second step of holography is reconstructing the original object from the captured interference pattern, which is done by backpropagation of the hologram to the object plane. The propagation to the object plane is explained by the Fresnel–Kirchhoff integral,⁴¹ described as

$$\mathbf{u}_o(\xi, \eta) = \frac{i}{\lambda} \iint U_R^*(x, y)\mathbf{I}(x, y) \frac{\exp(-ik\rho)}{\rho} dx dy, \quad (4)$$

where $\rho = \sqrt{(x - \xi)^2 + (y - \eta)^2 + z^2}$ is the distance between the hologram plane and the object plane and z is the perpendicular distance between the hologram and the reconstruction plane. The incident plane wave at the object location is obtained as “1” by selecting the optical axis along the propagation of a digital hologram. Then, Eq. (4) is rewritten as

$$\mathbf{u}_o(\xi, \eta) = \frac{i}{\lambda} \iint \mathbf{I}(x, y) \frac{\exp(-ik\rho)}{\rho} dx dy. \quad (5)$$

The reconstructed complex object $\hat{\mathbf{u}}_o(\xi, \eta)$ at a distance z from the image sensor using angular spectrum wave propagation⁴¹ is represented as

$$\hat{\mathbf{u}}_{\text{obj}}(\xi, \eta) = F^{-1} \left[F(\mathbf{I}) \cdot \exp\left(-\frac{2\pi iz}{\lambda}\right) \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \right], \quad (6)$$

where F and F^{-1} are Fourier transform and inverse Fourier transform, respectively, and f_x and f_y are the spatial frequency coordinates. Amplitude and phase are extracted from the reconstructed complex object.

The resolution of the LDIHM depends on several factors: the NA, sample-to-source distance, the degree of the optical coherence (temporal and spatial), and the pixel size of the camera.^{7,42} The temporal coherence length ΔL_c of the source is calculated as $\Delta L_c = \left(\frac{2 \ln 2}{\pi}\right) \cdot \frac{\lambda^2}{n \Delta \lambda}$, where n is the refractive index of the medium, λ is the wavelength, and Δ is the spectral width of the illuminating source. The spatial coherence diameter at the sample plane is proportional to $D_{\text{coh}} = \lambda \cdot z_1 / D$, where D is the aperture size at the illumination plane.

The resolution imposed by the temporal coherence of illumination is given as $\theta_{\text{max}} \leq \arccos\left(\frac{z_2}{z_2 + \Delta L_c}\right)$. The resolution imposed by spatial coherence is given by $\theta_{\text{max}} \leq \arctan\left(\frac{0.61\lambda z_1 / D}{z_2}\right)$. Based on these calculations, for a $z_2 \sim 1 \text{ mm}$ and spectral bandwidth $\Delta \lambda \sim 20 \text{ nm}$, the resolution (Δx) is close to the pixel size $\sim 1.67 \mu\text{m}$.

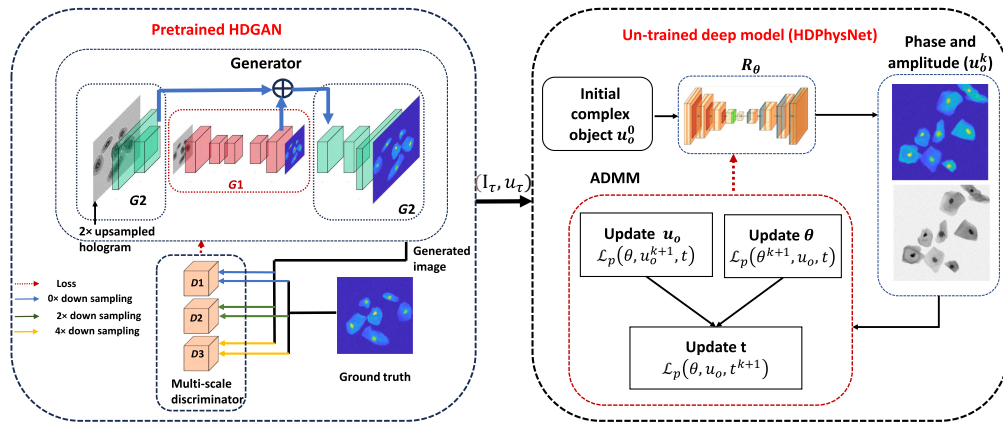


Fig. 3 Schematic of the proposed method: a semi-trained deep framework for high-resolution hologram reconstruction.

3 Methods

The low-resolution hologram is recorded under the proposed LDIHM setup, and the high-resolution complex object information (amplitude and phase) is obtained using the proposed semi-trained framework. The proposed semi-trained framework consists of two components: the pre-trained HDGAN and the HDPhysNet. The HDGAN is pre-trained on the simulated dataset to reconstruct a high-definition phase image from a low-resolution diffraction pattern. The HDGAN³⁹ can increase the image resolution by 2× or 4× without losing fine details. The HDPhysNet is an extension of our earlier work DIP-RED³¹ in which different denoisers are plug-and-play into physics-aware untrained networks. Here, instead of denoisers, the high-resolution phase generated by the HDGAN is plugged as the learned prior knowledge to further fine-tune the reconstruction through the physics-aware untrained model. This approach combines the benefits of both pre-training and physics-aware task-specific training to enhance the model's performance. Each component of Fig. 3 is described in Secs. 3.2 and 3.3.

