Research Article | Vol. 30, No. 24 / 21 Nov 2022 / Optics Express 43398
Optics EXPRESS

Physics-informed neural network for phase imaging based on transport of intensity equation

XIAOFENG WU,[1] ZILING WU,[1,2] SIBI CHAKRAVARTHY SHANMUGAVEL,[1] HANG Z. YU,[3] AND YUNHUI ZHU[1,*]

[1]Department of Electrical & Computer Engineering, 1185 Perry St., Virginia Tech, Blacksburg, VA 24061, USA
[2]Current address: Singapore-MIT Alliance for Research and Technology (SMART) Centre, 1 Create Way, Singapore 117543, Singapore
[3]Department of Materials Science and Engineering, 445 Old Turner St., Virginia Tech, Blacksburg, VA 24061, USA
*yunhuiz@vt.edu

Abstract: Non-interferometric quantitative phase imaging based on Transport of Intensity Equation (TIE) has been widely used in bio-medical imaging. However, analytic TIE phase retrieval is prone to low-spatial frequency noise amplification, which is caused by the illposedness of inversion at the origin of the spectrum. There are also retrieval ambiguities resulting from the lack of sensitivity to the curl component of the Poynting vector occurring with strong absorption. Here, we establish a physics-informed neural network (PINN) to address these issues, by integrating the forward and inverse physics models into a cascaded deep neural network. We demonstrate that the proposed PINN is efficiently trained using a small set of sample data, enabling the conversion of noise-corrupted 2-shot TIE phase retrievals to high quality phase images under partially coherent LED illumination. The efficacy of the proposed approach is demonstrated by both simulation using a standard image database and experiment using human buccal epitehlial cells. In particular, high image quality (SSIM = 0.919) is achieved experimentally using a reduced size of labeled data (140 image pairs). We discuss the robustness of the proposed approach against insufficient training data, and demonstrate that the parallel architecture of PINN is efficient for transfer learning.

1. Introduction

In recent years, quantitative phase imaging has been widely applied in bio-imaging [1], drug screening [2], object localization [3], and security scanning [4] as a label-free imaging modality. The wide-spread interest stems from the fact that transparent objects with little or no absorption can impose strong phase delay to the illuminating beam, resulting in significantly enhanced imaging contrast [5]. While historically many early phase imaging techniques have been developed based on the interferometric configuration with light sources of high temporal and spatial coherence [6–8], non–interferometric techniques have been developed using partially coherent illumination over the years [9]. This makes quantitative phase imaging applicable using a wide range of light sources, such as light-emitting diode [10], X-ray [11], and neutron sources [12]. Among the various phase retrieval approaches including contrast transfer function (CTF) [13], structured illumination [14] and speckle-based imaging [15], transport of intensity equation (TIE) retrieves phase profiles using a through-focus intensity profile stack under partially coherent illumination [9,16–22]. Owing to its experimental simplicity, TIE has been easily integrated into a variety of optical microscopes [17], X-ray imaging setups [20], and electron imaging microscopes [23].

#462844 https://doi.org/10.1364/OE.462844
Journal © 2022 Received 3 May 2022; revised 23 Sep 2022; accepted 26 Sep 2022; published 10 Nov 2022

Optics EXPRESS

However, there has been a constant struggle to improve the robustness and accuracy of TIE-based phase retrieval [24–28]. For example, low spatial frequency noise, or haze noise, is commonly found in TIE phase imaging due to the ill-posedness at the origin of the spectral transfer function [16,27]. In addition, TIE phase retrieval relies on the curl-free assumption, which causes ambiguities for absorptive phase objects [25]. To remove the artifacts and suppress the cloud-like low-frequency noise, multiple through-focus measurements [22,26] have been proposed but at the expense of increased experimental complexity involving dozens of repeated image acquisitions at precisely controlled translation distances. Computationally, the phase retrieval results can be improved by regularization methods such as Tikhonov [29–31], iteratively regularized Gauss-Newton (IRGN) [32], total variation (TV) [33–35], fields of experts (FoEs) [36], and block-matching and 3D filtering (BM3D) [37]. Despite demonstrated success for different imaging implementations, it has been reported that such hand-crafted regularization terms may not always accurately represent the features of the objects being imaged [38]. The performance is also sensitive to the imaging noise level [39–41], let alone the high computational cost of the multi-objective optimization involved in regularization.

Recent developments in deep learning techniques have demonstrated great potential for ill-posed image reconstruction [42–46], including low-dose tomographic reconstruction, computational optics, under-sampled magnetic resonance imaging (MRI), etc. Typically, these problems are solved in an end-to-end manner, where a deep neural network (DNN) with millions of parameters is optimized using a large data set. Learned object statistics are used as the prior regularization that transfers future corrupted reconstructions into high-fidelity estimates [47–52]. Such end-to-end DNN approach [53] is purely data-driven which requires large-scale well-annotated data sets to enable efficient supervised learning. In addition, the lack of physics knowledge of the imaging system [54,55] may not guarantee the reconstruction results to be physically accurate [56]. As a result, it is difficult to inform the DNN to optimally correct for the bias embedded in the forward model [57,58], often leading to over-fitting problems. DNN also acts as a low-pass filter due to the inverse-square power spectral density (PSD) law [59], which suppresses the high frequency components of phase images [49]. This insufficiency is addressed by some recent studies that fused the physics-aware forward operator modeling different optical systems trained with deep neural network for holographic image reconstruction [60] and phase retrieval [61]. Apart from supervised learning, the untrained network incorporated with physical knowledge which requires no training data set has been explored in phase imaging [62] and phase unwrapping [63].

To leverage the benefits of both physics knowledge and data-driven approaches, here we integrate the Fresnel forward operator and TIE inverse model into a cascaded convolutional neural network (also known as algorithm unrolling) for phase retrieval.While including prior knowledge from the physics model is not guaranteed to provide benefits to the inversion problem, especially when modeling is difficult and inaccurate [64], the well-established forward and inverse models of H and H_{inv} for the TIE phase retrieval problem here can indeed provide useful physics information that regularizes the data training. The physics-informed neural network (PINN) prevents over-fitting by unrolling a physics model-based data consistency layer in parallel with an iterative regularization term in each nonlinear reconstruction block. It should be noted that each training block corresponds to one stage in iterative optimization scheme, thus the optimal results can be obtained. Following a similar philosophy, cascaded neural networks have demonstrated successes in non-phase imaging processes, such as MRI reconstructions [65], mask-based lensless imaging [66], tomographic reconstructions [54], and 3D holographic particle imaging [56]. The selection of regularization model requires residual structure to capture enough spatial information but the fusion of large-scale deep neural networks in iterative block will aggravate the training burden and even result in over-fitting. To this end, we first train a pruned Unet++ [67,68] for regularization; through nested and dense skip connections, Unet++ reduces the semantic gap between the encoder and decoder sub-networks, and thereby enhances the training efficiency.

Optics EXPRESS

The proposed PINN framework is demonstrated for TIE phase retrieval via both simulation and experiment, achieving high fidelity performance in removing low-frequency noise and correcting reconstruction artifacts. This framework also demonstrates robustness against insufficient training data. Moreover, the proposed PINN framework can be conveniently used for phase imaging of similar objects in different imaging schemes via transfer learning, by updating the physics models in the data consistency layer.

The rest of the paper is organized as follows: we introduce the TIE-based phase imaging theory and formulate the PINN framework in Section 2. We demonstrate phase retrieval using PINN, and compare the simulation and experiment results quantitatively with end-to-end DNNs in Section 3. In addition, we explore the performance of phase retrieval with reduced training data and discuss the benefits of applying the PINN in transfer learning in Section 3. Discussion and conclusion are given in Section 4 and Section 5 respectively.

2. Method

2.1. Theory of TIE phase imaging and absorptive phase imaging

Phase objects can be detected by measuring the induced diffraction pattern. In the near field, diffraction is governed by Fresnel propagation. Under the paraxial approximation, TIE [69,70] provides a quantitative relationship between the diffraction intensity differential along the propagation direction and the lateral phase profile:

$$-k\frac{\partial I(x,y;z)}{\partial z} = \nabla_\perp\cdot(I(x,y;z)\nabla_\perp\varphi(x,y;z)), \tag{1}$$

where x and y are the lateral coordinates in the plane perpendicular to the optical axis z, I denotes the intensity profile, φ is the phase profile on the lateral plane, and $\frac{\partial I}{\partial z}$ is the intensity derivative along the axis of propagation. $k = \frac{2\pi}{\lambda}$ is the wave number of the optical wave, λ is the wavelength, and $\nabla_\perp = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y}$ is the two dimensional gradient operator on the x-y plane. For the simple case of a pure phase object under uniform illumination I_0, the in-focus intensity profile $I(x,y;z=0) = I_0$. Equation (1) is solved using Fourier transform, which gives

$$\varphi = \mathcal{F}^{-1}\left\{\frac{1}{4\pi^2(u^2+v^2)}\mathcal{F}\left\{\frac{-k}{I_0}\frac{\partial I(x,y;z)}{\partial z}\right\}\right\}, \tag{2}$$

where u, v are spectral coordinates in the Fourier domain. $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the Fourier transform and inverse Fourier transform operators, respectively. Figure 1 shows a typical experiment setup for TIE phase imaging. In the experiment, the intensity derivative $\partial I/\partial z$ at $z = 0$ can be conveniently approximated by the finite difference between the 2-shot measurements of $I(x,y;d)$ and $I(x,y;0)$ (abbreviated as I_d and I_0 below) as,

$$\left.\frac{\partial I(x,y;z)}{\partial z}\right|_{z=0} \approx \frac{I_d - I_0}{d}. \tag{3}$$

Combining with Eq. (2), we obtain the 2-shot TIE phase retrieval $\varphi_0 = H_{inv}(I_0, I_d)$.

