ADVANCED PHOTONICS

Three-dimensional tomography of red blood cells using deep learning

Joowon Lim,* Ahmed B. Ayoub, and Demetri Psaltis
École Polytechnique Fédérale de Lausanne, Optics Laboratory, Lausanne, Switzerland

Abstract. We accurately reconstruct three-dimensional (3-D) refractive index (RI) distributions from highly ill-posed two-dimensional (2-D) measurements using a deep neural network (DNN). Strong distortions are introduced on reconstructions obtained by the Wolf transform inversion method due to the ill-posed measurements acquired from the limited numerical apertures (NAs) of the optical system. Despite the recent success of DNNs in solving ill-posed inverse problems, the application to 3-D optical imaging is particularly challenging due to the lack of the ground truth. We overcome this limitation by generating digital phantoms that serve as samples for the discrete dipole approximation (DDA) to generate multiple 2-D projection maps for a limited range of illumination angles. The presented samples are red blood cells (RBCs), which are highly affected by the ill-posed problems due to their morphology. The trained network using synthetic measurements from the digital phantoms successfully eliminates the introduced distortions. Most importantly, we obtain high fidelity reconstructions from experimentally recorded projections of real RBC sample using the network that was trained on digitally generated RBC phantoms. Finally, we confirm the reconstruction accuracy using the DDA to calculate the 2-D projections of the 3-D reconstructions and compare them to the experimentally recorded projections.

Keywords: optical diffraction tomography; three-dimensional imaging; machine learning; deep learning; image reconstruction; red blood cell; missing cone problem.

Received Jan. 3, 2020; accepted for publication Mar. 4, 2020; published online Mar. 24, 2020.

1 Introduction

When we look at a three-dimensional (3-D) object in a conventional microscopy, we can only see a two-dimensional (2-D) projection at one time. Therefore, we need more information in order to extract the 3-D shape from the 2-D measurement. If we make a holographic measurement where we record both amplitude and phase, measuring at different z planes is equivalent to a single measurement followed by the digital propagation to multiple planes. Therefore, with coherent detection, a z-scan does not provide extra information compared to the single 2-D recording. Another dimension that can be exploited is the illumination angle θ. The measurements in the x, y, θ dimensions can be converted to the 3-D spatial domain by defining the physical relationship between the illuminating fields at the different angles and the corresponding measurements. However, most of the time, the 2-D measurements are incomplete due to the limited numerical apertures (NAs) of the optics, resulting in an inversion process that is highly ill-posed.

Optical diffraction tomography (ODT) is a 3-D imaging method that utilizes multiple 2-D measurements acquired by changing the angle of illumination. The contrast mechanism in ODT is endogenous index. It, therefore, does not require external labeling. ODT provides 3-D refractive index (RI) distributions[1] that contain morphological and biochemical information, which have been widely used to study various biological samples, which are summarized in recent review papers.[2–5] Under the assumption of weak scattering, multiple 2-D measurements in (x, y, θ) can be directly inverted to yield the 3-D RI information in (x, y, z) using the Wolf transform,[6] which is the transformation that maps the spatial frequencies of the 2-D spectrum of the projections to the spatial frequencies of the 3-D spectrum of the object. However, direct inversion reconstruction methods based on the Wolf transform suffer from the missing cone problem—a consequence of the missing spatial frequencies that are not accessible due to the limited NAs of the optics.[7]

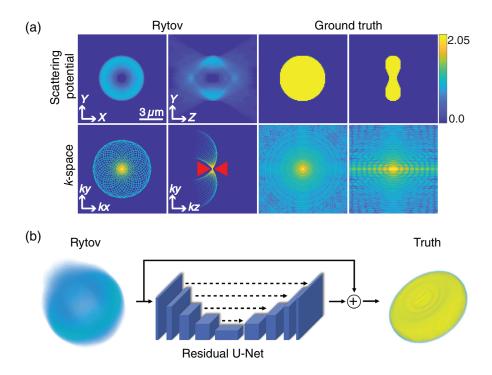*Address all correspondence to Joowon Lim, E-mail: limjoowon@gmail.com

The missing cone problem has been intensively investigated due to its importance.[7,8] Previous approaches are model-based iterative reconstruction (MBIR) schemes, which exploit regularizations based on our prior knowledge, such as non-negativity or sparsity constraints. In other words, MBIR schemes find a solution that is not only consistent with the measurements but also sparse in the regularization domain. The choice of regularization is critical. However, it requires extensive understanding of the characteristics of the forward models, including the degree of ill-posedness intertwined with the characteristics of the samples. This makes the problem challenging.