3.2 High-Definition Generative Adversarial Networks

The deep network is trained to learn the mapping of the input hologram (I) to the reconstructed phase (\emptyset) of the object (u_o). Traditional GAN models often struggle to generate high-resolution images due to the limitations of the network architecture and training process.⁴³ HDGANs³⁹ incorporate various techniques and modifications to enhance the generation of high-resolution images. We adopted HDGAN to generate high-resolution phase images from low-resolution holograms. With the help of a multiscale generator and discriminator, as shown in Fig. 3, the HDGAN can synthesize phase images of size 1024×1024 , which are more visually appealing than those generated by earlier GANs.

To generate high-resolution images, the generator (G) is divided into two sub-networks: $G = \{G_1, G_2\}$. We first train a residual network G_1 , which operates at a resolution of 512×512 . Then, another residual network G_2 , which is a local enhancer appended to G_1 , and the two networks are trained jointly on high-resolution images.³⁹ Specifically, the feature map of the generator, G_1 , is added to the feature map of the local enhancer (G_2), which ensures the propagation of local information to the global generator. To differentiate the high-resolution real and synthesized images, multiscale discriminators (D_k) are proposed to operate at different image scales. In this study, we used three discriminators of the same architecture, i.e., patchGAN and image scales are obtained by down-sampling them by a factor of 2 and 4. In addition to the multiscale discriminator, the feature matching loss is also calculated at multiple layers of the discriminator.

Using a two-level generator effectively aggregates global and local information for image generation. Also, the multi-scale discriminators are trained at three different scales, which encourages the generator to produce finer details with global consistency. Training for the HDGAN requires a longer time, which is however a one-time effort, leading to high-speed

reconstruction after training. HDGAN can produce high-resolution phase images (ϕ_τ) from low-resolution holograms, which is highly desirable for lensless holography.

3.3 Physics-Aware Semi-trained Deep Framework

Deep image prior (DIP) is the first untrained network implemented; it incorporates a physical model (forward propagation) of the image formation process in the loss function.⁴⁴ Here, the loss is calculated as the difference between the captured hologram and the generated hologram (generated using the forward model). The untrained network uses ADMM¹⁶ to optimize the network parameters. ADMM is a variable splitting approach that can incorporate various types of regularization or constraints to promote desirable properties in the reconstructed hologram. ADMM allows for distributed optimization of large problems by dividing them into smaller and more manageable chunks. It is the best way to plug the learned prior into a loss function as a regularization. It is worth noting that the hologram used is a high-resolution hologram (I_τ) that is obtained from HDGAN for further processing. The detailed description of the proposed HDPhysNet, shown in Fig. 3, is as follows.

3.3.1 Initialization

The first input, (\mathbf{u}_o^0), is obtained by backpropagating the high-resolution hologram (I_τ) to the object plane using the propagation kernel. The phase generated by HDGAN, ϕ_τ , is combined with a constant amplitude of \mathbf{u}_o^0 to form a trained complex object, \mathbf{u}_τ . Other parameters such as network parameters, θ , are set randomly; the Lagrange multipliers vector, \mathbf{t} , is set to 0; and $k = 0$.

3.3.2 Loss function

Network R_θ learns to fine-tune the initial reconstruction (\mathbf{u}_o^0) using physics consistency and the learned prior (\mathbf{u}_τ). The loss function using ADMM for the proposed HDPhysNet is defined as

$$\mathcal{L}_{\text{HDPhysNet}}(\theta, \mathbf{u}_o, \mathbf{t}) = \arg \min_{\mathbf{u}_o, \theta} \left\{ \frac{1}{2} \|HR_\theta(\mathbf{u}_o) - I_\tau\|_2^2 + \frac{\alpha}{2} \mathbf{u}_o^T (\mathbf{u}_o - \mathbf{u}_\tau) + \frac{\beta}{2} \|\mathbf{u}_o - R_\theta(\mathbf{u}_o) - \mathbf{t}\|_2^2 \right\}, \quad (7)$$

where function $H(\cdot)$ maps complex object ($\hat{\mathbf{u}}_o$) to the hologram I' . α and β are regularization parameters. In Eq. (7), the first term is the data fidelity term, which enforces physics consistency by estimating the error between the captured hologram and the generated hologram. The second term is regularization in which the output of the HDGAN is used to regulate the reconstructed object. The gradient of regularization is $\mathbf{u}_o - \mathbf{u}_\tau$. The last term is proximity regularization, which forces network output $R_\theta(\hat{\mathbf{u}}_o)$ to be close to $\mathbf{u}_o - \mathbf{t}$.