For a general absorptive phase object, I_0 is not a constant and the Poynting vector can be decomposed into the circulation and divergence terms by $I_0\nabla_\perp\varphi = \nabla_\perp \times C + \nabla_\perp D$. Inserting into Eq. (1), it is found that only the divergence term contributes to the intensity differential and can be solved by $D = -k\nabla_\perp^{-2}\partial I(x,y;z)/\partial z$. An approximate phase retrieval [25] can be obtained based solely on the divergence term via

$$\varphi = \mathcal{F}^{-1}\left\{\frac{1}{4\pi^2(u^2+v^2)}\mathcal{F}\left\{-k\nabla_\perp\cdot\left(\frac{1}{I_0}\nabla_\perp D\right)\right\}\right\}, \tag{4}$$

which is used as the inverse operator H^*_{inv} of 2-shot TIE phase retrieval in absorptive case.

Optics EXPRESS

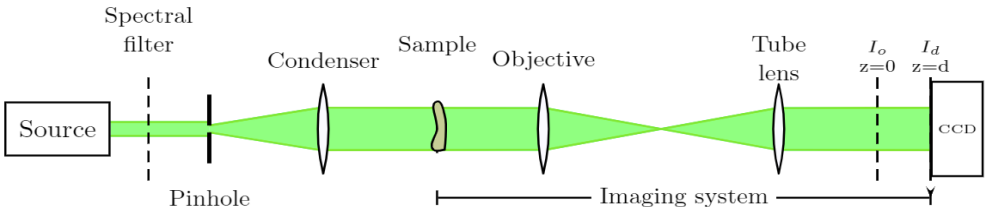

Fig. 1. Experimental setup.

In both cases, when $u^2 + v^2 \rightarrow 0$, inversion by Eq. (2) and Eq. (4) suffers from low frequency noise amplification around the singularity point. Regularization such as Tikhonov can be used to avoid the division-by-zero instability, but it usually suppresses the low spatial frequency components at the same time. On the other hand, finite differential approximation via Eq. (3) imposes a trade-off between low signal-to-noise ratio (SNR) associated with small d and non-linearity associated with large finite distance. These problems have been addressed by increasing the number of measurements in the through-focus intensity stack [22,26,39,40,71–73] but with increased data acquisition burden. Here, we seek to solve the problem via a physics informed learning approach, which uses the simple 2-shot experimental acquisition process.

2.2. *Proposed TIE phase retrieval approach via physics informed learning*

Phase retrieval quality can be improved by the prior knowledge of the phase objects such as edge sparsity, non-negativity, etc. The non-linear optimization process involves the minimization of the objective function $J(\varphi)$

$$\hat{\varphi} = \arg\min_{\varphi} J(\varphi) \equiv \arg\min_{\varphi} \frac{1}{2} \parallel H(\varphi) - I_0 \parallel_2^2 + \gamma R(\varphi), \tag{5}$$

where $\parallel H(\varphi) - I_0 \parallel_2^2$ is the data fidelity term, $R(\varphi)$ is the regularizer, and parameter γ controls the weight balance between them. For TIE-based phase retrieval, we formulate the forward operator H to obtain the defocused diffraction intensity I_d, by first using the Fresnel diffraction integral to obtain the optical field in the near field at the propagated distance $d = z$ [74], and then taking the square of the complex field to yield the defocused intensity I_d:

$$I_d = \left| \sqrt{I_0} e^{i\varphi} * \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} \right|^2, \tag{6}$$

where $*$ denotes convolution. The regularized optimization in Eq. (5) can be solved via the iterative gradient descent algorithm. Starting from the initial phase retrieval $\varphi_0 = H_{inv}(I_0, I_d)$ obtained by the 2-shot TIE inversion process in Eq. (2), the phase profile φ_i in each of the following ith iteration is updated via

$$\varphi_i = \varphi_{i-1} - \alpha \nabla_{\varphi} J(\varphi_{i-1}) = \varphi_{i-1} - \alpha H_{inv}(I_0, I_d) - \alpha\gamma \nabla_{\varphi} R(\varphi_{i-1}), \tag{7}$$

where α is the iteration step size. The iteration stops when marginal gain diminishes at the local minimum.

The performance of the regularized optimization is critically influenced by the choice of regularizers and hyper-parameters in R. Instead of using off-the-shelf regularizers as demonstrated in some of the previous TIE retrieval algorithms [75], in this work, a deep learning-based framework is employed to identify the prior features of the objects from the training data. As

Optics EXPRESS

shown in Fig. 2, the iterative optimization process in Eq. (7) is formulated as a cascaded neural network [76–81]

$$\varphi_i = \varphi_{i-1} - \alpha H_{inv}(I_0, I_d) - f_{\mathrm{CNN}}(\varphi_{i-1}), \tag{8}$$

where we employ the same gradient descent calculation of data fidelity term as Eq. (7), and reformulate the gradient descent of R using a convolutional neural network (CNN) f_{CNN}. A total number of N iterations is chosen in advance to keep the architecture of the cascaded neural network fixed. An optimal regularizer f_{CNN}, parameterized by f, is trained by minimizing the mean square error between the network output φ_f and the corresponding ground truth φ_{gt} via

$$\hat{f} = \arg\min_f \frac{1}{2K} \sum_{k=1}^{K} ||\varphi_f^k - \varphi_{gt}^k||^2, \tag{9}$$

where k indexes a training data set with K samples in total. Once trained, the cascaded neural network in Eq. (8) is used to convert the corrupted 2-shot phase retrieval results φ_0 into high fidelity phase images based on the combination of physics models and feature learning [79].

2.3. Architecture of PINN

As shown in Fig. 2, the proposed PINN is built with a cascaded network containing n blocks. Each block contains a feature-learning CNN and a physics-informed data consistency layer in parallel. We incorporate the physics forward model $H(\cdot)$ and the inverse model $H_{inv}(\cdot)$ into the data consistency layer. Note that the inversion algorithm is the 2-shot TIE phase retrieval from Eq. (2) or Eq. (4), which is not the transpose of the forward operator as originally used in the iterative gradient descent algorithm. For the feature-learning f_{CNN}, we use a pruned U-net, called Unet++, to identify structures in various scales. Built of convolution layers with both short and long residual connection pathways, Unet++ is efficient in fusing feature maps of different encoder and decoder levels. Compared with previous model-driven works [76,79], which only use plain skip connections between encoder and decoder networks, Unet++ shows advantages of preserving fine-grained details, therefore presents better results for small structures. In addition, the pruned structure reduces the total network parameters, enabling its fusion into the cascaded blocks and adaption to small-scale data sets. Specifically, Unet++ L^3 [67] is chosen in our scheme, which is built with three components of encoder, decoder, and concatenation layer. Three multi-scale levels of down-sampling and up-sampling are included. In each layer, we implement a 3x3 kernel size in the convolution layer, followed by a batch normalization (BN) and a non-linear activation function (ReLU). The filter number of the seven layers is 32, 64, 128, 256, 128, 64 and 32, respectively; skip connections are established between the feature maps of encoder and decoder at the same level. Dense convolution blocks and skip pathways are added to the first two levels of Unet++. With nested and dense skip connections, each decoder node is carried with not only the feature maps from up-sampled output of the lower dense block but also with the feature maps from intermediate dense convolution block and the original same-scale feature maps from the encoder. We use dropout with a rate of 0.5 in the last two layers to avoid over-fitting. The cascaded neural network contains a total of $N = 3$ iterations; the number 3 is chosen to ensure data convergence and to contain computational load. f_{CNN} is allowed to be parameterized differently in each iteration to improve the overall performance.