Recently, deep neural networks (DNNs) have been successful in various optical applications, such as enhancement of the transverse resolution,[9] phase retrieval from intensity measurements,[10,11] digital staining,[12,13] classification/segmentation based on holographic/tomographic measurements,[14–16] and others.[10,17,18] There are some previous demonstrations of the benefits of applying DNNs to the reconstruction of RI values in ODT.[19–21] As far as we know, nobody has succeeded before in using DNNs to reconstruct arbitrary 3-D RI distributions from limited angle measurements taking diffraction and multiple scattering into account.

In this paper, we describe a method based on DNNs to solve the long-standing missing cone problem and demonstrate it using red blood cell (RBC) samples. Despite the potential capacity of DNNs, the lack of the ground truth prevents us from applying DNNs on the ODT reconstruction, unlike other DNN optical imaging applications, such as digital staining or phase retrieval where we can access the target images. Our approach relies on the formation of digital phantoms followed by accurate digital models, which provide the 2-D measurements. The digital 2-D projections are used to form a rough 3-D image of the object using the Wolf transform under the Rytov approximation.[6] By training a DNN with the pairs of images from the Wolf transform and the corresponding digital phantoms, we can learn the distortions introduced due to the incomplete measurements in a data-driven way.

2 Main Idea

We demonstrate the method by using RBC samples that are highly affected by the missing cone problem. The shape of RBCs is flat and biconcave showing narrow dimple regions at the center, which requires high-frequency components along the optical axis to fully resolve the structures.[22,23] In Fig. 1(a), we observe that cross sections of the Rytov reconstruction are underestimated and elongated along the z axis when compared with the corresponding sections of the ground truth. The k-space representation of the Rytov reconstruction can be considered as the low-pass filtered version of the k-space of the ground truth under the weak scattering assumption. Looking at the k-space of the ground truth, the frequency components are more broadly distributed in the k_z axis compared to ones in the k_y axis since the sample is broad in the y axis but has the narrow biconcave shape in the z axis. While high-frequency components are required to fully resolve the thin structure, most of them are lost because they are inaccessible due to the limited NAs as indicated by the red triangles. This results in the high distortions in the final Rytov reconstruction. In general, Rytov reconstructions of RBCs show holes in the middle making it hard to retrieve meaningful information, such as cell volume, surface, and RI values.
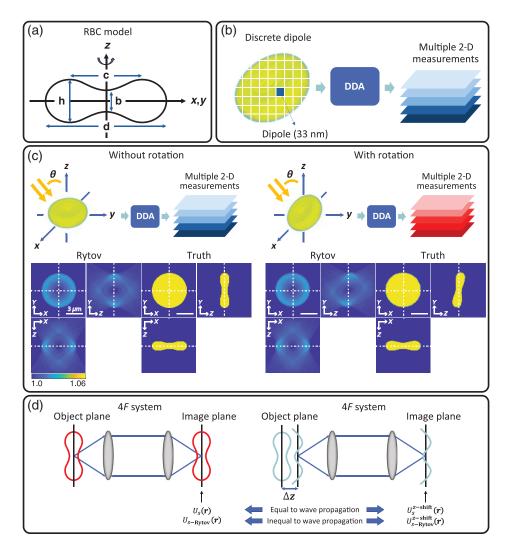

Fig. 1 The missing cone problem and overall scheme of the main idea. (a) Demonstration of the missing cone problem for a single RBC. The left two columns show the Rytov reconstruction and the right two columns show the ground truth. The first row displays the scattering potential, which can be converted to RI distributions, and the second row displays the k-spaces corresponding to the first row. (b) Overall scheme of the network.

A DNN can be trained to recover those missing high frequencies, which are especially important for forming tomograms of RBCs. The network reconstructs the original RBC with the Rytov reconstruction as the initial condition in the training of the DNN, as shown in Fig. 1(b). We refer to the network as TomoNet throughout this paper. The input to the network is the Rytov reconstruction of an RBC and the output is the enhanced image of the same RBC. The input is also relayed directly to the output of the network where it is summed with the correction calculated by the DNN. Therefore, the network learns to extract the difference between the input and the output. In other words, given the low-pass filtered input, the network synthesizes the missing high-pass filtered information using data-driven features from a large number of examples. By combining the low-pass filtered input with the high-pass synthesized output from the network, we can achieve the full resolutions in the transverse and axial planes.

