Supporting Information for

Deep learning-based holographic polarization microscopy

Tairan Liu[1,2,3†], Kevin de Haan[1,2,3†], Bijie Bai[1,2,3†], Yair Rivenson[1,2,3], Yi Luo[1,2,3], Hongda Wang[1,2,3], David Karalli[1], Hongxiang Fu[4], Yibo Zhang[1,2,3], John FitzGerald[5], and Aydogan Ozcan[1,2,3,6,*]

[1]Electrical and Computer Engineering Department, University of California, Los Angeles, CA 90095, USA
[2]Department of Bioengineering, University of California, Los Angeles, CA 90095, USA
[3]California NanoSystems Institute, University of California, Los Angeles, CA 90095, USA
[4]Computational and systems biology department, University of California, Los Angeles, CA 90095, USA
[5]Division of Rheumatology, Department of Internal Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA
[6]Department of Surgery, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA
[†]Equally contributing authors
*Corresponding author: Aydogan Ozcan
ozcan@ucla.edu

E-mail addresses:
Tairan Liu: liutr@ucla.edu
Kevin de Haan: kdehaan@ucla.edu
Bijie Bai: baibijie@ucla.edu
Yair Rivenson: rivensonyair@ucla.edu
Yi Luo: yluo2016@ucla.edu
Hongda Wang: hdwang@ucla.edu
David Karalli: davidkaralli@ucla.edu
Hongxiang Fu: hongxiang23@ucla.edu
Yibo Zhang: zybmax@ucla.edu
John Fitzgerald: JFitzgerald@mednet.ucla.edu
Aydogan Ozcan: ozcan@ucla.edu

Seven pages of supporting information, including two supplementary figures.

Holographic image reconstruction

Free space propagation. Free space wave propagation[1] is performed using the following equation:

$$\mathrm{FT}\{U(x,y;z)\} = \mathrm{FT}\{U(x,y;0)\} \cdot H(f_x, f_y; z) \tag{S1}$$

where $\mathrm{FT}\{\cdot\}$ is the 2D Fourier transform operation, $U(x, y; z)$ is the light field at an axial distance z, and $H(f_x, f_y; z)$ is defined as:

$$H(f_x, f_y; z) = \begin{cases} 0, & \left(\frac{\lambda f_x}{n}\right)^2 + \left(\frac{\lambda f_y}{n}\right)^2 > 1 \\ \exp\left[j2\pi \frac{n}{\lambda} z \sqrt{1 - \left(\frac{\lambda f_x}{n}\right)^2 - \left(\frac{\lambda f_y}{n}\right)^2} \right], & \text{o.w.} \end{cases} \tag{S2}$$

where f_x, f_y are the spatial frequencies along x and y directions, respectively, λ is the wavelength, and n is the refractive index of the medium.

Autofocusing. To accurately estimate the distance between the image sensor and the sample, an edge sparsity criterion (Tamura of the gradient (ToG)) based autofocusing is used[2]. First, we propagate the unfocused hologram to a series of z distances. At each distance, the ToG is computed, and the distance with the highest ToG is treated as the focusing distance. If the unfocused hologram is an amplitude only hologram (e.g. raw hologram captured using the image sensor), the propagation algorithm assumes zero phase by default.

Multi-height phase recovery. To eliminate twin-image related artifacts we used an iterative multi-height phase recovery algorithm[3–5]. Raw holograms from 8 sample to sensor distances with ~10 μm step size were collected. The hologram at the 1st height was padded with a zero-phase channel, and then propagated to the 2nd height. After the axial propagation, the amplitude channel values are averaged with the measured amplitude channel at the 2nd height, where the propagated phase channel was kept unchanged. This process is continued until it has been propagated to the 8th height, then back to the 1st height, which defines one iteration. The missing phase channel of a given hologram is typically recovered after 10-30 iterations.

Shift-and-add pixel super resolution. A super resolution algorithm[6] was used to improve the resolution of each hologram prior to the multi-height phase recovery step. Thirty-six low-resolution raw holograms at different lateral positions (6x6 with ~0.3733 μm spacing along the x and y directions) were collected at each sample-to-sensor distance to generate these high-resolution holograms. To accurately estimate the relative distance between the low-resolution holograms, a correlation-based method was used[7]. With the distance of these shifts being accurately obtained, the shift-and-add super resolution algorithm was used to fuse the low-resolution holograms into a high-resolution corresponding hologram.

SCPLM mathematical model derivation

The design of the single-shot computational polarized light microscopy (SCPLM) setup[8] is

demonstrated in the main text Figure 7(c). A birefringent sample of interest is illuminated using left-hand circularly polarized (LHCP) light, which is obtained by placing a left-hand circular polarizer behind an LED illumination source. The transmitted light is recorded using a pixelated polarizer camera. As shown in Figure 7(a), every four pixels of the camera sensor utilize four different directional polarizing filters (0°, 90°, 45°, and 135°) and form a repeated pattern across the sensor.