The proposed cascaded network is trained using Adam optimizer [82], with the suggested parameters the first-moment exponential decay rate $\beta_1 = 0.9$ and the second-moment exponential decay rate $\beta_2 = 0.999$. The initial learning rate is set as $\gamma = 5e^{-4}$. The mini-batch size is set as 1 in all training processes. We initialize the network weights using a Gaussian distribution with a mean and a standard deviation of 0 and 0.02, respectively. Each sample image is cropped into patches to build a data set with the dimension of 512×512 pixels. The initial trade-off value of

Optics EXPRESS

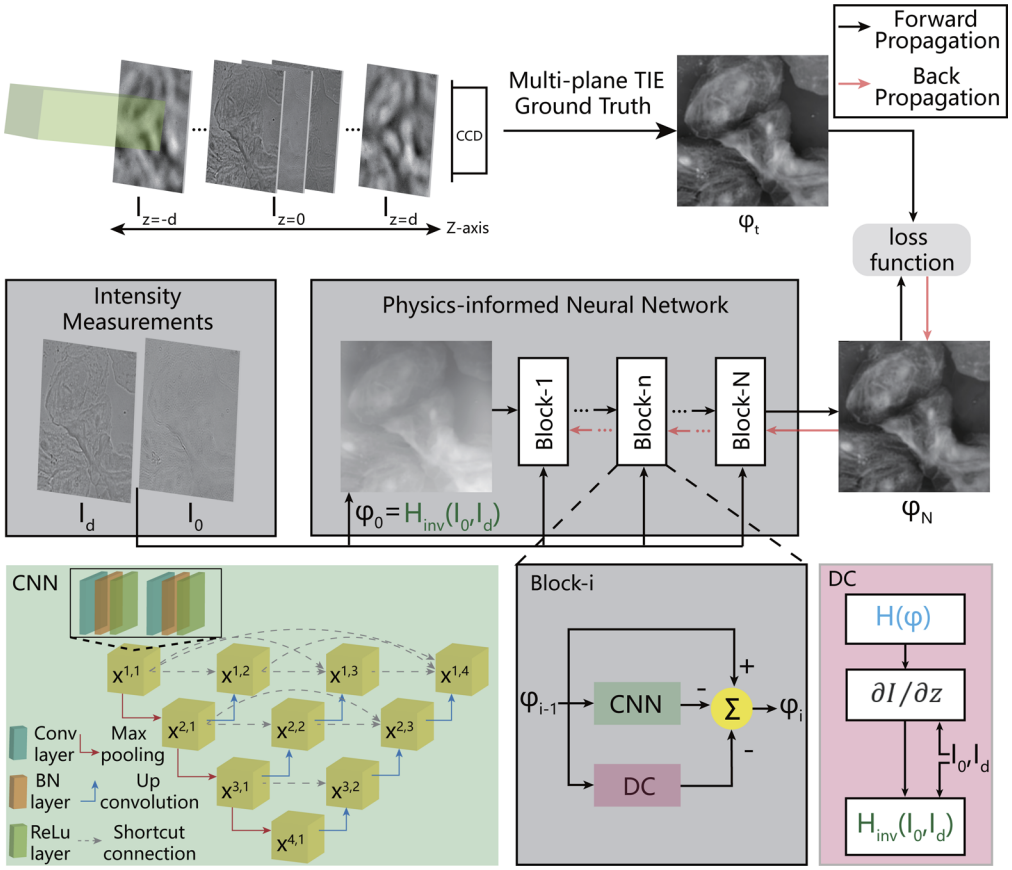

Fig. 2. The architecture of physics informed neural network (PINN) for quantitative phase retrieval. Through-focus intensity measurements and in-focus intensity measurements recorded at different distances are used to generate the ground truth phase profiles in the top row. Each of the N cascaded blocks is a parallel structure made of a data consistency layer (DC) that incorporates the physics forward and inverse models $H(\cdot)$ and the non-transpose-based inversion operator $H_{inv}(\cdot)$, and a CNN layer that learns the features of the objects. The trained network is then used for high quality phase retrieval from low quality 2-shot TIE phase retrieval φ_0 and the intensity measurements. A Unet++ architecture is used as the CNN layer in our PINN.

Optics EXPRESS

the CNN regularization term and data consistency cost is set as $\alpha = 1e^{-2}$. All trained networks are implemented on a NVIDIA 2080 Ti GPU with 11GB memory using Pytorch.

2.4. Simulation and experimental parameters

To simulate TIE phase imaging, we use images from standard image datasets and crop them to arrays of 512×512 pixels. We assume that the illumination wavelength is 660 nm, and the defocused distance is $d = 4\mu m$. Setting the pixel resolution to 0.22 μm, each image is assumed to be of the size 0.112×0.112 mm^2.

To construct pure phase objects φ_{gt}, the pixel intensity of the sample image is linearly mapped into phase from 0 to 3.5 rad, which composes a data set with a similar phase shift as the cheek cells used in the experiment in Sec.3.2 for PINN reconstruction. Note that since TIE phase retrieval responses to the gradient of the phase, there is no phase-unwrapping problem, and can be applied to phase shifts larger than 2π [83]. The in-focus intensity I_0 is assumed to be a uniform illuminated array of 512×512 with every pixel owning the same intensity value. The diffraction pattern I_d is calculated via the forward operator H consisting of Fresnel integral and taking the square of complex amplitude. White Gaussian noise with SNR 20 dB is applied to each simulated Fresnel diffraction pattern I_d and I_0. Inverting the noise corrupted intensity patterns, 2-shot TIE phase retrieval φ_0 is obtained using $H_{inv}(I_0, I_d)$ in Eq. (2).

To construct phase object with absorption, we introduce a strong-varying absorption profile I_0, and use its square root as the amplitude of the input wavefront, associating with the constructed phase. The defocused intensity I_d is then calculated with the same forward operator H and noise corrupted with Gaussian noise of 20 dB SNR. Phase φ_0 is retrieved via H^*_{inv} in Eq. (4).

Experimentally, we use a broadband LED and pass the light through a pinhole. We collimate the beam using a condenser lens (recall Fig. 1). A narrow band spectral filter with a central wavelength of 660 nm is inserted in the collimator to provide a quasi-monochromatic beam. The images of the specimen are collected using an objective lens (20x magnification, NA=0.45) and are mapped on to the CCD camera with a 4.4 μm pixel size, resulting in a pixel size of 0.22 μm on the image. We use human cheek cells as the phase object. The initial φ_0 is obtained from the 2-shot measurement at $d = 0$ and $d = 4\mu$m via $H_{inv}(I_0, I_d)$.

The phase retrieval images are quantitatively validated using the structural similarity index (SSIM) and normalized peak signal-to-noise ratio (PSNR) as,

$$SSIM(\hat{y}, y) = \frac{(2 \times \mu_{\hat{y}}\mu_y + C_1)(2 \times \sigma_{\hat{y}y} + C_2)}{(\mu_{\hat{y}}{}^2 + \mu_y{}^2 + C_1)(\sigma_{\hat{y}}{}^2 + \sigma_y{}^2 + C_2)}, \tag{10}$$

$$PSNR(\hat{y}, y) = 10 \times \log_{10}\left(\frac{MAX(y)^2}{\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(y(i,j) - \hat{y}(i,j))^2}\right). \tag{11}$$

In SSIM metric, μ_y represents the mean of y, $\sigma_{\hat{y}}$ denotes the variance of y, C_1 and C_2 are small constants, and $\sigma_{\hat{y}y}$ specifies the covariance function of y and $\hat{y}$. The $MAX(y)$ represents the maximum possible pixel of y defined in PSNR.

3. Results

3.1. Simulated TIE phase imaging and absorptive phase imaging

The proposed PINN is first validated via simulation. The 500 images featuring scenery in Kaggle data set [84] are used as the training set and test set of phase objects for simulation. In addition, four images from USC-SIPI data set [85] are also used as test data. In this simulation, 80% of the sampled images (400) are randomly selected for training, the rest (104) are used for testing. We simulate for both the TIE phase retrieval on pure phase objects and the TIE absorptive phase retrieval, using non-uniform circular absorptive profile as shown column 1 in Fig. 3(c).

Optics EXPRESS

Figure 3(a) shows the phase retrieval performance for a test image from the same database of Kaggle. Comparing the 2-shot phase retrieval φ_0 in column 1 and φ_{gt} in column 4, we can clearly see severe corruption featured by low-frequency noise. The amplified noise is efficiently removed by the proposed PINN approach, yielding a high-fidelity phase retrieval that has accurately restored the fine details. Out of all the 104 testing samples, the average SSIM and PSNR metrics are significantly increased from 0.478 and 13.38 to 0.817 and 21.84 from φ_0 to φ_{PINN}, demonstrating the efficacy of the proposed physics-informed learning approach.

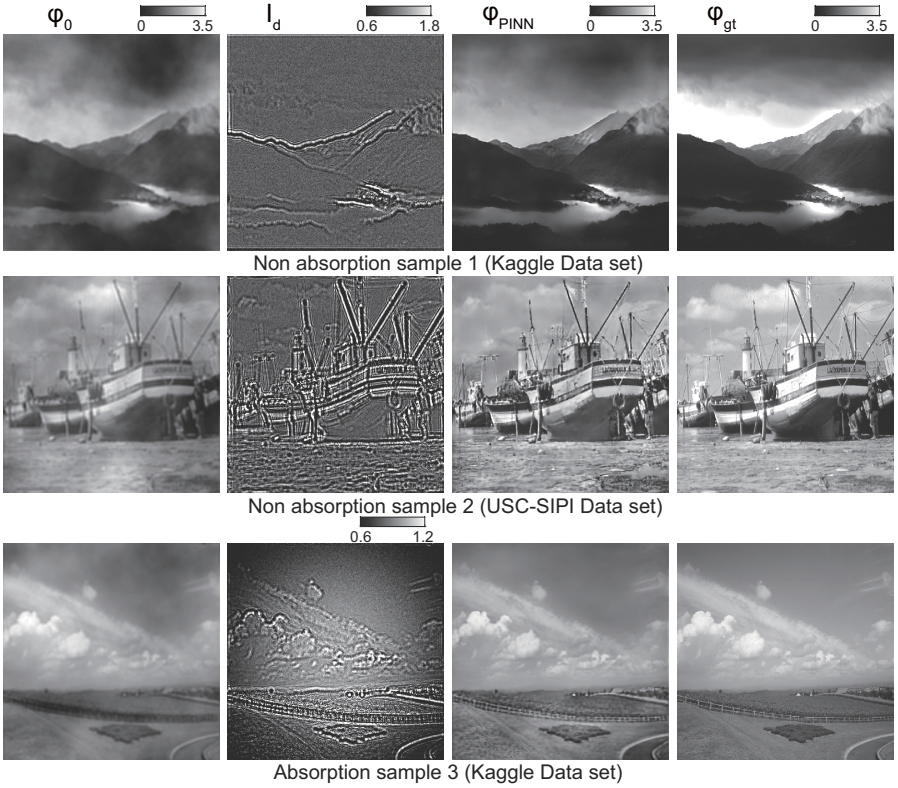

Fig. 3. Simulation results for 2-shot TIE phase retrieval. The 2-shot TIE inputs φ_0 (column 1), defocused diffraction pattern I_d (column 2), phase reconstruction φ_{PINN} (column 3) and ground truth phase profiles φ_{gt} (column 4) are shown. Row 1 and 2 shows the results of pure phase objects, implemented on one test sample image (scenery) in Kaggle data set, and on a different type of image (ship) from USC-SIPI data set, respectively. Row 3 shows the results of absorptive phase imaging based on sample images in Kaggle data set.