For training, we digitally generated many RBCs that are different in shape and RI value using the RBC model, as shown in Fig. 2(a).[24] For detailed information of the generation of different RBCs, we refer interested readers to Appendix A. Then, each RBC served as a sample for DDA simulations to generate accurate synthetic measurements, as shown in Fig. 2(b).[25,26] A total of 40 uniformly spaced measurements were acquired by scanning on a circular pattern while maintaining a fixed illumination angle of 36 deg with respect to the z axis. The RBC phantoms generated using the model [Fig. 2(a)] originally lie in the xy plane. We implemented the various orientations of RBCs that can occur by randomly rotating each sample in the yz and xy planes. The DDA method was then used to calculate the 2-D projections of each 3-D phantom for each of the 40 illumination angles [Fig. 2(c)]. These calculations were used to form 3-D reconstructions using the Rytov method, which served as the input to the network.[6] Each Rytov reconstruction was paired with the corresponding synthetic RBC that was used to generate

Fig. 2 Dataset generation. (a) RBC model parameters. (b) Synthetic measurements generation using the DDA. (c) Generation of synthetic measurements for two RBCs: one RBC lying in the xy plane and the same RBC but randomly rotated. The pairs of the Rytov reconstructions and the ground truth RBCs are presented. The scale represents the normalized RI, which is calculated by dividing the RI values of a sample by the RI of the background. (d) Schematic description of the z-shift variant property of the Rytov measurement.

the calculations. Figure 2(c) shows two example pairs, one without rotation and the other with rotation.

For each RBC pair, we want to augment the dataset by shifting each example in all the axes. To do so, it is important to consider the shift properties of the Rytov reconstruction along each axis. We start from the integral solution to the Helmholtz equation:

$$U_s(\mathbf{r}) = \int_V F(\mathbf{r}')U(\mathbf{r}')G(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}', \qquad (1)$$

where $F(\mathbf{r}) = k^2/4\pi[n(\mathbf{r})^2/n_0^2 - 1]$ is the scattering potential of a sample whose RI distributions are $n(\mathbf{r})$ when immersed in a medium whose RI is n_0, given a wavenumber for the wavelength λ in vacuum, $k = 2\pi n_0/\lambda$. Here $G(\mathbf{r}-\mathbf{r}') = e^{1ik|\mathbf{r}-\mathbf{r}'|}/|\mathbf{r}-\mathbf{r}'|$ is the Green's function of the 3-D Helmholtz equation. The $U_i(\mathbf{r})$ and $U(\mathbf{r})$ are the incident and total electric fields, respectively, and the $U_s(\mathbf{r})$ is the scattered electric field, $U_s(\mathbf{r}) = U(\mathbf{r}) - U_i(\mathbf{r})$. The term U_s, on the left-hand side, can be measured at the image plane, as shown in Fig. 2(d). It is intuitive to see that moving the sample in the xy plane results in the same shift in the plane of the measurement of the scattered field. When the sample is translated in z, however, the measured scattered field will be the propagated version of the original unshifted measurement. Assuming that the sample is weakly scattering, the Rytov approximation uses the phase of the field itself, and Eq. (1) can be rewritten as follows:

$$U_{s-\mathrm{Rytov}}(\mathbf{r}) = U_i(\mathbf{r}) \log \frac{U(\mathbf{r})}{U_i(\mathbf{r})} = \int_V F(\mathbf{r}')U_i(\mathbf{r}')G(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}'. \qquad (2)$$

The left term of Eq. (1), $U(\mathbf{r}) - U_i(\mathbf{r}) = [e^{\log U(\mathbf{r})/U_i(\mathbf{r})} - 1]U_i(\mathbf{r})$, is replaced with the first Taylor expansion of it, $U_i(\mathbf{r}) \log U(\mathbf{r})/U_i(\mathbf{r})$. It, therefore, loses the propagation property of the scattered field. We refer this term as $U_{s-\mathrm{Rytov}}$. In other words, we must recalculate $U_{s-\mathrm{Rytov}}^{z-\mathrm{shift}}$ when an object is shifted in the z axis and the result of this calculation is different than distally propagating the field $U_{s-\mathrm{Rytov}}$.

Taking these properties into consideration, we augmented the set of training examples by generalizing shifted versions of the original pairs. For the shift in the xy plane, we added an xy-shifted version of each pair in addition to the original pair (without any shift). The shift was randomly selected during training. For the shift in the z axis, after generating the 40 projections for an RBC centered at $z = 0$, we digitally propagated the simulated measurements, U and U_i, to four different z planes ($-2\Delta z$, $-\Delta z$, $+\Delta z$, and $+2\Delta z$) and calculated the corresponding $U_{s-\mathrm{Rytov}}$ values at each plane. This was followed by their Rytov reconstructions to obtain examples of RBCs shifted along the z axis. In this work, Δz was set to 122 nm, which corresponds to one pixel of reconstruction grid. Rytov reconstructions were paired with the shifted RBCs in the z axis.

