Research Article | Vol. 31, No. 3 / 30 Jan 2023 / Optics Express 4094
Optics EXPRESS

Multiple-scattering simulator-trained neural network for intensity diffraction tomography

ALEX MATLOCK,[1,2,*] JIABEI ZHU,[1] AND LEI TIAN[1]

[1]Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215, USA
[2]Current address: Laser Biomedical Research Center, G.R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[*]amatlock@mit.edu

Abstract: Recovering 3D phase features of complex biological samples traditionally sacrifices computational efficiency and processing time for physical model accuracy and reconstruction quality. Here, we overcome this challenge using an approximant-guided deep learning framework in a high-speed intensity diffraction tomography system. Applying a physics model simulator-based learning strategy trained entirely on natural image datasets, we show our network can robustly reconstruct complex 3D biological samples. To achieve highly efficient training and prediction, we implement a lightweight 2D network structure that utilizes a multi-channel input for encoding the axial information. We demonstrate this framework on experimental measurements of weakly scattering epithelial buccal cells and strongly scattering *C. elegans* worms. We benchmark the network's performance against a state-of-the-art multiple-scattering model-based iterative reconstruction algorithm. We highlight the network's robustness by reconstructing dynamic samples from a living worm video. We further emphasize the network's generalization capabilities by recovering algae samples imaged from different experimental setups. To assess the prediction quality, we develop a quantitative evaluation metric to show that our predictions are consistent with both multiple-scattering physics and experimental measurements.

1. Introduction

Imaging complex 3D biological sample remains an outstanding challenge in quantitative phase imaging (QPI) due to the inherent trade-off between physical model accuracy and computational efficiency for traditional techniques. The most efficient 3D QPI techniques utilize single-scattering models, such as the first Born or Rytov approximations, providing computationally efficient closed-form solutions for volumetric recovery [1,2]. Both interferometry-based [1,2] and intensity-based [3–6] 3D QPI methods using these approximations are successful in applications including immuno-oncology [7], cytopathology [8], and stem cell research [9]. However, the single-scattering model often underestimates the refractive index (RI) of strongly scattering samples [5,10–13] preventing the quantitative analysis of complex biology including tissue biopsies and organoids. Recent efforts have shown improved RI estimates can be achieved using multiple-scattering model based iterative reconstruction algorithms [10,11,13–15]. This improved accuracy requires greater computation times that limit real-time evaluation of dynamic samples and large-scale objects. In addition, the reconstructions are still limited by the missing-cone problem due to limited angular coverage in the measurement [12].

Here, we overcome this tradeoff between computational efficiency and reconstruction accuracy by melding closed-form single-scattering solutions with a fast and generalizable deep learning (DL) model. We illustrate this synergistic approach on our recently developed high-speed annular intensity diffraction tomography (aIDT) system, a computational 3D phase imaging modality using oblique illumination from a ring LED illuminator to encode 3D phase into intensity measurements (Fig. 1(a)) [5]. Our prior work demonstrated efficient 3D phase recovery with >10Hz volume rates based on a linear single-scattering model (Fig. 1(a)) [5]. To maintain fast

#477396 https://doi.org/10.1364/OE.477396
Journal © 2023 Received 3 Nov 2022; revised 22 Dec 2022; accepted 29 Dec 2022; published 20 Jan 2023

Optics EXPRESS

and quantitative 3D phase imaging, we combine aIDT with the proposed DL model for live sample imaging with minimal artifacts on both weakly and strongly scattering samples.

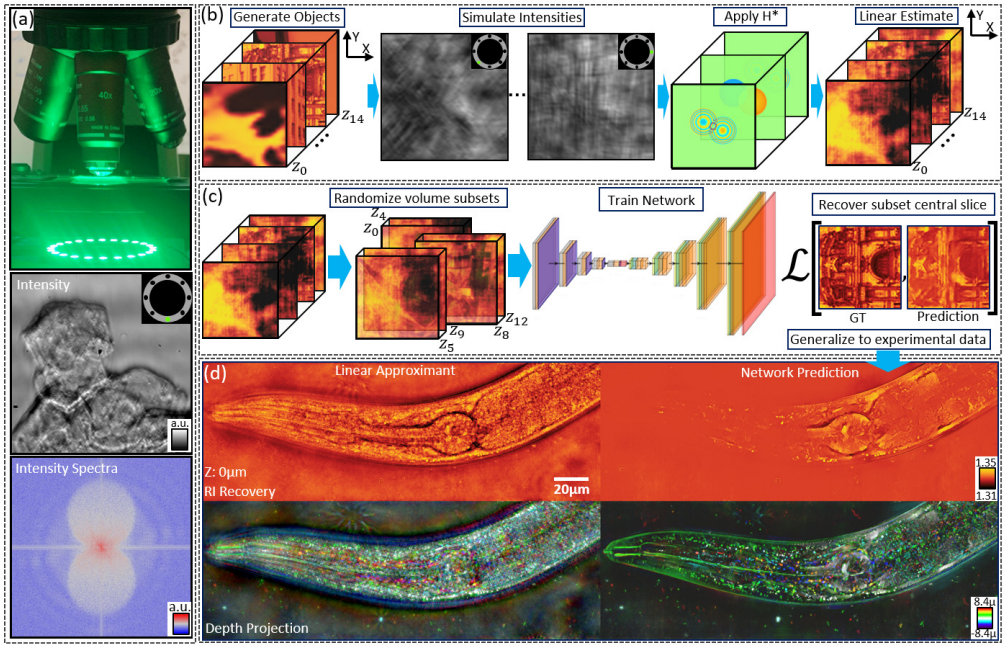

Fig. 1. Multiple-scattering simulator-trained neural network for intensity diffraction tomography. (a) aIDT imaging setup (top) with example intensity image (middle) and its spectra (bottom) under single-LED oblique illumination (LED pattern shown as an inset). (b) Simulation process for training data generation. Natural images are randomly sorted into volumes with randomized RI. Intensity images are simulated using the SSNP multiple-scattering model. Linear approximants are computed using the single-scattering aIDT model for the network inputs. (c) Training process for IDTnet. The simulated object volumes are randomly segmented into five-slice subsets during training and fed into the network to recover the central slice. (d) Example output from IDTnet on a *C. elegans* worm compared with the single-scattering reconstruction (Top: in-focus RI slice; Bottom: color-coded depth projections).

DL has revolutionized the fields of computational microscopy with its ability to efficiently solve complex nonlinear inverse problems [16]. Existing DL techniques for QPI utilize different learning strategies from full "end-to-end" models for direct inversion common in 2D QPI methods [17–19] to "guided" learning with embedded physical models or priors often used for 3D QPI [20–27]. The DL techniques for 3D QPI successfully improved RI predictions on red blood cells [21], high-contrast manufactured samples [24,28], and complex 3D biological samples [22,23,26]. However, the generalizability of supervised DL techniques [21,24,28] is limited due to the similarities between the training and testing data. For biological applications where features of interest vary between sample types, the limited robustness against sample variations hinders their broad application. The "untrained" DL techniques [22,23,26] use a deep neural network (DNN) to parameterize the RI distribution and/or the priors and require "training" on each new sample, a process that translates poorly to large-scale time-series applications. Furthermore, all existing DNN for 3D QPI utilize 3D network structures and contain a large number of trainable parameters that further complicate the computational requirements.