3.3.3 Optimization

ADMM is the variable splitting approach that updates the parameters θ , \mathbf{t} , and \mathbf{u}_o , sequentially. θ is updated by fixing \mathbf{t} , and \mathbf{u}_o is given as

$$\mathcal{L}_\theta(\theta^{k+1}, \mathbf{u}_o, \mathbf{t}) = \arg \min_{\mathbf{u}_o, \theta} \left\{ \frac{1}{2} \|HR_\theta(\mathbf{u}_o) - I_\tau\|_2^2 + \frac{\beta}{2} \|\mathbf{u}_o - R_\theta(\mathbf{u}_o) - \mathbf{t}\|_2^2 \right\}. \quad (8)$$

The effect of the second term in Eq. (7) is negligible while updating θ as \mathbf{u}_o is kept constant. The next input to the network \mathbf{u}_o^{k+1} is updated using the steepest descent method, which takes the gradient of the second and the third term in Eq. (7) as θ and \mathbf{t} are constants.

$$\mathcal{L}_{\mathbf{u}_o}(\theta, \mathbf{u}_o^{k+1}, \mathbf{t}) = \mathbf{u}_o^k - c[\alpha(\mathbf{u}_o^k - \mathbf{u}_\tau) + \beta(\mathbf{u}_o^{k+1} - R_\theta(\mathbf{u}_o^k) - \mathbf{t})]. \quad (9)$$

Here, the residual image is calculated to penalize the object estimate based on learned regularization and proximity regularization. c is the step size that ensures the steepest descent. Finally, the Lagrange multiplier vector \mathbf{t} is updated by keeping θ and \mathbf{u}_o constant:

$$\mathcal{L}_t(\theta, \mathbf{u}_o, \mathbf{t}^{k+1}) = \mathbf{t}^k - (\mathbf{u}_o^{k+1} + R_\theta(\mathbf{u}_o^k)). \quad (10)$$

The use of the variable splitting method allows for dealing with the error term and the regularization separately; hence, Eqs. (8) and (9) can be executed in parallel to reduce the convergence time.

3.3.4 Convergence/stopping criteria

As discussed, ADMM updates three variables, θ , \mathbf{t} , and \mathbf{u}_o , sequentially. Hence, we need a stopping criterion to avoid overfitting. We adapted the stopping criteria from Ref. 45, given as $\max\{\epsilon_1, \epsilon_2, \epsilon_3\} \leq \text{tol}/3$,

where $\epsilon_1 = \|\mathcal{L}_\theta(\theta^{k+1}, \mathbf{u}_o, \mathbf{t}) - \mathcal{L}_\theta(\theta^k, \mathbf{u}_o, \mathbf{t})\|_2/\sqrt{n}$, $\epsilon_2 = \|\mathcal{L}_{u_o}(\theta, \mathbf{u}_o^{k+1}, \mathbf{t}) - \mathcal{L}_{u_o}(\theta, \mathbf{u}_o^k, \mathbf{t})\|_2/\sqrt{n}$, and $\epsilon_3 = \|\mathcal{L}_t(\theta, \mathbf{u}_o, \mathbf{t}^{k+1}) - \mathcal{L}_t(\theta, \mathbf{u}_o, \mathbf{t}^k)\|_2/\sqrt{n}$ for $\mathbf{u}_o \in \mathbb{R}^n$. The tolerance level is decided based on the phase SNR value of the reconstructed phase image. We observed that $\text{tol} \approx 10^{-3}$ is sufficient for the convergence of HDPhysNet to avoid overfitting of the interference-related noise.

4 Materials

4.1 Sample Preparation

Biological samples were collected from the hospital using approved standard protocols accepted through the Institutional Internal Review Committee (IITH/IEC/2024/02/12). Cervical cells of healthy patients were extracted and suspended in a Carnoy's fixative (methanol: glacial acetic acid = 3:1) solution without any staining procedure. The buffer solution was then dropped on the microscopic slide and imaged under LDIHM. Similarly, the blood sample of a healthy patient was collected and centrifuged to separate red blood cells (RBCs) from the blood. RBCs were diluted with distilled water and dropped on the slide for imaging.

4.2 Dataset

The cervical cell dataset was obtained from the ISBI 2014 dataset,⁴⁶ which was designed for the purpose of a segmentation challenge contest (segment overlapping cells). The diffraction patterns are generated with different possible propagation distances used during experiments, e.g., $z = \{0.8, 0.9, 1, 1.1, 1.2, 1.3\}$ mm and the wavelength of the illuminating source, $\lambda = 627$ nm. A phase value of 4 radians was added during the simulation process. The generated hologram and corresponding phase image were stored in different folders with the same name. Out of 2048 images, 1800 images were used for training, 148 for validation, and 100 for testing. It is important to note that the imaging parameters (propagation distance and wavelength) for test images are not the same as those used in the training dataset.

4.3 Network Architecture

We used two different networks: HDGAN for the trained model and UNet for the untrained model. HDGAN has a generator and discriminator networks. Note that we have two generators $G = \{G_1, G_2\}$. This network contains two stride-2 convolutions, several residual blocks, and two 1/2-strided convolutions.³⁹ The multiscale discriminator has three discriminators working on different image scales. Each discriminator has the same architecture. It has two convolutional layers that work on patches; hence it is called patchGAN. The UNet architecture is adopted from our previous work.³¹

We implemented our untrained model with UNet architecture using the PyTorch framework on a GPU workstation with two NVIDIA GeForce GTX 1080ti 11 GB graphics cards. The HDGAN took 1 day and 18 h for training on 1750 images of 512×512 pixels with 200 epochs. The inference time of trained HDGAN for a hologram of 512×512 pixels is ~ 2 s. The reconstruction of a complex object by UNet takes ~ 3 min with 700 epochs on simulation data and ~ 5.3 min with 1000 epochs on experimental holograms. The dataset preparation and data plotting are done in MATLAB R2022a.