Figure 3(b) shows the performance of phase retrieval on a different type of image (ship) from a different database USC-SIPI. The trained PINN is directly applied to the simulated intensities, generated based on the new ship image. Some of the sharp-edged features, such as letters and pillars, are not commonly seen in the scenery images in Kaggle. As a result of changed image feature distribution, the performance of the phase retrieval shows some degree of aberration around those areas. Nevertheless, the severe low-frequency noise in the 2-shot retrieval in column 3 is efficiently removed.

Figure 3(c) shows the phase retrieval performance on absorptive phase objects. A spherical non-uniform absorption is applied to the phase objects, and phase retrieval estimate is obtained by the inverse operator H^*_{inv} defined in Eq. (4). As shown in column 1, in addition to the low-frequency noise that corrupts the image, the coupling of strong absorption gradient and

Optics EXPRESS

strong phase gradient also introduces curl effect with aberration in places such as the wood fence. Such curl effect is again efficiently repaired by the PINN, resolving in clear fence structures in column 3. The quantitative evaluation is also implemented using 104 test absorptive phase examples, where SSIM is increased from 0.446 to 0.798 and PSNR is improved from 12.39 to 20.80.

Overall, the simulation trials demonstrate that the proposed PINN can be efficiently trained for 2-shot TIE phase retrieval with a small dataset (400). The training takes around 4 hours and the phase retrieval with the trained PINN only takes 0.5 second. The trained PINN can be transferred to images of different types with robust performance. We also investigate the performance on absorptive phase imaging, and find that the curl induced error can be efficiently corrected by the PINN.

3.2. Experimental phase imaging of human cheek cells

We next implement the phase imaging of human cheek cells experimentally via the proposed PINN approach. In this experiment, we train a network to translate phase retrieval from 2-shot measurements based on the reconstruction using multi-stack defocused measurements. The purpose of the experiment is to obtain high quality phase retrieval with reduced experimental complexity. Several different neural network frameworks are used to achieve the same purpose and are compared with our proposed PINN. While the multi-stack reconstruction is by no means the ground truth of the phase object, it has been widely considered as high-quality phase retrieval [75,86], especially when the number of measurements of the defocused intensity is high. In our case of human cheek cells, where the ground truth is hard to obtain, it is not uncommon to use multi-stack reconstruction as the target for the training of the neural network [75,86]. Here, a through-focus intensity stack is collected at the defocus distances $d = \pm 2\mu m$, $\pm 4\mu m$, $\pm 8\mu m$, $\pm 16\mu m$, $\pm 32\mu m$, $\pm 64\mu m$, $\pm 128\mu m$, $\pm 256\mu m$. These exponentially spaced defocused intensity images are used to obtain the target phase profile φ_t using Gaussian process regression TIE [72].

We build a database with 200 image sets consisting of I_d, I_0 and φ_t. Out of the 200 image sets, 70% of the samples (140) are randomly selected for training, 15% (30) are randomly selected for validation and the rest 15% (30) are used for testing. For comparison, two other deep learning networks, including the widely used U-net based imaging translation network [48] (abbreviated as U-net) and an end-to-end multi-channel deep-learning network [83] (abbreviated as dTIE), are employed for the same phase retrieval problem and are trained with the same training, validation and testing image sets. The U-net based imaging translation utilizes six levels of residual blocks, followed by two additional standard residual blocks, where corrupted image inputs (2-shot TIE phase retrieval φ_0) are directed translated into enhanced phase image. On the other hand, dTIE adopts a deep end-to-end Resnet structure, and uses the three intensity measurements I_{-d}, I_0 and I_d as the inputs. For fair comparison of our work, we maintain the same network structure in U-net and dTIE with constant learning rate of 0.001, defocused distance ($d = 4\mu m$), input shape of 512x512 with learning batch size of 4 trained for 100 epochs.

Figure 4 shows the phase retrieval results. The target phase profile φ_t as well as the enhanced phase retrieval results φ using the four different neural networks, the proposed PINN in this work, PINN without DC layer, U-net, and dTIE, are shown in Fig. 4(a). As shown in Fig. 2, the input φ_0 suffers from the cloud-like low frequency noise. The noise and artifacts have been suppressed in various degrees by the deep neural networks. While the dTIE network improves the image quality by successfully removing the haze noise, the output φ_{dTIE} contains inadequately preserved features from the inputs, such as the diffraction fringes from the defocused intensity patterns I_d, and reduced contrast introduced by the blurring from φ_0. U-net successfully resolves the fine details of the cells, but reduces the overall image contrast because of the residual of low-frequency noise. We also compare the proposed cascaded structure with DC layer (PINN-DC) and without DC layer (PINN-non-DC). The PINN-non-DC is able to reconstruct most textures and shapes

Optics EXPRESS

of the phase profile but behaves poor at capturing the sharp edge information. In contrast, the proposed PINN-DC resolves the phase profile with the sharpest details and the highest accuracy. A plausible reason is that the prior knowledge on the object introduced by the DC layer can be learned by the neural network. Figure 4(d) shows the loss function with the training and validation data sets. Note that the loss function for the validation data set stabilizes after 80 epochs, which indicates that the proposed PINN is able to converge with this phase retrieval problem. Also, it provides evidence that the current total number of 100 training epochs is suitable to avoid over-fitting. The line profiles of the dashed yellow line in Fig. 4(b) are compared in Fig. 4(e). We see that dTIE (purple dashed line) and PINN-nonDC reduces the contrast, while the U-net curve (yellow dotted line) drifts upwards due to residual low-frequency noise. The result of PINN (green dashed line) fits the target phase profile (pink solid line) best.

The quantitative performance merits (SSIM and PSNR) shown in Table 1 confirm the visual observation. The PINN scores a highest 0.919 in SSIM and 25.23 in PSNR, a significant improvement from 0.588 of SSIM and 16.63 of PSNR for φ_0. Improvement by U-net and dTIE are not as high as that of PINN. Note that the supreme performance of PINN is based on a least number of parameters of 6.6 M ($\times 10^6$), which includes networks in all three cascaded iterations. On the other hand, the numbers of parameters contained in U-net and dTIE are 10.2M and 40.7M, respectively. It takes 2 hours 13 minutes to train the PINN with our current implantation on a single NVIDIA 2080 Ti GPU. Once trained, it only takes 0.54 second to retrieve the phase profile. This is evident that with additional physics information, the proposed PINN is able to achieve high quality phase retrieval results with an efficient compact network structure.

Table 1. Comparison of PSNR (dB) and SSIM for the reconstructed phase profiles using different methods. FACTOR FORMAT : MEAN ± STD.

	TIE retrieval	PINN	U-net	dTIE
Param	–	**6.6*M***	10.2*M*	40.7*M*
PSNR	16.63 ± 3.75	**25.23 ± 3.65**	23.44 ± 3.66	22.84 ± 2.95
SSIM	0.588 ± 0.105	**0.919 ± 0.034**	0.899 ± 0.045	0.877 ± 0.041

3.3. Performance robustness against deficient training data

To test the robustness against deficient training data, we investigate the phase retrieval performance using a varying number of training sample images. Human cheek cells are used as the phase objects for the investigation. The original training data set containing 140 image pairs used in Sec. 3.2 is adopted as the benchmark level of 100%. We then construct four data sets containing 35, 70, 105, 140 image pairs each randomly sampled from the 170 image pool, corresponding to 25%, 50%,75%, and 100% of the benchmark data size. These data sets are used to train the neural networks of PINN as well as the U-net and dTIE introduced in Sec. 3.2. The phase retrieval results are shown in Fig. 5, where we also quantitatively compare the performance of the three deep learning networks via SSIM and PSNR. To evaluate the uncertainty, we repetitively sample the training data sets from the data pool and average the performance merits among 10 independent training sessions for every data set size.