3 Method

3.1 Network Training

We trained a U-Net-type DNN in the regression manner using the following weighted l_2-norm as the cost function:[27,28]

$$\mathrm{Error}(\mathbf{x}_{\mathrm{recon}}, \mathbf{x}_{\mathrm{true}}) = \frac{\|\mathbf{x}_{\mathrm{recon}} - \mathbf{x}_{\mathrm{true}}\|_2^2}{\|\mathbf{x}_{\mathrm{true}}\|_2^2}, \qquad (3)$$

where $\mathbf{x}_{\mathrm{recon}}$ is the output from the network given $\mathbf{x}_{\mathrm{Rytov}}$, and $\mathbf{x}_{\mathrm{true}}$ is calculated from the ground truth RI contrast. Here, $\mathbf{x}$ represents the RI contrast multiplied by a scalar value, which is calculated as $c(\mathbf{n} - n_0)$, where $\mathbf{n}$ represents sample RI distributions and n_0 is the RI of medium. The scalar c was introduced for normalization of values; c can be either 40 for $\mathbf{x}_{\mathrm{Rytov}}$ or 20 for $\mathbf{x}_{\mathrm{true}}$. Negative components of input and output of the network were discarded. We implemented the network using PyTorch (1.2.0) and compute unified device architecture toolkit (10.0) on a desktop computer (Intel Core i7-6700 CPU, 3.4 GHz, 32 GB RAM) with a graphic processing unit (GPU, GeForce GTX 1070). The network was trained using the Adam optimizer with the learning rate of 1×10^{-4}, and it decayed half after every 10 epochs.[29] The mini-batch size was 8 and the total number of epoch was 50.

Figure 3 describes the network structure. It is very similar to the U-Net proposed in Refs. 27 and 28, except for slight modifications.[30,31] The input is skip-connected and summed to the output of the network. Therefore, the network learns the residual difference between the Rytov reconstruction and the ground truth.[28] All biases in the convolutional layers were set to zero and fixed. Zeros were padded for convolution layers of which kernel sizes are bigger than 1 so that the dimensions stay equal before and after the convolutions. The negative slope of leaky rectified linear unit (RELU) was set to 0.01. For the normalization layer, affine transform was turned off. For the transpose convolutional layers, the kernel size was set to $6 \times 6 \times 6$ with the zero padding of $2 \times 2 \times 2$ and the stride of $2 \times 2 \times 2$.

3.2 Experiment

The optical setup is described in Fig. 4.[32] It includes a diode pumped solid state 532 nm laser. The laser beam was first spatially filtered using a pinhole spatial filter. A beamsplitter was used to split the input beam into a sample beam and a reference beam. The sample beam was directed onto the sample at different angles of incidence using a reflective liquid crystal on silicon spatial light modulator (SLM) (Holoeye) with a pixel size and resolution of 1080×1920 pixels. Different illumination angles were obtained by projecting blazed gratings on the SLM. In the experiment presented here, a blazed grating with a period of 25 pixels was rotated a full 360 deg. Two $4F$ systems between the SLM and the sample permitted filtering of higher orders reflected from the SLM (due to limited fill factor and efficiency of the device) as well as magnification of the SLM projections onto the sample. Using a $100\times$ oil immersion objective lens with NA 1.4 (Olympus), the incident angle on the sample corresponding to the grating was 36 deg. The magnification of the illumination side was defined by the $4F$ systems we used before the sample. A third $4F$ system after the sample includes a $100\times$ oil immersion objective lens with NA 1.45 (Olympus). The sample and reference beams were collected on a second beamsplitter and projected onto a scientific CMOS (sCMOS) camera (Neo, Andor) with a pixel size and resolution of 2150×2650 pixels.

Blood sampling was performed by terminal intracardiac puncture on wild-type Balb/cByJ adult mice, in agreement with the Swiss legislation on animal experimentation (authorization number VD3290). RBCs were then isolated from blood plasma by centrifuging using Eppendorf-Centrifuge 5418 at 400 rpm
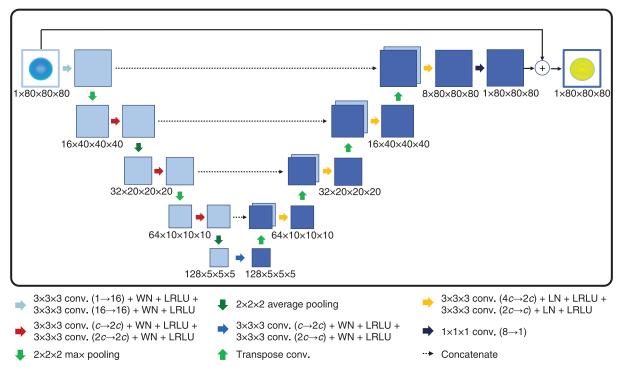
Fig. 3 Schematic description of the network structure. Here c represents the number of channels written at each block. WN, weight normalization; LRLU, leaky RELU; and LN, layer normalization.
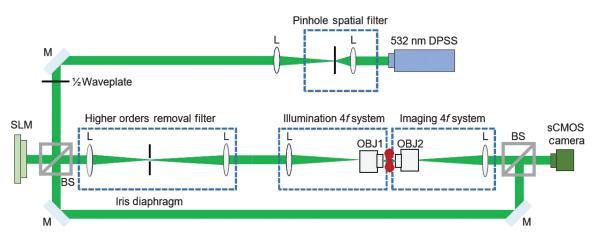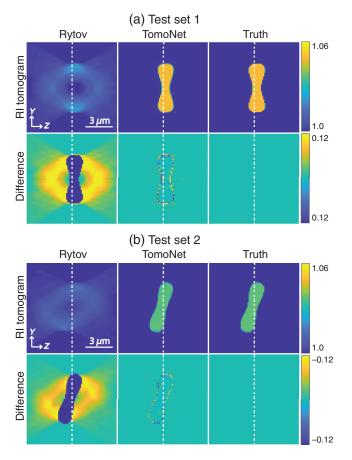