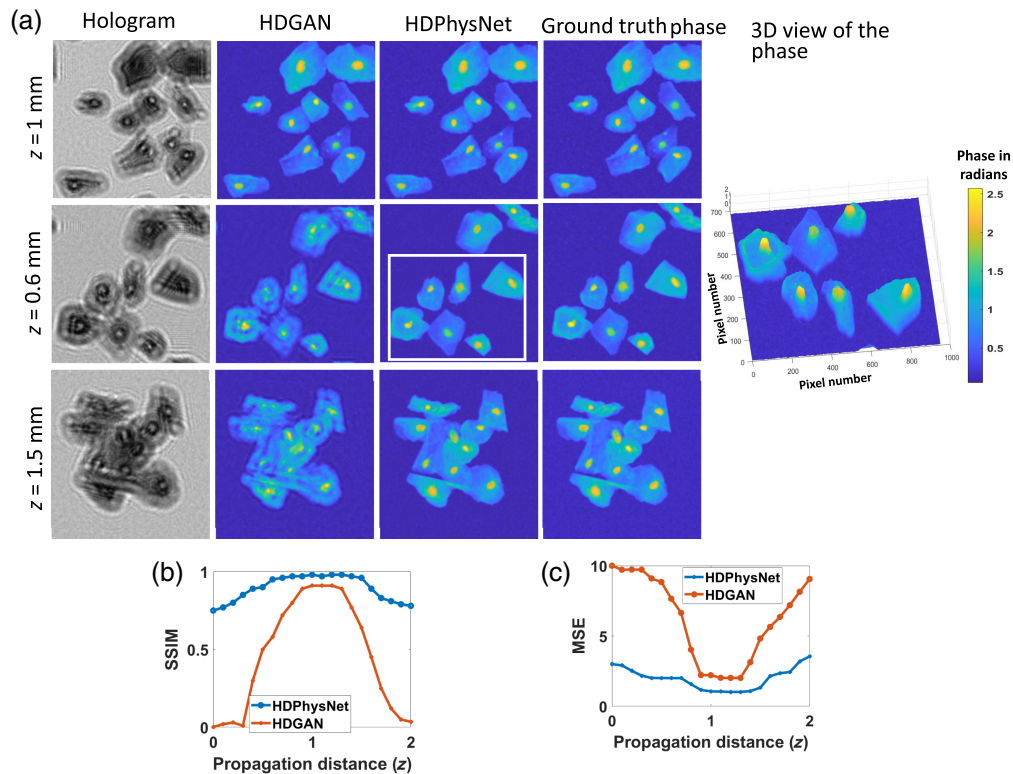


Fig. 4 Simulation results: (a) phase reconstruction by HDGAN and HDPhysNet by varying simulation parameters such as propagation distance (Δz in mm). (b) SSIM plot (c) MSE plot.

5 Result and Discussion

Generative networks can learn to generate high-definition hologram reconstructions by learning intricate details, complex patterns, and textures and extracting meaningful representations from the data. We used HDGAN³⁹ to generate a high-resolution phase from the low-resolution input hologram. The network is trained on the computer-generated holograms. The trained networks are not aware of the imaging physics, which results in low interpretability of unknown samples. The HDPhysNet combined the capabilities of the trained network and the physics of holography to improve the performance on unknown samples and change in imaging parameters (propagation distance and wavelength). To demonstrate this claim, the HDGAN is trained on the cervical cell dataset, which is simulated on experimental parameters, as given in Sec. 4.2. The inference is performed on the samples by varying the propagation distance and the illumination wavelength. It can be observed from Fig. 4(a) that the performance of HDGAN degrades for propagation distances beyond the range given during the training, but HDPhysNet shows consistent performance against the change of these parameters. The first row of Fig. 4(a) indicates that the negligible refinement by HDPhysNet is required on the results of HDGAN as the propagation distance is seen during training. The second and third rows show that the HDGAN could retrieve the structural features from the diffraction pattern with an unseen propagation distance. This initial feature extraction is beneficial for further reconstruction process but needs refinement. HDPhysNet uses this phase as *a priori* knowledge and plugs it with the loss function of the physics-aware network to further fine-tune the reconstruction.

The average structural similarity index metric (SSIM) and mean squared error (MSE) are calculated on 100 test images by HDGAN (trained)³⁹ and HDPhysNet (semi-trained). Phase images with SSIM ~ 1 and very low MSE indicate the highest structural similarity with the ground truth phase image with less error. Figures 4(b) and 4(c) show the SSIM and MSE of HDGAN and HDPhysNet with respect to the propagation distance, respectively. Here, for simplicity, we plotted the SSIM and MSE of only 21 images with varying propagation distances. These graphs indicate that the width of the peak performance is increased by the HDPhysNet

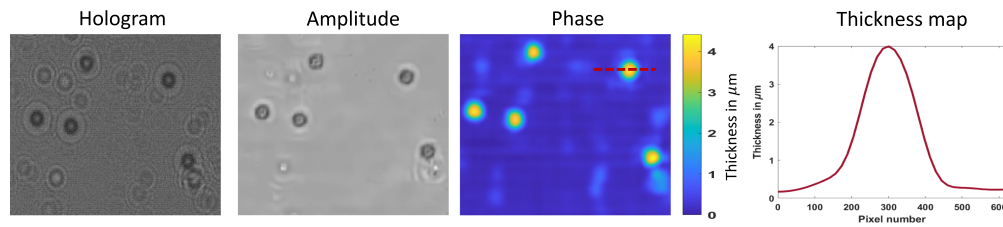


Fig. 5 Phase characterization and thickness profile of microbeads ($4 \mu\text{m}$).

over HDGAN because of the physics-based constraints that guide the learning process, making it capable of reconstructing defocused holograms. The results show that the SSIM is increased by 0.07 over HDGAN and 0.04 over DIP-RED. The MSE is decreased by 1.7 over HDGAN and 1.2 over DIP-RED. The proposed HDPhysNet leverages the general representations learned during pre-training using a limited dataset while fine-tuning the reconstruction using physical consistency.