Here, we overcome these existing limitations by leveraging a multiple-scattering simulator and a lightweight DNN architecture to achieve efficient object recovery on complex biological

Optics EXPRESS

samples. First, we develop a physical model simulator-based training approach bypassing the need for experimentally acquiring diverse training datasets with accessible ground truth. To facilitate strongly scattering feature recovery, we generate simulated object volumes using a fast and accurate Split-Step Non-Paraxial (SSNP) multiple-scattering IDT model [12] (Fig. 1(b)). We promote model generalizability by generating training data volumes from diverse natural images from multiple open-source databases. Using the single-scattering based 3D estimates as the DNN's input, we train the network to enhance the 3D RI recovery. Second, our DNN, termed IDTnet, features a lightweight 2D "U-Net" structure to perform 3D reconstruction (Fig. 1(c)). We achieve efficient learning by feeding five consecutive axial slices selected randomly from each larger object volume as a multi-channel input and predict only the central object slice (Fig. 1(c)). We show this approach efficiently encodes the depth and diffraction information and suppresses missing-cone-induced artifacts in a scalable and computational efficient manner. In combination, we show that IDTnet can be generalized to enhance the reconstructions on both weakly and strongly scattering complex 3D biological samples.

We experimentally demonstrate IDTnet's superior generalization capability by recovering live epithelial buccal cells, *C. elegans* worm samples (Fig. 1(d)), and fixed algae samples acquired using different experimental setups. We further highlight IDTnet's robustness by making time-series predictions on a living, dynamic worm. To quantitatively assess the reliability of IDTnet's reconstructions, we first compare IDTNet with SSNP reconstructions on both weakly and strongly scattering samples. We subsequently adapt an image-space based evaluation procedure by feeding the network predicted RI into the multiple-scattering SSNP forward model and comparing the calculated intensity and experimental measurements. Even for "unseen" illumination angles that are unused during model training or prediction, the calculated intensities from our predictions match well with experimental data. We show a 2-3× error reduction using IDTnet over the linear model. Our result highlights that leveraging large-scale multiple-scattering modeling can obviate major overhead in physical data acquisition and train a reliable, highly generalizable DL model for computational 3D phase imaging.

2. IDTnet design

To optimally combine aIDT's reconstruction pipeline with a DL model, the network architecture must satisfy four key properties: 1) preserve the modality's speed, 2) provide flexible volume size recovery, 3) suppress scattering and missing-cone induced artifacts, and 4) recover both weak and strongly scattering samples. This requires IDTnet to learn efficient object predictions robust to scattering strengths without sacrificing aIDT's fast acquisition speed. Here, we satisfy these parameters through the tailoring of our datasets and network training procedure.

Our data generation, training process, and network architecture are summarized in Fig. 1. Generating a DNN robust for different samples requires a sufficiently diverse training dataset. We achieve this requirement by leveraging an accurate and computationally efficient multiple-scattering model [12] and *simulated* object volumes from diverse open-source natural image datasets. As shown in [29], training with diverse images generally improves a network's generalization capability. Since capturing many diverse biological samples with paired ground truth for training is infeasible, this use of open-source datasets provides readily accessible high-entropy images to train a generalizable network for complex biological specimen recovery.

Randomly selecting natural images from these datasets, we stack the images into 3D volumes and assign random RI values to create weak and strongly scattering samples. We then apply our SSNP beam-propagation model [12] to generate intensity images with the same physical parameters as the experimental aIDT setup (NA=0.65) (Fig. 1(b)). Using SSNP provides easy, rapid generation of a large diverse dataset for training IDTnet while the use of natural images helps prevent overfitting [29]. With the simulated intensity images, we then compute linear model

Optics EXPRESS

approximants using the aIDT model, which become the network inputs for training. Details on this simulation process are found in Supplement 1.

To reconstruct samples with large depth ranges, we introduce a randomized sampling procedure to the training process. Due to the undersampling of the object's 3D Fourier information in aIDT, the linear reconstruction exhibits anisotropic, axially varying, and object-dependent missing-cone artifacts throughout the volume. To suppress these 3D artifacts, we input randomly selected five slice subvolumes from each extended object volume to IDTnet and predict only the subvolume's central slice using a modified 2D U-Net (Fig. 1(c)). This design presents 3D information as a multi-channel input (i.e. feature maps) from which IDTnet can extract 3D information with an efficient and easily trainable 2D network. By randomizing the subvolume input during training, IDTnet learns object recovery based on the subvolume's inter-slice relations and generalizes to larger volumes. Details on the training procedure and IDTnet architecture are provided in Supplement 1. Once trained, IDTnet performs slice-wise reconstructions of an object volume of arbitrary size. This multiple-scattering simulator trained network is directly generalizable to experimental data such as the *C. elegans* worm in Fig. 1(d).

3. Results

3.1. Weakly scattering sample recovery

We first apply IDTnet to weakly scattering epithelial buccal cells in aqueous media ($n_0 = 1.33$) (Fig. 2) measured using a 40×, 0.65NA objective. Figure 2(a) compares the single-scattering (lower left) and IDTnet reconstructions (upper right) as depth-coded projections. Figure 2(b) shows a volume rendering of a cell cluster from the IDTnet reconstruction with maximum intensity projections (MIP) in XY, YZ, and XZ. Comparisons for different reconstructed slices are shown in Fig. 2(c). Figure 2(d) compares XZ MIP through the reconstructed 3D Fourier space by the linear model and IDTNet.

Compared with the single-scattering based linear model, IDTnet's result shows significant artifact suppression and object feature enhancement. Shown in the projection of Fig. 2(a) and the cross-sections of Fig. 2(c), the linear reconstruction generates strong missing-cone artifacts corrupting the reconstructed object features. While lateral cross-section images show cell edges (Fig. 2(c), white brackets) and native bacteria (Fig. 2(c), blue arrows) are visible with the linear estimate, the artifacts reduce feature visibility and confound the true morphology of cellular structures (Fig. 2(c), white boxes). In contrast, IDTnet maintains or improves recovery of these features (Fig. 2(c), white circles) and removes the missing-cone artifacts as shown in the XZ and YZ cross-sections. IDTNet shows occasional high-frequency feature loss (Fig. 2(c), green arrows). We attribute the feature loss to inherent limitations of IDTNet in correcting high-frequency errors from small RI variations obscured by larger missing cone artifacts, as discussed in Supplemental 1. Despite these factors, we still observe drastically improved performance in object recovery for IDTNet over the linear model.