Fig. 4 Schematic for the experimental setup. M, mirror; L, lens; OBJ, objective lens; and BS, beamsplitter.

for 3 min. RBCs were then fixed using glutaraldehyde with concentration of 0.25% in phosphate-buffered saline (PBS) followed by centrifuging for 1 min and washing three times with PBS to remove any fixation reagents traces. To ensure strong adhesion between the RBC and the coverslip, coverslip was coated with 0.1% poly L-lysine diluted in PBS with molecular weight ranging between 1000 and 5000 gm/mol.

4 Results

4.1 Synthetic Data

Results obtained with the TomoNet are displayed in Fig. 5 for two different RBCs. Here, we only present centered RBCs without shifts in the xy plane. The first row shows RI reconstructions of Rytov, TomoNet, and the ground truth. The second row displays the difference map from the ground truth (reconstruction − the ground truth). In other words, blue regions in the difference map display underestimated parts and yellow regions show elongated regions. As expected, the Rytov reconstructions underestimate the RI values and elongate the RI distributions along the optical axis. Especially, the central dimple region of the RBC is significantly deteriorated. This is because the dimple region is thin and requires high frequencies for its reconstruction. By contrast, TomoNet shows excellent reconstruction results since it estimates accurately the values of these high frequencies from the data in the training set. In

(a) Test set 1

Rytov   TomoNet   Truth

(b) Test set 2

Rytov   TomoNet   Truth

Fig. 5 Reconstruction results using two examples from the test datasets. (a) Results for an RBC without rotation and (b) results for another RBC with rotation. The scale represents the normalized RI, which is calculated by dividing the RI values of a sample with the RI of background.

other words, the TomoNet implements super-resolution for 3-D samples revealing spatial details beyond the classical resolution limit. We quantitatively assessed the accuracy of the TomoNet over the Rytov reconstruction by calculating the following metric:
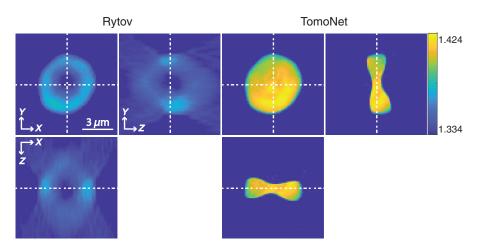
$$\text{error}(\Delta \boldsymbol{n}_{\text{recon}}, \Delta \boldsymbol{n}_{\text{true}}) = \frac{\|\Delta \boldsymbol{n}_{\text{recon}} - \Delta \boldsymbol{n}_{\text{true}}\|_2^2}{\|\Delta \boldsymbol{n}_{\text{true}}\|_2^2}, \quad (4)$$

where $\Delta \boldsymbol{n}_{\text{recon}}$ is the reconstructed RI contrast and $\Delta \boldsymbol{n}_{\text{true}}$ is the ground truth RI contrast. Here, $\Delta \boldsymbol{n}$ represents the RI contrast, which is defined as $(\boldsymbol{n} - n_0)$, where $\boldsymbol{n}$ represents the sample RI distributions and n_0 is the RI of medium. We discarded the negative values when calculating the error metric. The mean error values over all test RBCs are 0.5929 for the Rytov method and 0.0084 for the TomoNet, which confirms the improved performance of the network. The trained network accurately reconstructs RBCs and it does so in less than 10 ms on a GPU (GeForce GTX 1070).
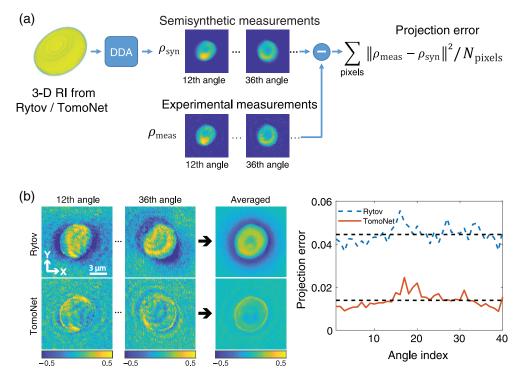
4.2 Experimental Data