It is worth noting that the GAN considers the phase image to be an intensity (real) image; hence, scaling (phase to intensity) and re-scaling (intensity to phase) are required at the beginning and end of HDGAN. Hence, we used another UNet architecture to handle complex data (amplitude and phase) and used the generated phase as the regularization for better feature extraction and early convergence. The phase characterization of HDPhysNet is performed using the polystyrene microbeads of size $4 \mu\text{m}$ with a refractive index of 1.68 and a refractive index of the medium (distilled water) of 1.33. The phase is directly proportional to the thickness (h) of the sample and is calculated as $h = \frac{p \cdot \lambda}{2\pi(n_1 - n_2)}$ where p is the estimated phase, λ is the wavelength of the illuminating source, and n_1 and n_2 are the refractive indices of the sample and surrounding media, respectively. The scaling factor is identified by mapping the phase of the microbeads to the thickness, as shown in Fig. 5. The average thickness calculated on beads is $4 \pm 0.05 \mu\text{m}$ with an error of $0.05 \mu\text{m}$.

The performance of HDPhysNet is observed on experimental hologram images under the proposed LDIHM. The large FOV (3840×2784) experimental holograms of the pap smear sample are captured under our LDIHM setup. Then, the region of interest is cropped and resized to the size 512×512 from the large FOV hologram. The $2\times$ up-sampled phase and hologram are obtained from pre-trained HDGAN. Then, the untrained model uses this information to reconstruct the complex object. The reconstructed phase and amplitude of size 1024×1024 by HDPhysNet are shown in Fig. 6. A 3D view of the phase and the phase profile of the cell are shown in the last column of Fig. 6.

The reconstructed phase is compared with different existing single-shot hologram reconstruction methods proposed in the literature. The angular spectrum wave propagation (ASM) is the basis for numerical simulation and reconstruction of the hologram, and it is used in many hologram reconstruction methods.⁴¹ We compared the reconstruction by HDPhysNet with

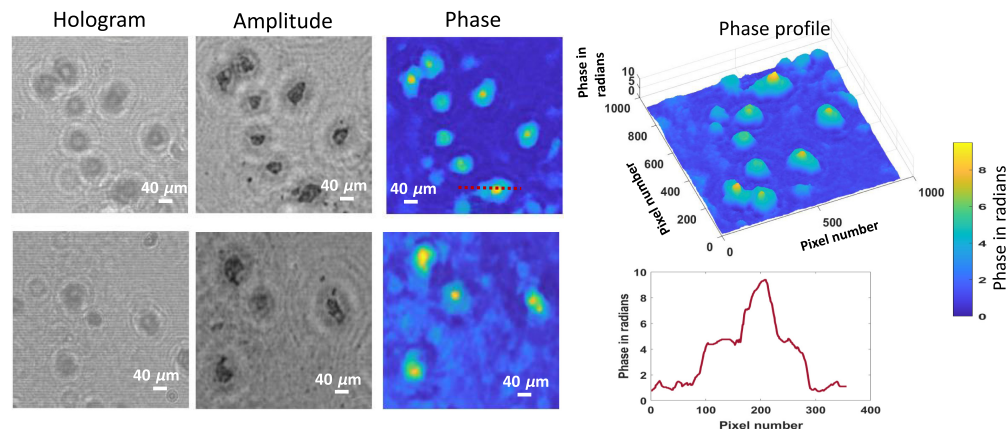


Fig. 6 Experimental cervical cells reconstruction; amplitude and phase reconstruction by the proposed method.