The IDTnet's improvements are also evident in the reconstructed 3D Fourier coverage shown as the XZ MIP in Fig. 2(d). IDTnet provides better filled and smoother Fourier information, especially around the origin, which in turn suppresses the missing-cone artifacts in the real space. Furthermore, the similar recovered RI values of the cell's edges and bacteria between the linear and IDTnet reconstructions suggest the network recovers the correct quantitative values for the weakly scattering sample. We further evaluate this quantitative recovery in sections 3.5 and 3.6. These results provide evidence that IDTnet generalizes well to weakly scattering, experimentally measured biological samples with complex cellular morphology.

Optics EXPRESS

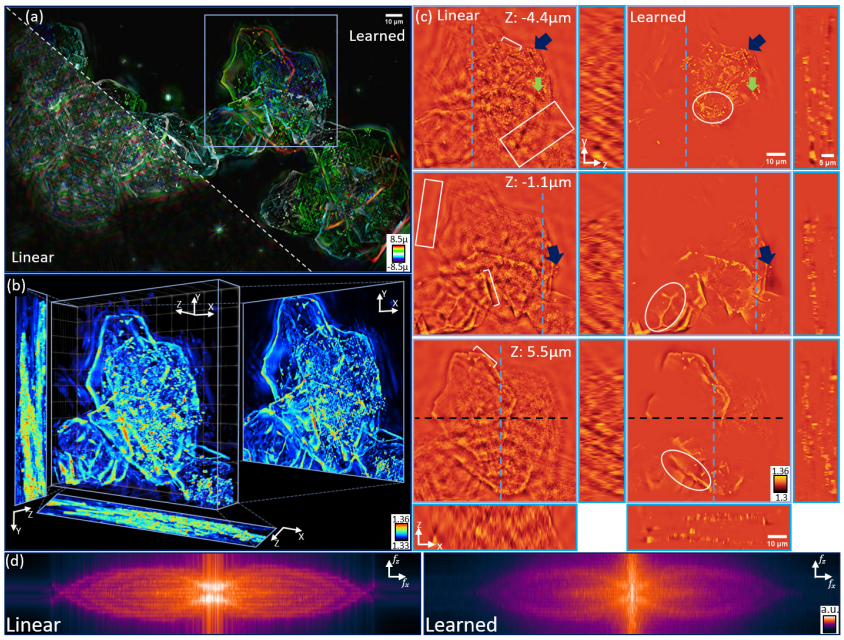

Fig. 2. Results of weakly scattering epithelial buccal cells. (a) Color-coded depth projections of reconstructions from the IDT linear model (lower left) and IDTnet (upper right). (b) 3D rendering of the IDTnet reconstruction from the blue subregion in (a). (c) RI slice reconstructions, XZ, and YZ cross-sections from the linear (left) and IDTnet (right) reconstructions. White boxes show poor depth sectioning in the linear model that is improved with IDTnet; blue arrows highlight native bacteria features; white circles show enhanced cell edges in the IDTnet result; white brackets highlight cell edges; green arrows indicate high-frequency feature loss regions. (d) XZ MIP of the reconstructed 3D Fourier coverage for the linear model (Left) and IDTnet (Right). IDTnet expands axial Fourier recovery near the origin that helps reduce missing-cone artifacts.

3.2. *IDTnet generalizes to different experimental setups*

We further evaluate IDTnet's ability to generalize to different IDT setups and sample media. In Fig. 3, We evaluate a *spirogyra* sample in an aqueous medium (n_0 = 1.33) measured with a 10×, NA = 0.25 objective in an aIDT setup different from the IDTnet training. This study highlights IDTnet's ability to handle new imaging setups that carry stronger missing-cone artifacts than the training data. Comparisons of the linear and IDTnet reconstructions as color-coded depth projections are shown in Fig. 3(a) with a rendering of the IDTnet-recovered volume in Fig. 3(b). Figure 3(c) compares individual recovered RI slices of both reconstruction methods, and Fig. 3(d) compares MIP XZ cross-sections through the 3D Fourier spectra of the linear and learned *spyrogyra* reconstructions. Despite the different imaging setup, the IDTnet predictions show nearly artifact-free recovery of the algae sample with enhanced depth sectioning providing clear visualizations of the algae fragments (Fig. 3(b), XZ and YZ MIP). This is further highlighted in Fig. 3(c), where white arrows show the *spirogyra*'s helical structure is now well separated into different axial planes without the missing-cone artifacts. This separation is also visible in the color-coded depth projection of the algae helical structure in Fig. 3(a).

While IDTNet's recovered volumes show general improvements, fine features from the spyrogyra can be lost (Fig. 3(c) blue arrows). We attribute these losses to an inherent, sampling-dependent bias within the trained IDTNet. The original aIDT system oversamples the object's

Optics EXPRESS

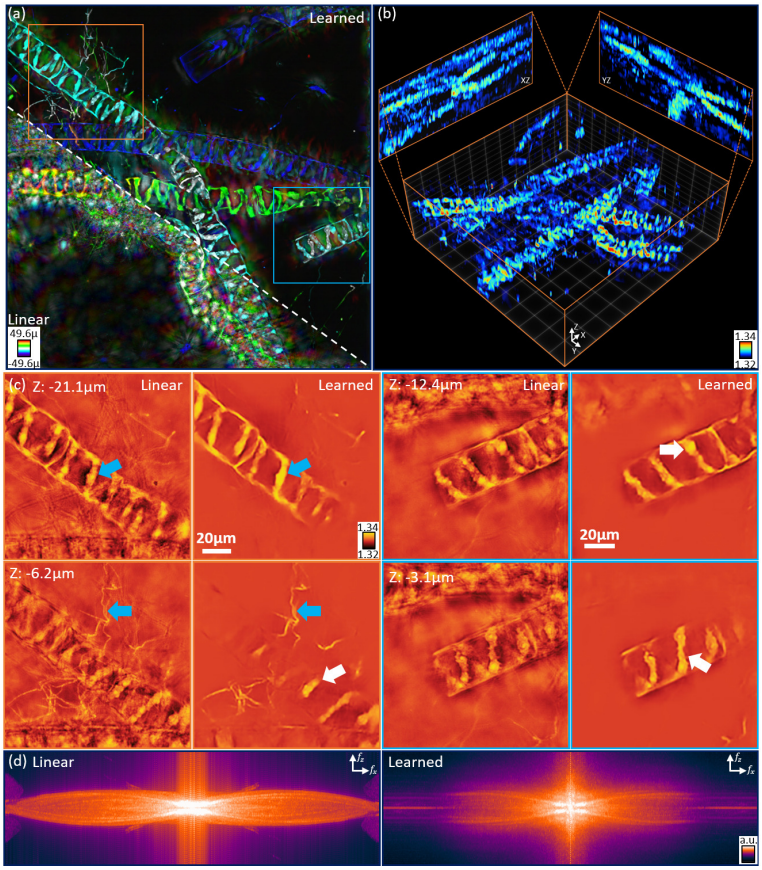

Fig. 3. Reconstructions of *Spirogyra* in a low-NA aIDT setup (NA=0.25). (a) Depth-coded projections comparing the linear and IDTnet reconstructions spanning 100μm. (b) 3D rendering of the IDTnet reconstruction with XZ and YZ MIPs. The artifact removal and strong feature recovery provide enhanced depth sectioning throughout the volume. (c) RI slice-wise comparisons of the linear and IDTnet reconstructions. The IDTnet's better depth sectioning removes defocused sample features providing greater visibility of in-focus structures (white arrows) and more faithful RI recovery than the linear model. Blue arrows highlight areas of high-frequency feature loss. (d) Linear (Left) and Learned (Right) XZ MIP through the reconstructed 3D Fourier spectra.