We applied the network trained with digital phantoms to the Rytov reconstruction of a mouse RBC formed from experimental measurements. In the experiment, the samples were circularly scanned at the illumination angle of 36 deg in 9-deg steps resulting in 40 measurements, matching the parameters we used to generate the digital training data. As shown in Fig. 6, the Rytov reconstructions using the measurements show severe distortions, especially at the dimple region as we also observed in the synthetic data. With the Rytov reconstruction as its input, the TomoNet reconstructs RI tomograms without those artifacts resulting in the biconcave morphology. We verified the great improvement in the quality of the reconstructions visible in Fig. 6, by using a quantitative method[33] to evaluate the reconstruction accuracy of 3-D objects when the ground truth is not accessible. This was possible by generating semisynthetic measurements.

As described in Fig. 7(a), following the reconstruction of the RI distributions from the experimental measurements, we generated semisynthetic 2-D projections using an accurate forward model such as the DDA at each illumination angle. By comparing the digital projections with the corresponding 2-D experimental measurements, the difference between them reflects how close the 3-D reconstruction is to the ground truth. It is noteworthy that we did not use the forward model involved in the reconstruction to generate the digital projections to be fair.

Rytov   TomoNet

Fig. 6 Reconstruction of mouse RBC from experimental data using the network trained on synthetic data. The images to the left show the Rytov reconstruction, which is the input to the network. The images to the right show the results of the TomoNet.

Fig. 7 Validation of the experimental result using semisynthetic measurements. (a) Overall scheme of semisynthetic measurement generation using DDA. (b) Phase difference maps for two randomly selected angles and the average maps for all angles. The color bars are in radians. Calculation of projection errors in retrieved phase information from experimental and semisynthetic measurements.

Figure 7(a) shows two examples of phase maps from digital projections. For each digital projection, we calculated the projection error map, the difference in phase information between experimental and simulated measurements, along with the mean projection error map over all angles. Figure 7(b) displays two randomly selected projection error maps as well as the mean projection error maps for the Rytov and the TomoNet. In the case of Rytov, we can clearly see the mismatch between experimental and digital projections in the mean projection error map. By contrast, the mean projection error map of TomoNet shows excellent consistency. We further quantitatively confirmed the improvement in performance of TomoNet over Rytov by calculating the metric, $\sum_{l=1}^{L} \|\rho_{\exp}^{l} - \rho_{\text{syn}}^{l}\|^2 / N$, where L is the total number of angles, N is the total number of pixels, and $\rho_{\exp}$ and ρ_{syn} are the phase maps from experimental and semisynthetic measurements, respectively. As shown in Fig. 7(b), the average of the metric shows twofold improvement of TomoNet over the Rytov method.

5 Conclusion

We presented a DNN approach for reconstructing tomograms of RBCs with greatly improved image quality and super-resolution capability, especially enhancing the axial resolution. We digitally generated various RBCs and used them to generate synthetic measurements using the DDA to overcome the lack of the ground truth. The network trained on the synthetic data accurately reconstructs RI distributions of RBCs resolving the problems caused by the missing cone problem. We applied the trained network on experimental data to utilize extracted

features from the synthetic datasets. Despite the lack of the ground truth for the experimental result, we further validated the result of the network using semisynthetic measurements, and it confirmed the great improvement.

In this work, we focused on one specific cell type, RBCs, since it is relatively easy to model them. More importantly, RBCs are highly distorted by the missing cone problem, which prevents us from retrieving meaningful information for various applications. However, we believe that the proposed scheme can be further extended to other types of sample by carefully designing phantoms to statistically capture information in the generated dataset.