Comparing phase retrieval images in Fig. 5, we see that while all three methods show performance degradation over deficient data, PINN is more robust, showing a slower deterioration. As shown in Fig. 5(b), even the phase retrieval result from PINN trained by 25% of the data set is able to resolve major structure features, such as the cell core and the boundary edges, and successfully removes the background haze noise. In contrast, the phase retrieval results from U-net and dTIE deteriorate considerately with 25% of the data set. As shown in the top image of Fig. 5(c), U-net is not able to completely remove the low frequency noise in the figure, leaving

Optics EXPRESS

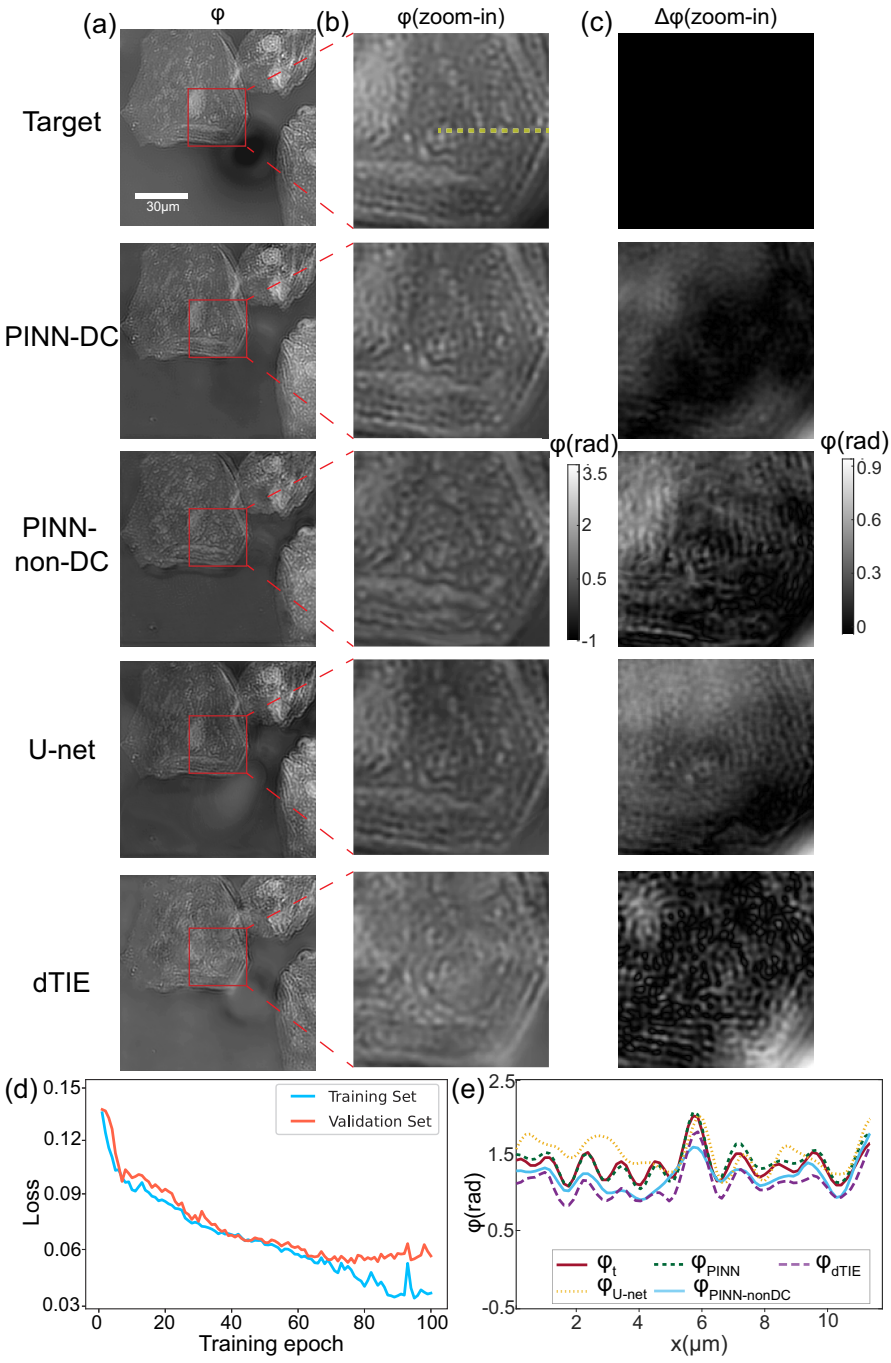

Fig. 4. Experimental phase retrieval results. (a) The phase profiles retrieved by Gaussian process regression TIE φ_t, PINN φ_{PINN}, PINN without DC $\varphi_{\mathrm{PINN-nonDC}}$, U-net $\varphi_{\mathrm{U-net}}$, and dTIE φ_{dTIE}. (b) The zoomed-in area near cell core in (a). (c) The error map as compared to the target φ_t. (d) The convergence of loss for the training and validation data sets during the training of PINN. (e) Comparison of the line profiles of the yellow dashed line labeled in (b).

Optics EXPRESS

the lower right part of the image untruthfully dark. Severe artifacts of boundary shadows are also observed. The phase reconstruction result of dTIE in the top row of Fig. 5(d) blurs the image and misinterprets some of the major structures, such as the central core.

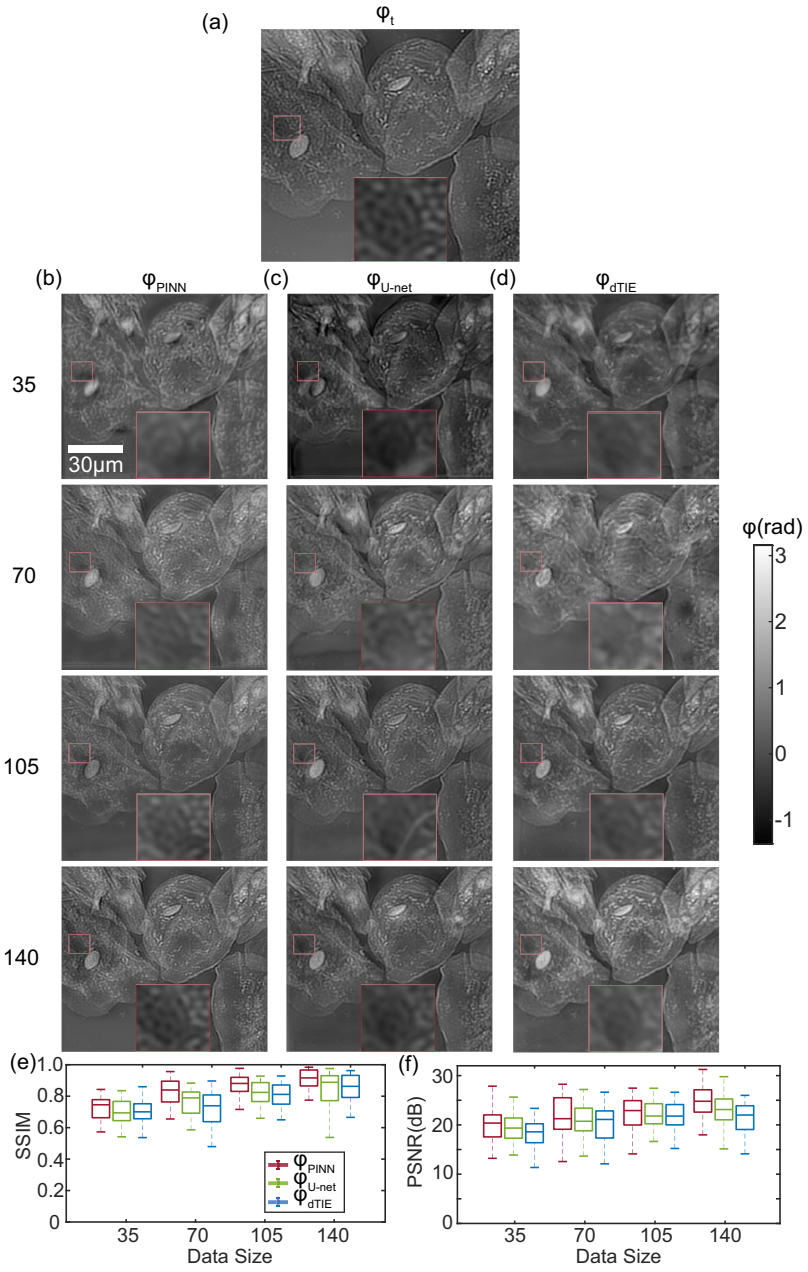

Fig. 5. Phase retrieval performance with deficient training data. (a) Target phase profiles φ_t. Phase retrieval results based on (b) PINN φ_{PINN}, (c) U-net $\varphi_{\text{U-net}}$ and (d) dTIE φ_{dTIE} using 35, 70, 105 and 140 image patches of the original training data. Quantitative merits of (e) SSIM and (f) PSNR are shown with boxplots indicating the average performance variances among 10 independent training sessions.

Optics EXPRESS

Quantitative analysis of SSIM and PSNR in Figs. 5(e) and (f) is consistent with the visual observation. PINN not only outperforms the non-physics informed networks when there is sufficient training data, it also demonstrates higher robustness against data deficiency. We find that when the training data set size is reduced to 50% (70 image sets), PINN (red box) yields a small drop (8%) of average SSIM from 0.919 to 0.846. In contrast, non-physics informed DNN approaches of U-net and dTIE deteriorate much faster. With 50% data size reduction, U-net (green box) yields a 13% drop of performance from 0.899 to 0.784, while dTIE (blue box) drops by 14% from 0.877 to 0.754. Notably, the high-performing PINN is built using a smaller number of parameters, which is 40% less than the U-net and 84% less than the dTIE networks (Table 1). Combining the compact network with a relatively small training data size, the PINN approach improves the efficiency of phase retrieval applications, despite having a cascaded structure.