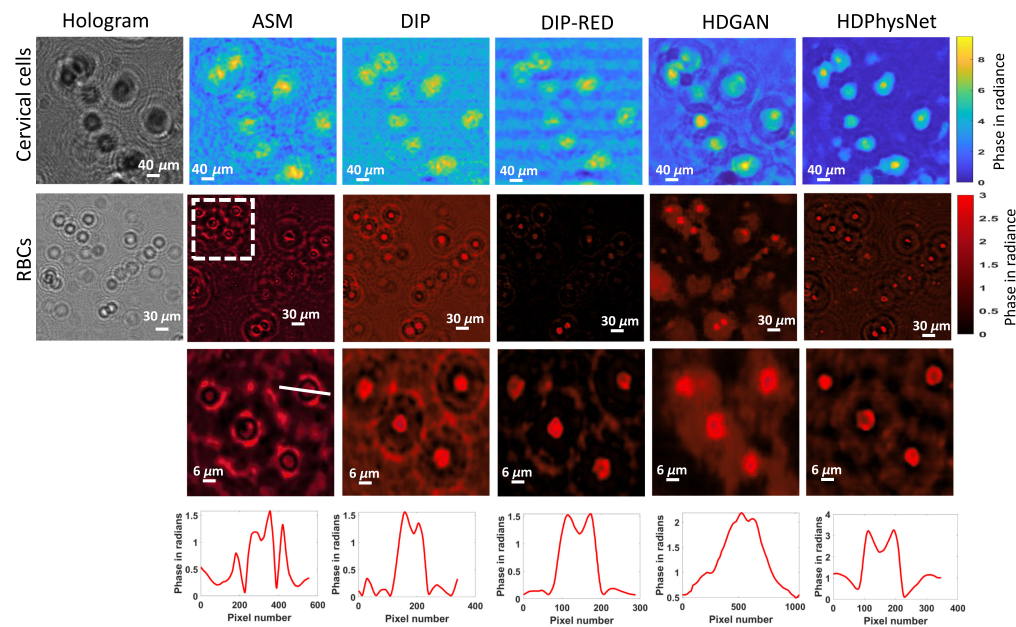


Fig. 7 Phase reconstruction on experimental holograms of biological samples: a comparative study of phase recovery by ASM,⁴¹ DIP,²⁵ DIP-RED,³¹ HDGAN,³⁹ and the proposed HDPhysNet. Color bars on the right represent phase in radians. We changed the color map of the RBC phase for better visibility.

trained (HDGAN) and untrained methods (DIP) separately. We also compared results with our previous work DIP-RED³¹ in which DIP is powered with regularization by denoising to remove interference-related noise and twin images. The denoiser used here is a total variation denoiser. It can be observed from Fig. 7 that DIP-RED has better performance compared with ASM, DIP, and HDGAN. The interpolation used in ASM, DIP, and DIP-RED to increase the visibility introduces some unknown artifacts. This problem is overcome using HDGAN in the proposed method, which upsamples phase and hologram by $2\times$ without losing finer details. Here, the $2\times$ upsampled hologram by the HDGAN is used for reconstruction by each method.

The interpretability of the HDPhysNet on unknown samples is checked on RBCs, which were not part of the training process. It can be observed from Fig. 7 that the results of the HDGAN are biased for RBCs due to the limited interpretability of HDGAN on unknown samples. A cytoplasm-like boundary is generated around the RBCs, which is refined by the HDPhysNet. However, the phase recovered by HDGAN is used in the loss function (of HDPhysNet) to obtain the next input to the network in each iteration. Hence, the phase recovered by HDGAN influences the final reconstruction. This can be improved by adding more data into the training dataset with different imaging samples and varying experimental parameters. The phase SNR (p-SNR) computes the ratio of phase values (P) of the signal and noise, in decibels, as $p - \text{SNR} = 10 \log_{10} \left[\frac{\max(P) - \mu}{\sigma} \right]$ where μ is the mean and σ is the covariance. The comparative analysis of the average phase-SNR by each phase recovery method is given in Table 1. It is observed that the P-SNR is increased by 8.2 dB over HDGAN and 9.8 dB over DIP-RED on the experimental data.

Table 1 Phase SNR (p-SNR) calculated on reconstructed phase images.

Sample	z_2 (mm)	ASM	DIP	DIP-RED	HDGAN	HDPhysNet
Simulations	1	43.7	58	58.3	61.2	63.6
Cervical cells	1.3	15.5	24	25.4	32	41.3
RBCs	1.05	12.7	21.8	28.3	25	38

It is worth noting that the sample-to-detector distance for each experiment was ~ 1 mm. While carrying out reconstructions, our reconstruction algorithm employs an autofocus algorithm⁴⁰ to calculate the exact propagation distance for the best image quality. Hence, within this range of propagation distance, i.e., 1 to 1.5 mm, the performance of our algorithm is very robust. Simulation data showing the performance of the algorithm versus propagation distance (0 mm to 2 mm) are illustrated in Fig. 4. The experimental results showing reconstructions of two different cell types (RBC and cervical cells) for varying propagation distances are illustrated in Fig. 7. Also, to prove the improved performance of the proposed HDPhysNet against the wavelength of the illuminating source, the detailed analysis is demonstrated in the [Supplementary Material](#). However, to prove the robustness of the HDPhysNet against unknown perturbations in imaging parameters and imaging samples, analysis is needed on a large number of samples.

We used two different deep networks in the hologram reconstruction process. The existing semi-trained approaches use a single network that is trained on a limited dataset along with physics consistency.^{33,38} The ability of HDGAN to generate high-definition phase images with a limited dataset is appealing over other networks. However, HDGAN has limitations in handling complex data types (amplitude and phase). To address this limitation, we implemented a physics-aware untrained UNet to reconstruct complex objects and used the phase generated by HDGAN as a learned prior. The use of ADMM for untrained network optimization allowed us to use the phase generated by HDGAN efficiently in the loss function without explicit differentiation of regularization. HDPhysNet combines the strengths of a trained deep model, such as its ability to learn complex patterns, with the physics-aware untrained network to give better results, which is difficult to achieve by either DL models (trained or untrained) alone. This can help overcome challenges such as limited hologram training data and the complexity of the holographic reconstruction process.