6 Appendix A: Dataset Generation

The shape of the surface of an RBC can be modeled by the following equation:

$$\rho^4 + 2S\rho^2 z^2 + z^4 + P\rho^2 + Qz^2 + R = 0, \quad (5)$$

where ρ is the radius in cylindrical coordinates ($\rho^2 = x^2 + y^2$) and S, P, Q, and R are the parameters derived from d, h, b, and c shown in Fig. 2(a).[24] To generate various RBCs, the d, h, and b values in microns were randomly selected from normal Gaussian distributions whose mean values were 7.65, 2.84, and 1.44 and standard deviations were 0.67, 0.46, and 0.47. c/d and the normalized RI values, $(n - n_0)/n_0$, were sampled from uniform distributions whose ranges were (0.56, 0.76) for c and (1.0355, 1.596) for the normalized RI.[24] To avoid nonrealistic shapes, several criteria were applied to limit the parameter

values in the following ranges: $h < 0.95 \times d/2$ and $b \leq h$. In addition, we limited the derived geometrical parameters such as cell volume (V μm^3), surface (S μm^2), and sphericity index (SI), $6\sqrt{\pi}V/S^{3/2}$, to fall within the normal ranges: $66 < V < 130$, $98 < S < 162$, and $0.494 < SI < 0.914$.[24] The cell surface was calculated using the equation, $\pi d[d/2 + 2h \times (\sinh^{-1} e)/e]$, where $e = 2\sqrt{9d^2 - 4h^2}/5h$.[34] Finally, RBC shapes were downsampled with a factor, 7.65/5.8, since the mean diameter of mouse RBCs is 5.8 μm, compared to the 7.65 μm for humans.[35] A total of 100 different RBCs were generated and each of them was randomly rotated in the yz plane (uniform distribution: $[0, \pi/6]$) and the xy plane (uniform distribution: $[0, 2\pi]$) resulting in 200 different RBCs (100 without rotation and 100 with rotation).

To generate synthetic measurements using the DDA, each RBC was illuminated at the incident angle of 36 deg for 40 angles, which were uniformly distributed on a circle. The illumination wavelength λ was 396 nm and the size of each dipole was set to $\lambda/12 = 33$ nm. The background medium in the simulation was air and the sample RI was set to the normalized RI. For 2-D measurements, the size of the grid was 256×256 with a pixel size of 99 nm. After that, the measurements were downsampled by cropping in k-space resulting in a pixel size of 122 nm. The original phantom defined using dipoles interpolated to a sampling grid that matched the pixel size of the measurements.

Measurements for 200 randomly generated RBCs were digitally refocused to five different planes resulting in 1000 pairs. Totally 800, 100, and 100 pairs were used for training, validation, and test, respectively. For the training and validation, we doubled the datasets by adding the randomly shifted sets on top of the original sets, resulting in 1600 and 200 pairs. The random shift varied at every iteration. For the rotated RBCs, after generating the measurements, we reversely rotated the Rytov reconstructions and the paired ground truth RBCs in the xy plane, resulting in rotations only in the yz plane to simplify the training. For the experimental data, since we do not know the rotation angle in the xy plane, we applied ellipsoidal fitting on a binary mask generated by applying Otsu thresholding[36] on the maximum projection map of Rytov. By analyzing the short and long axes, we extracted the orientation of the RBC. Since the Rytov reconstruction is shift-invariant in the xy plane, we simply interpolated in the xy plane for rotation.

7 Appendix B: Semisynthetic Simulation

The semisynthetic measurements were calculated using reconstruction results acquired from Rytov and TomoNet as samples for the DDA simulations. The pixel size of these reconstructions was 122 nm. Since the size of dipole was set to $\lambda/6n_0 = 67$ nm, where $\lambda = 532$ nm and $n_0 = 1.334$, the reconstructions were interpolated to a grid, one pixel of which was the size of a dipole. Then, we discretized the RI values as round($n/n_0 \times 1000$)/1000, and the negative values were discarded.