3.4. Application to four-shot TIE phase retrieval via transfer learning

So far, we have demonstrated that PINN is a prominent approach for 2-shot TIE phase retrieval. In this section, we show that a well-trained 2-shot PINN can be utilized to reconstruct the phase images of similar objects in a new acquisition setting via transfer learning [87,88]. Instead of training the neural networks from a random initialization, transfer learning focuses on utilizing the learned knowledge from one problem to solve a related problem. By doing so, the related problem is expected to be efficiently solved with less computation and training data. In imaging applications, a common situation is that the imaging setup is slightly changed while the objects of interest remain the same. For example, the alignment and focusing of the microscope may change from trial to trial; wear on the surface of the lenses may introduce unexpected scattering; the light source may lose coherence overtime, etc. In these cases, the forward models in the physical process are changed, but the features of the objects remain the same. In an end-to-end, non-transferred learning framework, the features introduced by the aberration from the imaging setup are mixed with the features of the objects, which are only identified (hopefully) after another computationally expensive learning process with the excess burden of collecting a large set of training data in the new setup.

The proposed PINN, on the other hand, is well suited for transfer learning in altered imaging settings. Benefiting from the parallel structure with separated data consistency layer and CNN-regularization layer, PINN can adapt to new imaging settings by updating the data consistency layer with the new physics models while transferring the learned features from a pre-trained CNN-regularization layer, as shown in Fig. 6. The latter is achieved by first transferring the network architecture and initialize the target model with a pre-trained CNN; the target model is then fine-tuned by additional training. In this way, the training process of phase imaging can be significantly accelerated.

To demonstrate the transfer learning performance of PINN phase retrieval, we transfer the PINN previously trained in Sec. 3 (2-shot TIE) to a 4-shot TIE phase retrieval problem. In the 4-shot TIE experiments, we capture four de-focused intensity measurements I_1, I_{-1}, I_2, I_{-2} at $d_1 = \pm 2\mu$m, $d_2 = \pm 4\mu$m, respectively. The intensity differential is obtained using the high-order finite difference method via Taylor expansion as [22],

$$\left.\frac{\partial I(x,y;z)}{\partial z}\right|_{z=0} \approx \sum_{i=-2}^{i=+2} \frac{a_i I_i}{\Delta z}, \tag{12}$$

where

$$a_i = \frac{4i(-1)^{i+1}}{(2+i)!(2-i)!}, \tag{13}$$

and $\Delta z = 2\mu$m. The 4-shot TIE forward and inverse models are then built by substituting the intensity differential Eq. (12) into Eq. (1) and Eq. (2) respectively, and are used to update the data consistency layer.

Optics EXPRESS

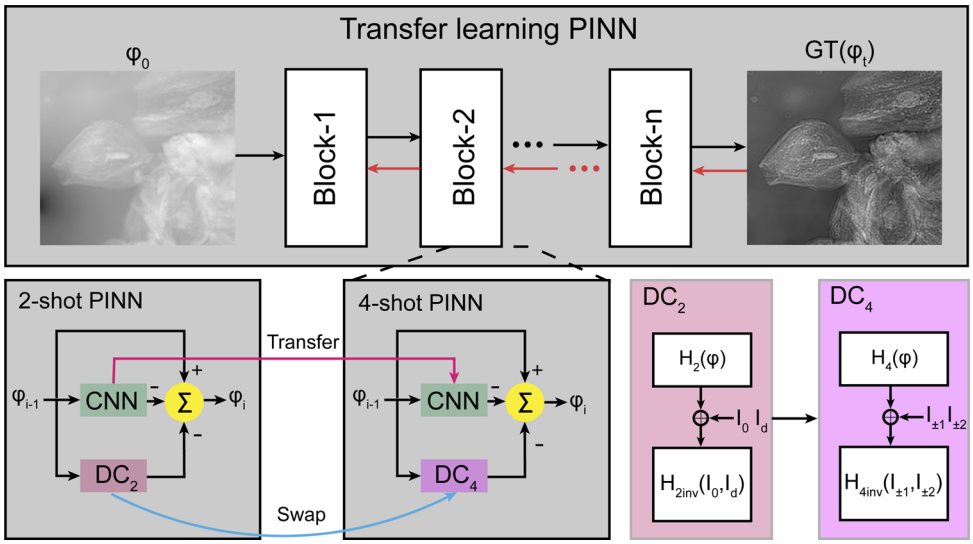

Fig. 6. Transfer learning scheme for PINN. A trained 2-shot PINN is used as the initialization for the training of a 4-shot TIE PINN, with the data consistency layer DC_2 replaced by the 4-shot model DC_4.

We then train a 4-shot PINN initialized by the pre-trained 2-shot PINN using 140 image pairs. The training performance is evaluated by the training loss between the retrieved phase objects and the corresponding target phase profiles in a validation set at different training epochs. In this case, we form a validation set containing 30 image sets, and the training loss is obtained by computing the mean square error among them. To compare, we also use the same set of data to train a randomly initialized 4-shot PINN without transfer learning. The performance comparison is summarized in Fig. 7

As shown in Fig. 7(b), the transfer learning approach (dashed red line) constantly outperforms non-transferred PINN (solid green line). In particular, transfer learning starts at a low loss level and converges very fast. For as little as 5 epochs (Point ii), the training loss of transfer-learned PINN is already on-par with that of non-transferred PINN at the end of training with 50 epochs (Point i). The final training loss for transfer learning with 50 epochs (Point iii) is also significantly lower than the final loss for non-transferred PINN at Point i. The corresponding phase retrieval results generated by PINN at Point i, ii and iii are shown in Fig. 7(c), which demonstrate the same trend. The image quality of transfer-learned PINN after 5 epochs (ii) is close to that of non-transferred PINN with 50 epochs (i); With a same number of 50 epochs, transfer learning (iii) achieves the best performance with a SSIM of 0.923, exceeding SSIM of 0.905 for non-transferred PINN.

Interestingly, we find that the transfer learning of PINN can be further accelerated by adopting a pre-trained 2-shot PINN with reduced epoch numbers. The transferred PINN initialized by a 2-shot PINN trained with 20 epochs shows almost an identical convergence performance, and obtains an SSIM of 0.923 by the end of 50 epochs. The results demonstrate that the lower weight parameters optimized in the first few epochs contain the most valuable knowledge to transfer. Overall, our results are evident that the learned features of manifold patches by PINN can be transferred to resolve similar imaging inversion problems with reduced computational cost.

Optics EXPRESS

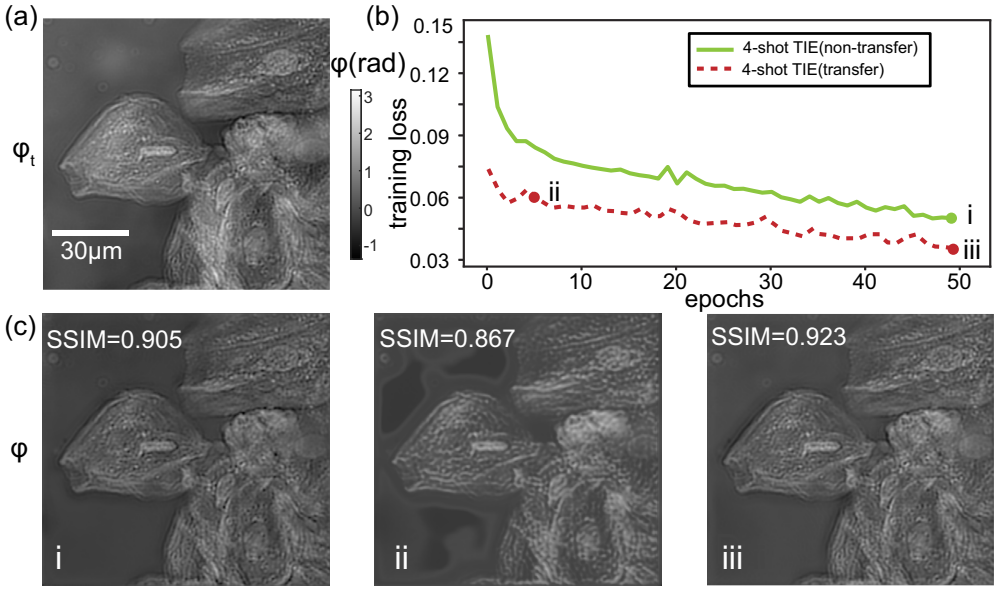

Fig. 7. 4-shot TIE phase retrieval using transfer learning. (a) The target phase profile φ_t, (b) the convergence curves based on training loss as a function of epoch number for non-transferred PINN (solid green line), and transfer learning PINN (dashed red line, initialized with the 2-shot PINN trained for 100 epochs). (c) Phase retrieval results φ_{PINN} for the various epoch stages i (non-transferred), ii (transferred) and iii (transferred) as indicated in (b).