LDIHM in combination with the proposed HDPhysNet can overcome the limits of the current cytology procedure at point-of-care with the following advantages:

- Label-free, single-shot holography with a large FOV results in the scanning of complex data in a short period.
- HDPhysNet is not sample-specific and, hence, can be successfully generalized to reconstruct unknown biological samples.
- HDPhysNet is stable against the perturbation of the imaging parameters and hence can overcome the stringent requirement of environmental stability.
- Portable and technology-driven microscopy, fast scanning, and early disease diagnosis will revolutionize healthcare in rural areas.

6 Conclusion

DL-based hologram reconstruction models are object-specific without any guarantee of performance on unknown practical data. The proposed HDPhysNet for hologram reconstruction is a hybrid approach that combines both DL techniques and the physics of holography. The HDGAN used for phase retrieval increases the resolution without losing finer details and inherently denoises the reconstruction for interference-related noise. The use of ADMM allows us to integrate the benefits of HDGAN into the physics-aware untrained deep network (UNet) to refine and improve the results. The simulation results show that the SSIM of the HDPhysNet is increased by 0.07 over HDGAN (trained) and 0.04 over the untrained method (DIP-RED). The phase-SNR is increased by 8.2 dB over HDGAN and 9.8 dB over DIP-RED on the experimental data. The performance of the HDPhysNet is independent of the experimental parameters and imaging sample, and therefore, it is best suited for point-of-care cytology applications. The inference time of the HDPhysNet can be reduced drastically by adding a small amount of experimental ground-truth data while training. We strongly believe that the HDPhysNet can be easily adapted to other phase imaging modalities by replacing the mapping function (i.e., the forward imaging model).

Disclosures

The authors declare no conflicts of interest.

Code and Data Availability

Code and data used in this paper can be obtained from the corresponding author upon reasonable request.

Acknowledgments

We wish to acknowledge the Kamineni Hospital authorities for providing us with the required data and data handling protocols for performing experiments.

References

1. K. Lee et al., “Quantitative phase imaging techniques for the study of cell pathophysiology: from principles to applications,” *Sensors* **13**(4), 4170–4191 (2013).
2. H. P. R. Gurram et al., “Digital holographic microscopy for quantitative and label-free oral cytology evaluation,” *Opt. Eng.* **59**(2), 024105 (2020).
3. V. P. Pandiyan and R. John, “Optofluidic bioimaging platform for quantitative phase imaging of lab on a chip devices using digital holographic microscopy,” *Appl. Opt.* **55**(3), A54–A59 (2016).
4. D. Gabor, “A new microscopic principle,” *Nature* **161**(4098), 777–778 (1948).
5. A. Ozcan and E. McLeod, “Lensless imaging and sensing,” *Ann. Rev. Biomed. Eng.* **18**, 77–102 (2016).
6. I. Moon and B. Javidi, “3-D visualization and identification of biological microorganisms using partially temporal incoherent light in-line computational holographic imaging,” *IEEE Trans. Med. Imaging* **27**(12), 1782–1790 (2008).
7. A. Greenbaum et al., “Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy,” *Nat. Methods* **9**(9), 889–895 (2012).
8. H. P. R. Gurram et al., “Nanometric depth phase imaging using low-cost on-chip lensless inline holographic microscopy,” *Opt. Eng.* **59**(10), 104105 (2020).
9. C. Fournier et al., “Inverse problem approaches for digital hologram reconstruction,” *Proc. SPIE* **8043**, 80430S (2011).
10. W. Zhang et al., “Twin image-free holography: a compressive sensing approach,” *Phys. Rev. Lett.* **121**, 093902 (2018).
11. Y. Shuai et al., “Digital inline holographic reconstruction with learned sparsifying transform,” *Opt. Commun.* **498**, 127220 (2021).
12. Z. Kai et al., “Learning deep CNN denoiser prior for image restoration,” in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, pp. 3929–3938 (2017).
13. A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imaging Sci.* **2**, 183–202 (2009).
14. J. M. Bioucas-Dias and M. A. T. Figueiredo, “A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration,” *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
15. P. Bao et al., “Lensless phase microscopy using phase retrieval with multiple illumination wavelengths,” *Appl. Opt.* **51**, 5486–5494 (2012).
16. B. Wahlberg et al., “An ADMM algorithm for a class of total variation regularized estimation problems,” *IFAC Proc. Vol.* **45**(16), 83–88 (2007).
17. A. Galande et al., “Quantitative phase imaging of biological cells using lensless inline holographic microscopy through sparsity-assisted iterative phase retrieval algorithm,” *J. Appl. Phys.* **132**(24), 243102 (2022).
18. K. Wang et al., “On the use of deep learning for phase recovery,” arXiv:2308.00942 (2023).
19. G. Ongie et al., “Deep learning techniques for inverse problems in imaging,” *IEEE J. Sel. Areas Inf. Theory* **1**(1), 39–56 (2020).
20. A. Lucas et al., “Using deep neural networks for inverse problems in imaging: beyond analytical methods,” *IEEE Signal Process. Mag.* **35**(1), 20–36 (2018).
21. J. Scarlett et al., “Theoretical perspectives on deep learning methods in inverse problems,” *IEEE J. Sel. Areas Inf. Theory* **3**(3), 433–453 (2022).
22. O. Gregory et al., “Deep learning techniques for inverse problems in imaging,” *IEEE J. Sel. Areas Inf. Theory* **1**, 39–56 (2020).
23. R. Yair et al., “Phase recovery and holographic image reconstruction using deep learning in neural networks,” *Light Sci. Appl.* **7**(2), 17141 (2018).
24. Q. Adnan et al., “Untrained neural network priors for inverse imaging problems: a survey,” *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(5), 6511–6536 (2022).
25. H. Li et al., “Deep DIH: single-shot digital in-line holography reconstruction by deep learning,” *IEEE Access* **8**, 202648–202659 (2020).
26. W. Fei et al., “Phase imaging with an untrained neural network,” *Light Sci. Appl.* **9**(1), 77 (2020).
27. S. Kumar, “Phase retrieval with physics informed zero-shot network,” *Opt. Lett.* **46**, 5942–5945 (2021).