bandwidth by 2.4× while the low-NA system exhibits slight undersampling at 1.6× the object's bandwidth. This finer sampling better confines the object bandwidth, observed when comparing Fig. 2(d) and Fig. 3(d), and imposes an implicit constraint on the maximum pixel size of all recoverable features in the object space. Because IDTNet learns feature recovery and missing cone removal based on the *relative* size information provided within the input subvolume slices, this network tends to remove lateral frequency features below this learned minimum pixel size. This limitation, along with IDTNet's sensitivity loss to small and high-frequency RI changes discussed in Supplement 1, lead to this resolution loss in the *spyrogyra* sample. In spite of these limitations, IDTNet still improves the recovery of low spatial-frequency features and removes missing cone artifacts in the *spyrogyra* sample, as evident in the stronger spectral content near the learned reconstruction's origin (Fig. 3(d)). Furthermore, our results suggest that other imaging systems matching the relative sampling conditions of the primary aIDT system could still utilize

Optics EXPRESS

the trained IDTNet without any resolution loss. These results show that IDTnet still provides good generalization to different imaging setup configurations and different sample types. Additional results shown in Supplement 1 demonstrate similar improvements are seen on diatom algae under high-resolution aIDT systems with different imaging media.

3.3. Strongly scattering sample recovery

To evaluate IDTnet's capabilities on stronger scattering samples, we apply the trained network to a *C. elegans* worm sample in Fig. 4. A 3D rendered reconstruction is shown in Fig. 4(a) with the central RI slice in Fig. 4(b) and zoom-in regions of tissue structures including lipid droplets in Fig. 4(c), the terminal pharyngeal bulb with grinder in Fig. 4(d), and the buccal cavity in Fig. 4(e). Immediately apparent in the IDTnet's prediction is the enhanced clarity and RI contrast of the worm's tissue structures. IDTnet removes missing-cone artifacts, improves the RI prediction, and recovers worm features across the entire segment in Fig. 4(b) with improved lipid droplet visibility in Fig. 4(c). Our IDTnet result shows fine, continuous features are recoverable through the volume such as the grinder from the worm's digestive tract (Fig. 4(d)) and the pharyngeal epithelium (Fig. 4(e)). While these features are also recovered using the linear model, IDTnet significantly improves the depth sectioning and clarity of the reconstruction. This is particularly evident with the buccal cavity centered at 5μm whose missing-cone artifacts have been nearly completely removed from the central slice (Fig. 4(b) purple square, 4(e)). IDTNet does show limited recovery capabilities in the worm's mid-section, where features appear to be reduced in value or removed entirely (Fig. 4(b), red box). Within this region, the linear model's reconstruction exhibits numerous scattering features that are difficult to separate as in-focus features or missing cone artifacts from adjacent slices. As discussed in Supplement 1, this condition can cause failures in IDTNet to resolve high-resolution features. Despite this loss, these results still highlight IDTnet's capabilities in improving the linear model's reconstruction and generalizing to strongly scattering multi-cellular organisms.

3.4. Dynamic sample recovery

A key objective for combining aIDT with IDTnet is to maintain fast reconstruction of complex samples for imaging living dynamic biological samples. To demonstrate this capability, we apply IDTnet to *C. elegans* time-series measurements from [5]. Results are shown for specific time points in Fig. 5 and the video reconstruction is provided in Visualization 1.

Figure 5 highlights the wealth of information recovered by IDTnet from the complex, dynamic biological sample. IDTnet provides clear visualizations of the worm's movement through the entire 2.5-minute measurement period shown in Fig. 5(a). IDTnet enhances the depth sectioning across all temporal measurements, which is particularly evident in the well-separated features in the color-coded depth projection of Fig. 5(b). During this time period, IDTnet provides recovery of the digestive tract (white brackets) and lipid droplets (white circles) in Fig. 5(c–e) with complex internal organ features clearly recovered in Fig. 5(d) (black oval). Figure 5(e) shows new feature recovery previously outside of the initial time-series measurement's FOV including muscle walls (black bracket) and the worm's vulva (white arrow).

IDTnet's enhanced recovery of such features in temporal data highlights its utility for dynamic imaging. Despite training on simulated natural images, the IDTnet's generalization recovers complex biological features consistently in time. This result opens the possibility for IDTnet's application to studying temporal dynamics of biological samples with significantly enhanced feature recovery over conventional model-based IDT techniques.

3.5. Reconstruction quality assessment

To evaluate IDTNet's reliability in reconstructing experimental data, we compare reconstructions of biological samples using the aIDT linear model, SSNP [12], and IDTNet in Fig. 6. We first