4. Discussion

In this section, we discuss phase retrieval approaches using deep neural network beyond the two methods of dTIE and Unet used for comparison in the previous section. Compared with iterative methods which optimize the phase retrieval with deep priors-based generic regularizers [89], a trained PINN can significantly reduce the reconstruction time of phase retrieval. Regarding deep learning strategies, the most widely used network structures for phase retrieval are based on Resnet [90], DenseNet [91] or their variations. In such end-to-end phase reconstruction scheme, the large-scale data set with input estimates and targets are always required to reconstruct accurate phase results. Another promising method used for phase imaging is generative adversarial network (GANs) [51], which builds two separated neural networks known as the generator and discriminator to improve model performance. Note that these deep learning strategies are entirely data-driven, meaning that they depend only on the network parameters and data set and is ignorant to the physics of the imaging system. On the other hand, physics-aware model is usually much less dependent on the training data. As demonstrated by recent advances of untrained neural network [62,63], such physics-aware architecture can be iteratively optimized based on only a few or even a single measurement, thereby greatly reducing the constraint of big data collection. Alternatively, recent studies of unpaired training connect physics model with CycleGAN [92] and demonstrate success of image-to-image translation with cyclic constraints. This efficient approach is also recently applied to address phase reconstruction with incorporated forward model [60], leading to ultra-fast and accurate phase reconstruction. The potential difficulty of CycleGAN-based works arises from the training of two GANs. In other words, the balance of generator and discriminator with consistency constraints may be challenging in each case study. In comparison, our proposed PINN follows an iterative optimization scheme and combines the

Optics EXPRESS

physics image formation model and learning-based techniques in each iteration. The easy-training PINN can produce high-quality phase retrieval results with reduced data set, and the incorporated DC term is flexible to fit training samples collected in different environments.

5. Conclusions

To conclude, we have demonstrated that the physics-informed neural network (PINN) is capable of embedding physics knowledge in parallel with deep neural networks for high quality TIE-based phase retrieval. With the physics information, the proposed PINN successfully removes the low-frequency noise and corrects for curl induced errors commonly observed in TIE phase imaging. In the experiments presented, the proposed PINN is shown to perform better than two end-to-end deep-learning algorithms from literature, both in terms of imaging quality and robustness against insufficient training. It is also shown to be easily transferred to different types of objects or different imaging setups. By using PINN to integrate physics knowledge with the object prior features, high quality phase retrieval can be obtained using less training data and faster computational speed.

Funding. National Science Foundation (QMR-1825646).

Acknowledgments. The authors would like to acknowledge Dr. Yizheng Zhu in ECE Department at Virginia Tech for providing the optical lab space.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in [93].

References

1. Y. Cotte, F. Toy, P. Jourdain, N. Pavillon, D. Boss, P. Magistretti, P. Marquet, and C. Depeursinge, "Marker-free phase nanoscopy," Nat. Photonics **7**(2), 113–117 (2013).
2. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," Nat. Photonics **12**(10), 578–589 (2018).
3. P. Bon, N. Bourg, S. Lécart, S. Monneret, E. Fort, J. Wenger, and S. Lévêque-Fort, "Three-dimensional nanometre localization of nanoparticles to enhance super-resolution microscopy," Nat. Commun. **6**(1), 7764 (2015).
4. A. Olivo, K. Ignatyev, P. Munro, and R. Speller, "Design and realization of a coded-aperture based x-ray phase contrast imaging for homeland security applications," Nucl. Instrum. Methods Phys. Res., Sect. A **610**(2), 604–614 (2009).
5. M. Stampanoni, Z. Wang, T. Thüring, C. David, E. Roessl, M. Trippel, R. A. Kubik-Huch, G. Singer, M. K. Hohl, and N. Hauser, "The first analysis and clinical evaluation of native breast tissue using differential phase-contrast mammography," Invest. Radiol. **46**(12), 801–806 (2011).
6. E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," J. Opt. Soc. Am. **52**(10), 1123–1130 (1962).
7. E. N. Leith, J. Upatnieks, and K. A. Haines, "Microscopy by wavefront reconstruction," J. Opt. Soc. Am. **55**(8), 981–986 (1965).
8. P. Marquet, B. Rappaz, P. J. Magistretti, E. Cuche, Y. Emery, T. Colomb, and C. Depeursinge, "Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy," Opt. Lett. **30**(5), 468–470 (2005).
9. D. Paganin and K. A. Nugent, "Noninterferometric phase imaging with partially coherent light," Phys. Rev. Lett. **80**(12), 2586–2589 (1998).
10. J. Liao, Z. Wang, Z. Zhang, Z. Bian, K. Guo, A. Nambiar, Y. Jiang, S. Jiang, J. Zhong, M. Choma, and G. Zheng, "Dual light-emitting diode-based multichannel microscopy for whole-slide multiplane, multispectral and phase imaging," J. Biophotonics **11**(2), e201700075 (2018).
11. A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov, "On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation," Rev. Sci. Instrum. **66**(12), 5486–5492 (1995).
12. B. Allman, P. McMahon, K. A. Nugent, D. Paganin, D. L. Jacobson, M. Arif, and S. Werner, "Phase radiography with neutrons," Nature **408**(6809), 158–159 (2000).
13. P. Cloetens, W. Ludwig, J. Baruchel, D. Van Dyck, J. Van Landuyt, J. Guigay, and M. Schlenker, "Holotomography: Quantitative phase tomography with micrometer resolution using hard synchrotron radiation x rays," Appl. Phys. Lett. **75**(19), 2912–2914 (1999).
14. Y. Zhu and G. Barbastathis, "Superresolution tie phase imaging by structured illumination," in *Digital Holography and Three-Dimensional Imaging* (Optica Publishing Group, 2015), paper DT1A–6.
15. M.-C. Zdora, "State of the art of x-ray speckle-based phase-contrast and dark-field imaging," J. Imaging **4**(5), 60 (2018).

Optics EXPRESS

16. Y. Zhu, A. Shanker, L. Tian, L. Waller, and G. Barbastathis, "Low-noise phase imaging by hybrid uniform and structured illumination transport of intensity equation," Opt. Express **22**(22), 26696–26711 (2014).
17. S. S. Kou, L. Waller, G. Barbastathis, and C. J. Sheppard, "Transport-of-intensity approach to differential interference contrast (ti-dic) microscopy for quantitative phase imaging," Opt. Lett. **35**(3), 447–449 (2010).
18. C. Zuo, Q. Chen, L. Tian, L. Waller, and A. Asundi, "Transport of intensity phase retrieval and computational imaging for partially coherent fields: The phase space perspective," Opt. Lasers Eng. **71**, 20–32 (2015).
19. T. C. Petersen, V. J. Keast, and D. M. Paganin, "Quantitative tem-based phase retrieval of mgo nano-cubes using the transport of intensity equation," Ultramicroscopy **108**(9), 805–815 (2008).
20. M. Krenkel, M. Bartels, and T. Salditt, "Transport of intensity phase reconstruction to solve the twin image problem in holographic x-ray imaging," Opt. Express **21**(2), 2220–2235 (2013).
21. N. Yoneda, A. Onishi, Y. Saita, K. Komuro, and T. Nomura, "Single-shot higher-order transport-of-intensity quantitative phase imaging based on computer-generated holography," Opt. Express **29**(4), 4783–4801 (2021).
22. L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," Opt. Express **18**(12), 12552 (2010).
23. M. Beleggia, M. Schofield, V. Volkov, and Y. Zhu, "On the transport of intensity technique for phase retrieval," Ultramicroscopy **102**(1), 37–49 (2004).
24. D. Paganin, A. Barty, P. McMahon, and K. A. Nugent, "Quantitative phase-amplitude microscopy. iii. the effects of noise," J. Microsc. **214**(1), 51–61 (2004).
25. Y. Zhu, Z. Zhang, and G. Barbastathis, "Phase imaging for absorptive phase objects using hybrid uniform and structured illumination transport of intensity equation," Opt. Express **22**(23), 28966–28976 (2014).
26. C. Zuo, Q. Chen, Y. Yu, and A. Asundi, "Transport-of-intensity phase imaging using savitzky-golay differentiation filter-theory and applications," Opt. Express **21**(5), 5346–5362 (2013).
27. C. Zuo, J. Sun, J. Li, J. Zhang, A. Asundi, and Q. Chen, "High-resolution transport-of-intensity quantitative phase microscopy with annular illumination," Sci. Rep. **7**(1), 7654 (2017).
28. L. Lu, Y. Fan, J. Sun, J. Zhang, X. Wu, Q. Chen, and C. Zuo, "Accurate quantitative phase imaging by the transport of intensity equation: a mixed-transfer-function approach," Opt. Lett. **46**(7), 1740–1743 (2021).
29. M. Langer, P. Cloetens, and F. Peyrin, "Regularization of phase retrieval with phase-attenuation duality prior for 3-d holotomography," IEEE Trans. on Image Process. **19**(9), 2428–2436 (2010).
30. V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Non-linear iterative phase retrieval based on frechet derivative," Opt. Express **19**(23), 22809–22819 (2011).
31. B. Sixou, V. Davidoiu, M. Langer, and F. Peyrin, "Absorption and phase retrieval with tikhonov and joint sparsity regularizations," Inverse Problems Imaging **7**(1), 267–282 (2013).
32. S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage, "Regularized newton methods for x-ray phase contrast and general imaging problems," Opt. Express **24**(6), 6490–6506 (2016).
33. E. Bostan, E. Froustey, B. Rappaz, E. Shaffer, D. Sage, and M. Unser, "Phase retrieval by using transport-of-intensity equation and differential interference contrast microscopy," in *International Conference on Image Processing (ICIP)* (IEEE, 2014), pp. 3939–3943.
34. E. Bostan, E. Froustey, M. Nilchian, D. Sage, and M. Unser, "Variational phase imaging using the transport-of-intensity equation," IEEE Trans. on Image Process. **25**(2), 807–817 (2016).
35. L. Tian, J. C. Petruccelli, and G. Barbastathis, "Nonlinear diffusion regularization for transport of intensity phase imaging," Opt. Lett. **37**(19), 4131–4133 (2012).
36. S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," in *Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2 (IEEE, 2005), pp. 860–867.
37. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," IEEE Trans. Image Process. **16**(8), 2080–2095 (2007).
38. C. Zuo, J. Li, J. Sun, Y. Fan, J. Zhang, L. Lu, R. Zhang, B. Wang, L. Huang, and Q. Chen, "Transport of intensity equation: a tutorial," Opt. Lasers Eng. **135**, 106187 (2020).
39. R. Bie, X.-H. Yuan, M. Zhao, and L. Zhang, "Method for estimating the axial intensity derivative in the tie with higher order intensity derivatives and noise suppression," Opt. Express **20**(7), 8186–8191 (2012).
40. S. Zheng, B. Xue, W. Xue, X. Bai, and F. Zhou, "Transport of intensity phase imaging from multiple noisy intensities measured in unequally-spaced planes," Opt. Express **20**(2), 972–985 (2012).
41. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," IEEE Trans. on Image Process. **26**(7), 3142–3155 (2017).
42. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," IEEE Trans. on Image Process. **26**(9), 4509–4522 (2017).
43. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," IEEE J. Sel. Areas Inf. Theory **1**(1), 39–56 (2020).
44. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," Optica **6**(8), 921–943 (2019).
45. Z. Wang, J. Chen, and S. C. Hoi, "Deep learning for image super-resolution: A survey," IEEE Trans. Pattern Anal. Mach. Intell. **43**(10), 3365–3387 (2021).
46. M. Genzel, J. Macdonald, and M. Marz, "Solving inverse problems with deep neural networks-robustness included," IEEE Transactions on Pattern Analysis and Machine Intelligence (to be published).