28. Y. Yao et al., “AutoPhaseNN: unsupervised physics-aware deep learning of 3D nanoscale Bragg coherent diffraction imaging,” *NPJ Comput. Mater.* **8**, 124 (2022).
29. X. Chen et al., “DH-GAN: a physics-driven untrained generative adversarial network for holographic imaging” *Opt. Express* **31**(6), 10114–10135 (2023).
30. F. Niknam et al., “Holographic optical field recovery using a regularized untrained deep decoder network,” *Sci. Rep.* **11**(1), 10903 (2021).
31. A. Galande et al., “Untrained deep network powered with explicit denoiser for phase recovery in inline holography,” *Appl. Phys. Lett.* **122**(13), 133701 (2023).
32. G. Volpe et al., “Roadmap on deep learning for microscopy” ArXiv (2013).
33. L. Huang et al., “Self-supervised learning of hologram reconstruction using physics consistency,” *Nature Mach. Intell.* **5**(8), 895–907 (2023).
34. D. Yin et al., “Digital holographic reconstruction based on deep learning framework with unpaired data,” *IEEE Photonics J.* **12**(2), 1–12 (2019).
35. Y. Zhang et al., “PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets” *Opt. Express* **29**(13), 19593–19604 (2021).
36. Y. Yang et al., “HIONet: deep priors based deep unfolded network for phase retrieval,” *Digit. Signal Process.* **132**, 103797 (2022).
37. S. Venkatakrishnan et al., “Plug-and-play priors for model based reconstruction,” in *IEEE Glob. Conf. Signal Inf. Process.*, pp. 945–948 (2013).
38. U. S. Kamilov et al., “Plug-and-play methods for integrating physical and learned models in computational imaging: theory, algorithms, and applications,” *IEEE Signal Process. Mag.* **40**(1), 85–97 (2023).
39. T. C. Wang et al., “High-resolution image synthesis and semantic manipulation with conditional GANs,” in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, pp. 8798–8807 (2018).
40. O. Mudanyali et al., “Detection of waterborne parasites using field-portable and cost-effective lensfree microscopy,” *Lab Chip* **10**, 2419–2423 (2010).
41. T. Lатышевская and H.-W. Fink, “Practical algorithms for simulation and reconstruction of digital in-line holograms,” *Appl. Opt.* **54**, 2424–2434 (2015).
42. A. Greenbaum and A. Ozcan, “Maskless imaging of dense samples using pixel super-resolution based multi-height lensfree on-chip microscopy,” *Opt. Express* **20**, 3129–3143 (2012).
43. P. Isola et al., “Image-to-image translation with conditional adversarial networks,” in *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)* (2017).
44. U. Dmitry et al., “Deep image prior,” in *Proc. IEEE CVPR*, pp. 9446–9454 (2018).
45. S. H. Chan et al., “Plug-and-play ADMM for image restoration: fixed-point convergence and applications,” *IEEE Trans. Comput. Imaging* **3**(1), 84–98 (2016).
46. Z. Lu et al., “Evaluation of three algorithms for the segmentation of overlapping cervical cells,” *IEEE J. Biomed. Health Inform.* **21**(2), 441–450 (2015).

Ashwini S. Galande is a senior AI scientist at GE Healthcare, India. She received her PhD in biomedical engineering from IIT Hyderabad, India. She is working on the physics-aware deep networks for medical imaging. Her main research area is the interface of medical imaging physics, mathematics, and deep learning techniques.

Vikas Thapa is a PhD research scholar in the Department of Biomedical Engineering at Indian Institute of Technology (IIT), Hyderabad, India. He received his master’s degree from UPES, Dehradun, in 2018. His PhD research work is dedicated to deep learning based microscopic image analysis. He has hands-on experience with various biomedical devices and instruments. He is also actively participating in international and national conferences.

Aswathy Vijay received her MSc degree in physics from Cochin University of Science and Technology (CUSAT), Cochin, India. Currently, she is a PhD student in Medical Optics and Sensors (MOSLab) group at IIT Hyderabad. Her research focuses on digital holographic microscopy, lensless imaging, and optical instrumentation.

Renu John is a professor in the Department of Biomedical Engineering at IIT, Hyderabad, India. He received his PhD from IIT Delhi in 2006 and received his postdoc from Duke University, Durham, North Carolina, United States, and the University of Illinois Urbana-Champaign, Champaign, Illinois, United States, from 2006 to 2010. His research covers several key areas in biomedical optics, such as quantitative phase imaging and optical coherence imaging.