References

1. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Cambridge University Press, Cambridge, England (1999).
2. K. Lee et al., "Quantitative phase imaging techniques for the study of cell pathophysiology: from principles to applications," *Sensors* **13**(4), 4170–4191 (2013).
3. K. Kim et al., "Optical diffraction tomography techniques for the study of cell pathophysiology," *J. Biomed. Photonics Eng.* **2**(2), 020201 (2016).
4. D. Jin et al., "Tomographic phase microscopy: principles and applications in bioimaging," *J. Opt. Soc. Am. B* **34**(5), B64–B77 (2017).
5. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nat. Photonics* **12**(10), 578–589 (2018).
6. E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Opt. Commun.* **1**(4), 153–156 (1969).
7. J. Lim et al., "Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography," *Opt. Express* **23**(13), 16933–16948 (2015).
8. Y. Sung and R. R. Dasari, "Deterministic regularization of three-dimensional optical diffraction tomography," *J. Opt. Soc. Am. A* **28**(8), 1554–1561 (2011).
9. Y. Rivenson et al., "Deep learning microscopy," *Optica* **4**(11), 1437–1443 (2017).
10. A. Sinha et al., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
11. Y. Rivenson et al., "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Sci. Appl.* **7**(2), 17141 (2018).
12. Y. Rivenson et al., "Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning," *Nat. Biomed. Eng.* **3**(6), 466–477 (2019).
13. N. Borhani et al., "Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy," *Biomed. Opt. Express* **10**(3), 1339–1350 (2019).
14. Y. Jo et al., "Holographic deep learning for rapid optical screening of anthrax spores," *Sci. Adv.* **3**(8), e1700606 (2017).
15. J. Yoon et al., "Identification of non-activated lymphocytes using three-dimensional refractive index tomography and machine learning," *Sci. Rep.* **7**(1), 6654 (2017).
16. J. Lee et al., "Deep-learning-based label-free segmentation of cell nuclei in time-lapse refractive index tomograms," *IEEE Access* **7**, 83449–83460 (2019).
17. Y. Jo et al., "Quantitative phase imaging and artificial intelligence: a review," *IEEE J. Sel. Top. Quantum Electron.* **25**(1), 6800914 (2018).
18. J. Yoo et al., "Deep learning diffuse optical tomography," *IEEE Trans. Med. Imaging* (2019).
19. Y. Sun, Z. Xia, and U. S. Kamilov, "Efficient and accurate inversion of multiple scattering with deep learning," *Opt. Express* **26**(11), 14678–14688 (2018).
20. T. C. Nguyen, V. Bui, and G. Nehmetallah, "Computational optical tomography using 3-D deep convolutional neural networks," *Opt. Eng.* **57**(4), 043111 (2018).
21. A. Goy et al., "High-resolution limited-angle phase tomography of dense layered objects using deep neural networks," *Proc. Natl. Acad. Sci. U. S. A.* **116**(40), 19848–19856 (2019).
22. K. Kim et al., "High-resolution three-dimensional imaging of red blood cells parasitized by *Plasmodium falciparum* and *in situ* hemozoin crystals using optical diffraction tomography," *J. Biomed. Opt.* **19**(1), 011005 (2013).
23. Y. Kim et al., "Profiling individual human red blood cells using common-path diffraction optical tomography," *Sci. Rep.* **4**, 6659 (2014).

24. M. A. Yurkin et al., *Discrete Dipole Simulations of Light Scattering by Blood Cells*, Universiteit van Amsterdam [Host], Amsterdam (2007).

25. M. A. Yurkin and A. G. Hoekstra, "The discrete-dipole-approximation code ADDA: capabilities and known limitations," *J. Quant. Spectrosc. Radiat. Transfer* **112**(13), 2234–2247 (2011).

26. S. D'Agostino et al., "Enhanced fluorescence by metal nanospheres on metal substrates," *Opt. Lett.* **34**(15), 2381–2383 (2009).

27. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).

28. K. H. Jin et al., "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**(9), 4509–4522 (2017).

29. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).

30. T. Salimans and D. P. Kingma, "Weight normalization: a simple reparameterization to accelerate training of deep neural networks," in *Adv. Neural Inf. Process. Syst.*, pp. 901–909 (2016).

31. J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," arXiv:1607.06450 (2016).

32. A. B. Ayoub et al., "A method for assessing the fidelity of optical diffraction tomography reconstruction methods using structured illumination," *Opt. Commun.* **454**, 124486 (2020).

33. J. Lim et al., "High-fidelity optical diffraction tomography of multiple scattering samples," *Light: Sci. Appl.* **8**, 82 (2019).

34. I. Udroiu, "Estimation of erythrocyte surface area in mammals," arXiv:1403.7660 (2014).

35. K. Namdee et al., "Effect of variation in hemorheology between human and animal blood on the binding efficacy of vascular-targeted carriers," *Sci. Rep.* **5**, 11631 (2015).

36. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).

Joowon Lim received his BS and MS degrees in bio and brain engineering from Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2014 and 2016, respectively. He is pursuing his PhD in electrical engineering at the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His current research interests include model-based and deep-learning-based approaches for optical diffraction tomography reconstruction.

Ahmed B. Ayoub received his BS in electrical engineering from Alexandria University in Egypt in 2013. He received his MS in physics from the American University, Cairo, Egypt, in 2017. He is pursuing his PhD in electrical engineering at the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His current research interests include optical imaging and three-dimensional refractive index reconstructions for biological samples.

Demetri Psaltis received his BSc, MSc, and PhD from Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA. He is a professor of optics and the director of the Optics Laboratory at the École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. In 1980, he joined the faculty at the California Institute of Technology, Pasadena, California, USA. He moved to EPFL in 2006. His research interests include imaging, holography, biophotonics, nonlinear optics, and optofluidics. He has authored or coauthored over 400 publications in these areas. He is a fellow of the Optical Society of America, the European Optical Society, and SPIE. He was the recipient of the International Commission of Optics Prize, the Humboldt Award, the Leith Medal, and the Gabor Prize.