Optics EXPRESS

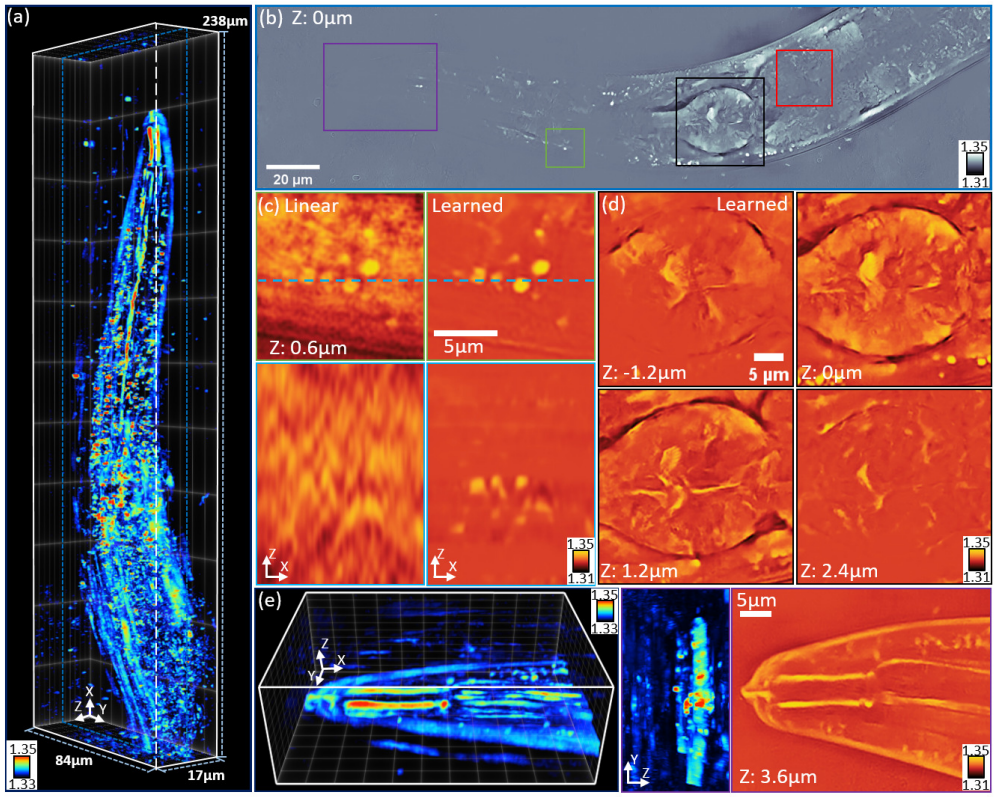

Fig. 4. Results of strongly scattering *C. elegans* worm sample. (a) 3D rendering of the full volume worm reconstruction. (b) Central slice reconstruction with outsets of lipid droplets (green), pharyngeal bulb (black) and buccal cavity (purple). Worm tissue structure loss is outlined in red. (c) Outset comparing the linear and IDTnet reconstructions of lipid droplets in an XY RI slice and XZ cross-sections. (d) Consecutive Z slices of the terminal pharyngeal bulb with clear recovery of the grinder organ. High-resolution features are recovered with IDTnet at 1.2μm axial steps through the volume. (e) 3D rendering, YZ MIP, and RI XY slice of the worm's buccal cavity. The buccal cavity becomes well-resolved at the 5μm axial plane with IDTNet.

evaluate the weak-scattering buccal cell reconstructions in Fig. 6(a) with corresponding Fourier spectra in Fig. 6(b). The linear and SSNP models exhibit missing-cone artifacts evident in the YZ cross-sections of Fig. 6. These artifacts are also visible from the missing Fourier information around the origin of Fig. 6(b). These missing-cone artifacts are reduced but not removed for SSNP through the use of iterative reconstruction with a total variation regularizer enforcing a piece-wise constant RI sample distribution [12]. This result leads to slightly stronger spectral content around the origin of the SSNP result spectra (Fig. 6(b)). By comparison, these features are largely removed with IDTNet to provide better visibility of the cell features. IDTNet's reconstruction shows good agreement with SSNP's recovery of the buccal cells in-focus features. Comparing the RI values reported between SSNP and IDTNet, we observe nearly equivalent, slightly higher RI values recovered with IDTNet compared to SSNP. We attribute this RI increase to IDTNet's dependency on the linear model's RI as input because the linear reconstruction also shows higher RI than the SSNP model. This result indicates that IDTNet can provide equivalent 3D RI distributions to SSNP with fewer missing-cone artifacts on weakly scattering samples. Evaluating the recovered IDTNet Fourier spectrum in Fig. 6(b), we observe strong

Optics EXPRESS

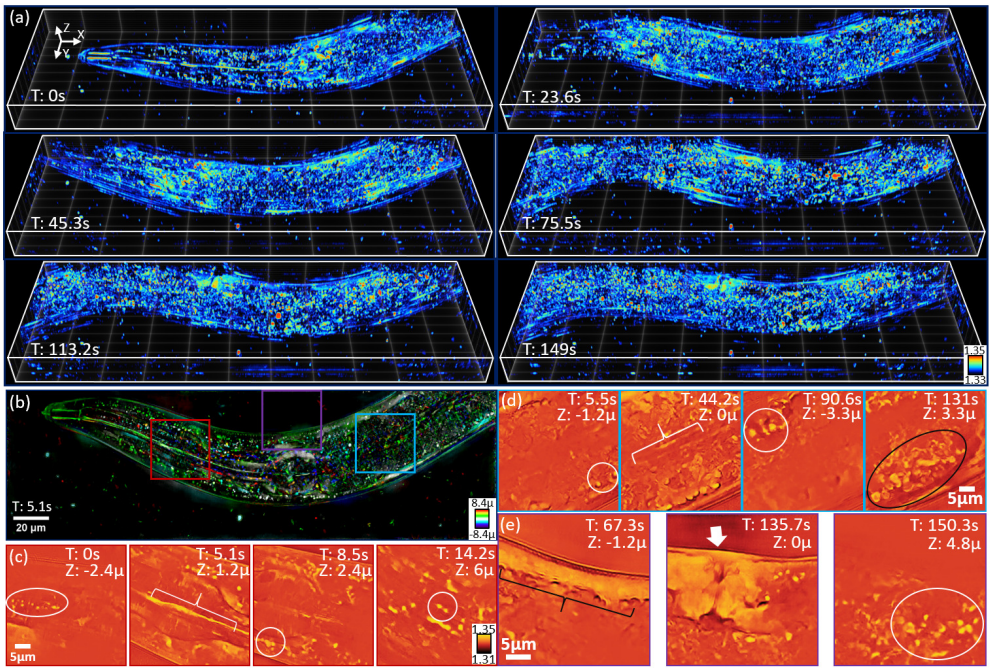

Fig. 5. Results of a dynamic *C. elegans* worm. (a) 3D renderings of IDTnet reconstructions across different time points of an aIDT longitudinal measurement. (b) Color-coded depth projection of the reconstruction at 5.1 seconds. IDTnet reconstruction shows minimal missing-cone artifacts and clear feature recovery. (c)-(e) Zoom-ins highlighting recovered *C. elegans* organs and tissues during video reconstruction. White circles: lipid droplets and high-resolution circular structures; white brackets: intestinal tract, black brackets: worm muscle wall; black circles: complex tissue features; white arrow: vulva and reproductive organs. Additional worm dynamics are shown in Visualization 1.

spectral content along the axial direction near the origin suggesting the recovery of low spatial frequencies and missing-cone removal.

Next, we compare reconstruction results on the strongly scattering *C. elegans* worm segment in Fig. 6(c),(d). First comparing the linear and SSNP models, we observe similar reconstructions with significant missing-cone artifact corruption (Fig. 6(c), YZ cross-sections). The Fourier XZ MIP also shows frequency loss near the origin indicating strong missing cone artifacts (Fig. 6(d)). Comparing IDTNet and SSNP, we see the network can still easily remove missing-cone artifacts to provide improved visibility of the worm's structures (Fig. 6(c)). Lipid droplets and higher RI structures are well-preserved with IDTNet as highlighted in the white circles of Fig. 6(c), but lower RI structures appear to be reduced or removed with IDTNet as well (Fig. 6(c), black squares). Due to the corrupted multiple-scattering model, we cannot easily determine whether IDTNet properly removes reconstruction artifacts or incorrectly removes weak RI features due to low sensitivity (Supplement 1).