Optics EXPRESS

47. J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," Adv. Opt. Photonics **10**(2), 409–483 (2018).
48. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," Optica **4**(9), 1117–1125 (2017).
49. M. Deng, S. Li, A. Goy, I. Kang, and G. Barbastathis, "Learning to synthesize: robust phase retrieval at low photon counts," Light: Sci. Appl. **9**(1), 36 (2020).
50. I. Kang, F. Zhang, and G. Barbastathis, "Phase extraction neural network (phenn) with coherent modulation imaging (cmi) for phase retrieval at low photon counts," Opt. Express **28**(15), 21578–21600 (2020).
51. Z. Wu, X. Wu, and Y. Zhu, "Structured illumination-based phase retrieval via generative adversarial network," Proc. SPIE **11249**, 112490L (2020).
52. K. Mom, B. Sixou, and M. Langer, "Mixed scale dense convolutional networks for x-ray phase contrast imaging," Appl. Opt. **61**(10), 2497–2505 (2022).
53. D. M. Pelt and J. A. Sethian, "A mixed-scale dense convolutional neural network for image analysis," Proc. Natl. Acad. Sci. **115**(2), 254–259 (2018).
54. H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "Learn: Learned experts' assessment-based reconstruction network for sparse-data ct," IEEE Trans. Med. Imaging **37**(6), 1333–1347 (2018).
55. S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, "Solving inverse problems using data-driven models," Acta Numerica **28**, 1–174 (2019).
56. N. Chen, C. Wang, and W. Heidrich, "Holographic 3d particle imaging with model-based deep network," IEEE Trans. Comput. Imaging **7**, 288–296 (2021).
57. H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "Nett: Solving inverse problems with deep neural networks," Inverse Prob. **36**, 065005 (2020).
58. B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," Magn. Reson. Med. **84**(6), 3172–3191 (2020).
59. V. A. Van der Schaaf and J. V. van Hateren, "Modelling the power spectra of natural images: statistics and information," Vision Res. **36**(17), 2759–2770 (1996).
60. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, "Phasegan: a deep-learning phase-retrieval approach for unpaired datasets," Opt. Express **29**(13), 19593–19604 (2021).
61. C. Bai, M. Zhou, J. Min, S. Dang, X. Yu, P. Zhang, T. Peng, and B. Yao, "Robust contrast-transfer-function phase retrieval via flexible deep learning networks," Opt. Lett. **44**(21), 5141–5144 (2019).
62. F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," Light: Sci. Appl. **9**(1), 77 (2020).
63. F. Yang, T.-A. Pham, N. Brandenberg, M. P. Lütolf, J. Ma, and M. Unser, "Robust phase unwrapping via deep image prior for quantitative phase imaging," IEEE Trans. on Image Process. **30**, 7025–7037 (2021).
64. J. M. Ede, "Deep learning in electron microscopy," Machine Learning: Science and Technology **2**, 011004 (2021).
65. F. Knoll, K. Hammernik, E. Kobler, T. Pock, M. P. Recht, and D. K. Sodickson, "Assessment of the generalization of learned image reconstruction and the potential for transfer learning," Magn. Reson. Med. **81**(1), 116–128 (2019).
66. M. R. Kellman, E. Bostan, N. A. Repina, and L. Waller, "Physics-based learned design: Optimized coded-illumination for quantitative phase imaging," IEEE Trans. Comput. Imaging **5**(3), 344–353 (2019).
67. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (Springer, 2018), pp. 3–11.
68. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation," IEEE Trans. Med. Imaging **39**(6), 1856–1867 (2020).
69. M. R. Teague, "Deterministic phase retrieval: a green's function solution," J. Opt. Soc. Am. **73**(11), 1434–1441 (1983).
70. M. R. Teague, "Irradiance moments: their propagation and use for unique retrieval of phase," J. Opt. Soc. Am. **72**(9), 1199–1209 (1982).
71. M. Soto and E. Acosta, "Improved phase imaging from intensity measurements in multiple planes," Appl. Opt. **46**(33), 7978–7981 (2007).
72. Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller, "Transport of intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes," Opt. Express **22**(9), 10661–10674 (2014).
73. J. Sun, C. Zuo, and Q. Chen, "Iterative optimum frequency combination method for high efficiency phase imaging of absorptive objects based on phase transfer function," Opt. Express **23**(21), 28031–28049 (2015).
74. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier, 2013).
75. E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, "Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network," Optica **7**(6), 559–562 (2020).
76. J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic mr image reconstruction," IEEE Trans. Med. Imaging **37**(2), 491–503 (2018).
77. H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," IEEE Trans. Med. Imaging **38**(2), 394–405 (2019).

Optics EXPRESS

78. K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated mri data," Magn. Reson. Med. **79**(6), 3055–3071 (2018).
79. J. Schlemper, S. S. M. Salehi, P. Kundu, C. Lazarus, H. Dyvorne, D. Rueckert, and M. Sofka, "Nonuniform variational network: Deep learning for accelerated nonuniform mr image reconstruction," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2019), pp. 57–64.
80. K. Zhang, L. V. Gool, and R. Timofte, "Deep unfolding network for image super-resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 3217–3226.
81. V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," IEEE Signal Process. Mag. **38**(2), 18–44 (2021).
82. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 (2014).
83. K. Wang, J. Di, Y. Li, Z. Ren, Q. Kemao, and J. Zhao, "Transport of intensity equation from a single intensity image via deep learning," Opt. Lasers Eng. **134**, 106233 (2020).
84. A. Rougetet, "Kaggle Landscape Pictures," https://www.kaggle.com/datasets/arnaud58/landscape-pictures. Accessed: 2022-08-06.
85. "The USC-SIPI Image Database," https://sipi.usc.edu/database/. Accessed: 2022-08-06.
86. F. Wang, A. Eljarrat, J. Müller, T. R. Henninen, R. Erni, and C. T. Koch, "Multi-resolution convolutional neural networks for inverse problems," Sci. Rep. **10**, 1–11 (2020).
87. J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" Adv. Neural Inform. Process. Syst. **27**, 3320–3328 (2014).
88. C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *International Conference on Artificial Neural Networks* (Springer, 2018), pp. 270–279.
89. S. Kumar, "Phase retrieval with physics informed zero-shot network," Opt. Lett. **46**(23), 5942–5945 (2021).
90. Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light: Sci. Appl. **7**(2), 17141 (2018).
91. Y. Wu, J. Wu, S. Jin, L. Cao, and G. Jin, "Dense-u-net: Dense encoder–decoder network for holographic imaging of 3d particle fields," Opt. Commun. **493**, 126970 (2021).
92. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 2223–2232.
93. X. Wu, S. Chakravarthy Shanmugavel, and Y. Zhu, "Physics-informed neural network for phase imaging based on transport of intensity equation," https://github.com/dongxixiaoqiang915/Physics-informed-neural-network (2022).