3.6. *Prediction reliability analysis*

While IDTnet's predictions on simulated objects show close agreement with the ground truth (Supplement 1) and our experimental results show similar feature recovery to multiple-scattering models, the network's reliability for experimental data remains an outstanding question. Several factors including noise, illumination angle, source homogeneity, imaging wavelength, and

Optics EXPRESS

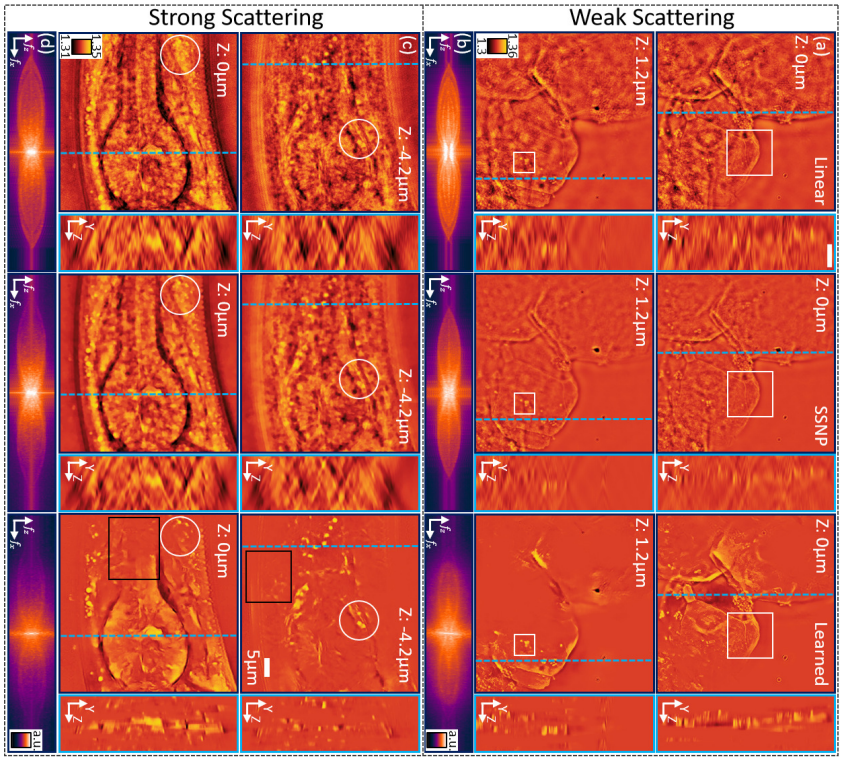

Fig. 6. Reconstruction comparison between the aIDT linear (Left), multiple-scattering SSNP (Middle), and IDTNet (Right) methods. (a) Weakly scattering epithelial buccal cell XY reconstructions (Dark blue) at $0\mu m$ (Top) and $1.2\mu m$ (Bottom) with corresponding YZ cross-sections (Light blue). White squares highlight similar recovered cell and bacteria structures across all models. (b) 3D Fourier spectrum XZ MIP of the linear, SSNP, and learned buccal cell reconstructions. (c) Strongly scattering *C. elegans* XY reconstructions (dark blue) at $-4.2\mu m$ (Top) and $0\mu m$ (Bottom) with YZ cross-sections (Light blue). White circles show recovered lipid droplets. The black square highlights potential feature removal or artifact correction with IDTNet. (d) 3D Fourier spectrum XZ MIP of the linear, SSNP, and learned *C. elegans* worm section reconstructions.

aberrations could result in additional and/or stronger reconstruction artifacts. These variations can generate unreliable object predictions that cannot be evaluated due to the lack of ground-truth information in experimental measurements. Understanding the network reliability is crucial for applying this pipeline to biological studies where artificial features could cause mis-classification of features and/or disease mis-diagnoses. To mitigate this issue, we develop an image space analysis metric to evaluate the reliability of IDTnet's predictions on experimental data.

Our image space metric expands upon the method proposed in [21]. Here, we implement the SSNP forward model [12], assuming perfectly calibrated illumination angles, LED intensities, and no pupil aberrations in the system, to generate intensity images from the reconstructed sample volumes that are then compared with the experimental measurements. With a sufficiently rigorous simulator, deviations between the simulated and experimental images can be related to errors in the predicted RI and structure of the recovered volumes. Here, we compare intensity images from illumination angles used in our reconstruction (seen) and those unused (unseen) in our models and training. Our idea is that if IDTnet "overfits" to the features recovered using seen

Optics EXPRESS

illuminations, the simulated images from *unseen* illuminations would exhibit increased error from network hallucinated features.

We first evaluate epithelial buccal cells across 24 illuminations as shown in Fig. 7(a) and (c). Here we used the main setup in Fig. 1(a) for acquisition described in detail in Supplement 1. To further evaluate the effect of experimental setup variations with this metric, we compute the metrics on diatom algae samples measured with a different LED array setup with 120 illuminations in Fig. 7(b) and (d). We apply the SSNP model for calculating the intensity images from the linear (blue) and IDTnet (orange) estimates and compare them with experimental measurements using pixel-wise absolute error maps (Fig. 7(a,b)) and the mean squared error (MSE) of each image at different illumination angles (Fig. 7(c,d)). Seen illuminations are in magenta with green LEDs highlighting the specific illuminations used for the images in Fig. 7(a,b).

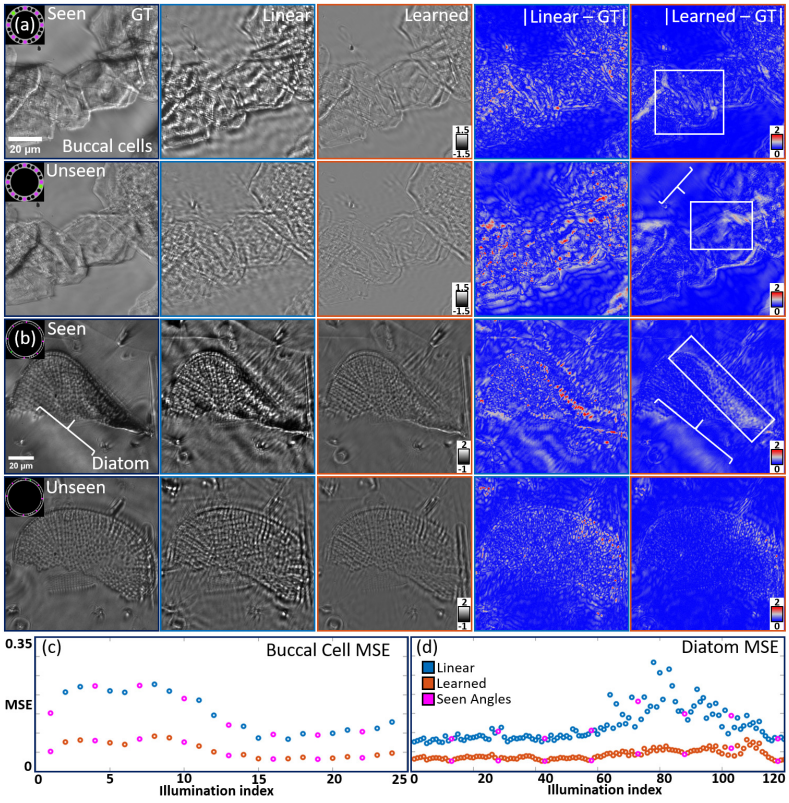

Fig. 7. Quantitative reliability analysis of network's predictions. We compare intensity images computed using the multiple-scattering model from the linear and learned object and compare them with the experimentally measured intensity images for (a) epithelial buccal cells and (b) a diatom sample. IDTNet's reconstructions show closer intensity image contrast and lower errors than intensity images computed from the IDT linear model-based reconstructions. The MSE between the computed and experimental measured images across both seen and unseen angles for (c) epithelial buccal cells and (d) a diatom sample. The results show consistently lower error using the learned model regardless whether the illuminations are used in the model training.

In both the cells and diatom sample, the IDTnet intensity images show closer contrast and lower error to the experimental data than the linear model. This result is consistent regardless of whether the illumination angle was used for the reconstruction (Fig. 7(a,b), Seen vs. Unseen). The main differences between the IDTnet and experimental images appears as low spatial frequency

Optics EXPRESS

loss (Fig. 7(a,b), white boxes) creating "flatter" images and background inhomogeneities in the experimental measurements (Fig. 7(a,b), white brackets). These errors are attributed to the linear approximant input lacking low spatial frequency information and the LED generating non-ideal plane wave illumination, respectively.

Evaluating the image-wise MSE in Fig. 7(c,d) further confirms IDTnet recovers the underlying sample volume. Plotted as a function of the illumination index, the MSE for the seen and unseen illumination angles show no substantial difference in error for both samples. Across all illuminations, the images from the IDTnet reconstructed volume show consistently lower error than the linear model, which is partially attributed to the removal of missing-cone artifacts. The dominant error in Fig. 7(d) instead results from residual illumination angle misalignments after implementing the calibration procedures of [5]. The misalignment is most visible in illuminations 70-90 for the diatom sample where the illuminations are close to the objective's 0.65 NA cut-off and are difficult to calibrate. While both linear and learned estimates exhibit elevated MSE values within this region, IDTNet shows smaller variations and provides consistently lower error across all illumination angles. This result suggests that IDTNet still provides a better object estimate than the linear aIDT reconstruction model.

4. Discussion and conclusion

Our results highlight the significant potential of DL in computational 3D phase imaging. With only simulated objects, we showed a lightweight 2D network can be trained following approximant-guided learning methods to recover the 3D phase of complex biological samples. Our IDTnet corrects not only missing-cone artifacts but also shows reconstruction accuracy improvements compared to linear models in simulation. We demonstrated improved volumetric RI recovery on experimental data acquired using several IDT setups. Finally, we showed that IDTnet can be readily applied to recover high-quality volumetric reconstructions of dynamic biological samples.

A main limitation of this approximant-guided learning approach is the network's reliance on the initial model-based object estimate for feature prediction. The single-scattering IDT model neglects contributions from the nonlinear scattering signals that become significant for strongly scattering samples [13]. This limits the information available for the network to learn from, which contributes to the network's failure to recover certain object features outside the linear model's Fourier support. This limitation could be alleviated through the incorporation of multiple-scattering model-based approximants as network inputs [28].

Our learned IDT approach holds promise for improving the image quality in low-cost optical setups. Recent works have developed low-cost, open-source optical imaging setups enabling affordable multi-modal imaging in a push for the "democratization" of science to the general public [30,31]. Particularly, recent work has shown that aIDT can be implemented in such setups enhancing both the capabilities of these platforms and accessibility to the imaging modality [30]. By using cheaper optical components, however, the reconstructed volume can degrade in quality. Because our learned approach generalizes well to different optical setups, the use of this lightweight framework on low-cost setups could drastically improve the reconstructions and potentially be implemented on portable computing devices for real-time processing [31,32]. In addition, IDTnet could also be extended to improve the multiplexed IDT for high-speed imaging [6]. This will bypass the tradeoff between reconstruction quality and acquisition speeds suffered by model-based reconstructions [6]. Furthermore, IDT has recently been integrated into a mid-infrared photothermal microscope enabling chemically specific 3D phase imaging [33]. Our IDTnet could be integrated in this setup to further improve the sensitivity and accuracy of the chemically specific phase reconstruction results.

Finally, the generalization of IDTnet is achieved by leveraging our multiple-scattering simulator, which shows the power of simulation-based training for applying DL for imaging in complex media applications. Similar simulator-based DL strategies have also been demonstrated recently

Optics EXPRESS

in holographic 3D imaging [34] and computational fluorescence microscopy [35,36]. Further development of this framework may impact many other emerging areas, such as such as imaging and light manipulation in complex media [37] and optical computing [38].

Funding. National Science Foundation (1840990, 1846784).

Acknowledgment. The authors thank Boston University Shared Computing Cluster for providing the computational resources.

Disclosures. The authors declare no conflicts of interest.

Data Availability. The experimental data is acquired from related publications [4,5]. The trained IDTNet code and example data can be found at our Github repository [39].

Supplemental document. See Supplement 1 for supporting content.

References

1. D. Jin, R. Zhou, Z. Yaqoob, and P. So, "Tomographic phase microscopy: Principles and applications in bioimaging," J. Opt. Soc. Am. B **34**(5), B64–B77 (2017).
2. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," Nat. Photonics **12**(10), 578–589 (2018).
3. J. A. Rodrigo, J. M. Soto, and T. Alieva, "Fast label-free microscopy technique for 3d dynamic quantitative imaging of living cells," Biomed. Opt. Express **8**(12), 5507–5517 (2017).
4. R. Ling, W. Tahir, H. Y. Lin, H. Lee, and L. Tian, "High-throughput intensity diffraction tomography with a computational microscope," Biomed. Opt. Express **9**(5), 2130–2141 (2018).
5. J. Li, A. C. Matlock, Y. Li, Q. Chen, C. Zuo, and L. Tian, "High-speed in vitro intensity diffraction tomography," Adv. Photonics **1**(06), 1 (2019).
6. A. Matlock and L. Tian, "High-throughput, volumetric quantitative phase imaging with multiplexed intensity diffraction tomography," Biomed. Opt. Express **10**(12), 6432 (2019).
7. V. Nandakumar, L. Kelbauskas, K. F. Hernandez, K. M. Lintecum, P. Senechal, K. J. Bussey, P. C. Davies, R. H. Johnson, and D. R. Meldrum, "Isotropic 3d nuclear morphometry of normal, fibrocystic and malignant breast epithelial cells reveals new structural alterations," PLoS One **7**(1), e29230 (2012).
8. Y. Park, M. Diez-Silva, G. Popescu, G. Lykotrafitis, W. Choi, M. S. Feld, and S. Suresh, "Refractive index maps and membrane dynamics of human red blood cells parasitized by plasmodium falciparum," Proc. Natl. Acad. Sci. **105**(37), 13730–13735 (2008).
9. P. A. Sandoz, C. Tremblay, S. Equis, S. Pop, L. Pollaro, Y. Cotte, G. F. Van Der Goot, and M. Frechin, "Label free 3d analysis of organelles in living cells by refractive index shows pre-mitotic organelle spinning in mammalian stem cells," BioRxiv, 407239 (2018).
10. J. Lim, A. B. Ayoub, E. E. Antoine, and D. Psaltis, "High-fidelity optical diffraction tomography of multiple scattering samples," Light: Sci. Appl. **8**(1), 1–12 (2019).
11. M. Chen, D. Ren, H.-Y. Liu, S. Chowdhury, and L. Waller, "Multi-layer born multiple-scattering model for 3d phase microscopy," Optica **7**(5), 394–403 (2020).
12. J. Zhu, H. Wang, and L. Tian, "High-fidelity intensity diffraction tomography with a non-paraxial multiple-scattering model," Opt. Express **30**(18), 32808–32821 (2022).
13. S. Chowdhury, M. Chen, R. Eckert, D. Ren, F. Wu, N. Repina, and L. Waller, "High-resolution 3d refractive index microscopy of multiple-scattering samples from intensity images," Optica **6**(9), 1211–1219 (2019).
14. L. Tian and L. Waller, "3d intensity and phase imaging from light field measurements in an led array microscope," Optica **2**(2), 104–111 (2015).
15. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," Optica **2**(6), 517–522 (2015).
16. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," Optica **6**(8), 921–943 (2019).
17. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," Optica **4**(9), 1117–1125 (2017).
18. Y. Xue, S. Cheng, Y. Li, and L. Tian, "Reliable deep-learning-based phase imaging with uncertainty quantification," Optica **6**(5), 618–629 (2019).
19. H. Chen, L. Huang, T. Liu, and A. Ozcan, "Fourier Imager Network (FIN): A deep neural network for hologram reconstruction with superior external generalization," Light: Sci. Appl. **11**(1), 254 (2022).
20. 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).
21. J. Lim, A. B. Ayoub, and D. Psaltis, "Three-dimensional tomography of red blood cells using deep learning," Adv. Photonics **2**(02), 1 (2020).
22. Z. Wu, Y. Sun, A. Matlock, J. Liu, L. Tian, and U. S. Kamilov, "Simba: Scalable inversion in optical tomography using deep denoising priors," IEEE J. Sel. Top. Signal Process. **14**(6), 1163–1175 (2020).

Optics EXPRESS

23. K. C. Zhou and R. Horstmeyer, "Diffraction tomography with a deep image prior," Opt. Express **28**(9), 12872–12896 (2020).
24. I. Kang, A. Goy, and G. Barbastathis, "Dynamical machine learning volumetric reconstruction of objects' interiors from limited angular views," Light: Sci. Appl. **10**(1), 1–21 (2021).
25. 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 (2017).
26. R. Liu, Y. Sun, J. Zhu, L. Tian, and U. S. Kamilov, "Recovery of continuous 3d refractive index maps from discrete intensity-only measurements using neural fields," Nat. Mach. Intell. **4**(9), 781–791 (2022).
27. A. Saba, C. Gigli, A. B. Ayoub, and D. Psaltis, "Physics-informed neural networks for diffraction tomography," arXiv, arXiv:2207.14230 (2022).
28. A. Goy, G. Rughoobur, S. Li, K. Arthur, A. I. Akinwande, and G. Barbastathis, "High-resolution limited-angle phase tomography of dense layered objects using deep neural networks," Proc. Natl. Acad. Sci. **116**(40), 19848–19856 (2019).
29. M. Deng, S. Li, Z. Zhang, I. Kang, N. X. Fang, and G. Barbastathis, "On the interplay between physical and content priors in deep learning for computational imaging," Opt. Express **28**(16), 24152–24170 (2020).
30. B. Diederich, R. Lachmann, S. Carlstedt, B. Marsikova, H. Wang, X. Uwurukundo, A. S. Mosig, and R. Heintzmann, "A versatile and customizable low-cost 3d-printed open standard for microscopic imaging," Nat. Commun. **11**(1), 5979 (2020).
31. J. T. Collins, J. Knapper, J. Stirling, J. Mduda, C. Mkindi, V. Mayagaya, G. A. Mwakajinga, P. T. Nyakyi, V. L. Sanga, D. Carbery, L. White, S. Dale, Z. J. Lim, J. J. Baumberg, P. Cicuta, S. McDermott, B. Vodenicharski, and R. Bowman, "Robotic microscopy for everyone: the openflexure microscope," Biomed. Opt. Express **11**(5), 2447–2460 (2020).
32. B. Diederich, P. Then, A. Jügler, R. Förster, and R. Heintzmann, "cellstorm—cost-effective super-resolution on a cellphone using dstorm," PLoS One **14**(1), e0209827 (2019).
33. J. Zhao, A. Matlock, H. Zhu, Z. Song, J. Zhu, B. Wang, F. Chen, Y. Zhan, Z. Chen, Y. Xu, X. Lin, L. Tian, and J.-X. Cheng, "Bond-selective intensity diffraction tomography," Nat. Commun. **13**(1), 7767 (2022).
34. W. Tahir, H. Wang, and L. Tian, "Adaptive 3d descattering with a dynamic synthesis network," Light: Sci. Appl. **11**(1), 42 (2022).
35. Y. Xue, Q. Yang, G. Hu, K. Guo, and L. Tian, "Deep-learning-augmented computational miniature mesoscope," Optica **9**(9), 1009–1021 (2022).
36. K. Yanny, K. Monakhova, R. W. Shuai, and L. Waller, "Deep learning for fast spatially varying deconvolution," Optica **9**(1), 96–99 (2022).
37. S. Yoon, M. Kim, M. Jang, Y. Choi, W. Choi, S. Kang, and W. Choi, "Deep optical imaging within complex scattering media," Nat. Rev. Phys. **2**(3), 141–158 (2020).
38. G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. Miller, and D. Psaltis, "Inference in artificial intelligence with deep optics and photonics," Nature **588**(7836), 39–47 (2020).
39. IDTNet: Multiple-scattering simulator-trained neural network for intensity diffraction tomography, https://github.com/bu-cisl/IDTNet.