The mathematical model of the SCPLM can be formulated by Jones calculus, where the optical components can be expressed as:

I. Input left-hand circularly polarized (LHCP) light:

$$\mathbf{E}_{\text{in}} = \frac{E_0}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix} \tag{S3}$$

where $i^2 = -1$, E_0 is the amplitude of the incident LHCP light, and LHCP is defined from the point of view of the receiver.

II. Birefringent sample:

$$\mathbf{M}_{\text{sample}} = \begin{bmatrix}\cos\frac{\eta}{2} + i\sin\frac{\eta}{2}\cos 2\theta & i\sin\frac{\eta}{2}\sin 2\theta\\ i\sin\frac{\eta}{2}\sin 2\theta & \cos\frac{\eta}{2} - i\sin\frac{\eta}{2}\cos 2\theta\end{bmatrix} \tag{S4}$$

where η and θ represent the retardance magnitude and the slow axis orientation of the birefringent object. Both η and θ are spatially varying.

III. Detector on the pixelated image sensor:

$$\mathbf{M}_{\text{detector}} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix},\ \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix},\ \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix},\ \text{or}\ \frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix} \tag{S5}$$

for four different directional polarizing filters at 0°, 45°, 90°, and 135°, respectively.

The output light field can be expressed as:

$$\mathbf{E}_{\text{out}} = \mathbf{M}_{\text{detector}}\mathbf{M}_{\text{sample}}\mathbf{E}_{\text{LHCP}} \tag{S6}$$

Consider a pixel with 0°-orientated polarizing filter, the corresponding light field can be formulated as:

$$\mathbf{E}^0_{\text{out}} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\cdot\begin{bmatrix}\cos\frac{\eta}{2} + i\sin\frac{\eta}{2}\cos 2\theta & i\sin\frac{\eta}{2}\sin 2\theta\\ i\sin\frac{\eta}{2}\sin 2\theta & \cos\frac{\eta}{2} - i\sin\frac{\eta}{2}\cos 2\theta\end{bmatrix}\cdot\frac{E_0}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix} \tag{S7}$$

Then the intensity value recorded by this 0°-orientated pixel can be written as:

$$I_0 \propto \left|\mathbf{E}^0_{\text{out}}\right|^2 = \mathbf{E}^0_{\text{out}}\cdot\mathbf{E}^{0\ *}_{\text{out}} \tag{S8}$$

Combining equation S7 and S8, we get:

$$I_0 = \frac{1}{2}I_{\max}(1 - \sin\eta\sin 2\theta) \tag{S9}$$

where $I_{\max}$ is a constant related to the intensity of the illumination light. Following the same

steps, the intensity values captured by other three different orientated pixels of the image sensor chip can be expressed as:

$$\begin{aligned} I_{45} &= \frac{1}{2} I_{\max}(1+\sin\eta\cos 2\theta), \\ I_{90} &= \frac{1}{2} I_{\max}(1+\sin\eta\sin 2\theta), \\ I_{135} &= \frac{1}{2} I_{\max}(1-\sin\eta\cos 2\theta) \end{aligned} \tag{S10}$$

Gathering equation S9 and S10, two auxiliary variables A and B can be defined as:

$$\begin{aligned} A &= \frac{I_{90}-I_0}{I_{90}+I_0} = \sin\eta\sin 2\theta \\ B &= \frac{I_{45}-I_{135}}{I_{45}+I_{135}} = \sin\eta\cos 2\theta \end{aligned} \tag{S11}$$

Finally, the retardance magnitude and the slow axis orientation of the birefringent object can be reconstructed as:

$$\begin{aligned} \eta &= \arcsin\sqrt{A^2+B^2} \\ \theta &= \frac{1}{2}\arctan\frac{A}{B} \end{aligned} \tag{S12}$$

DL-HPM mathematical model

The mathematical model of the deep learning-based holographic polarization microscope (DL-HPM) which is described by equation 7 in the main text can be expanded as:

$$\begin{aligned} \mathbf{E}_{\text{out}} = &\begin{bmatrix} \cos^2\beta & \cos\beta\sin\beta \\ \cos\beta\sin\beta & \sin^2\beta \end{bmatrix} e^{-\frac{i\pi}{4}} \begin{bmatrix} \cos^2\alpha + i\sin^2\alpha & (1-i)\cos\alpha\sin\alpha \\ (1-i)\cos\alpha\sin\alpha & \sin^2\alpha + i\cos^2\alpha \end{bmatrix} \\ &e^{-\frac{i\eta}{2}} \begin{bmatrix} \cos^2\theta + e^{i\eta}\sin^2\theta & (1-e^{i\eta})\cos\theta\sin\theta \\ (1-e^{i\eta})\cos\theta\sin\theta & \sin^2\theta + e^{i\eta}\cos^2\theta \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \end{aligned} \tag{S13}$$

where the illumination light is normalized to have an amplitude of 1, LHCP is defined from the point of view of the source, α represents the orientation of the fast axis of the $\lambda/4$ retarder with respect to the x-axis, β represents the orientation of the linear polarizer with respect to the x-axis, θ represents the orientation of the fast axis of the sample with respect to the x-axis and η represents the relative phase retardance.

In our optical design, we selected α to be 90° without a loss of generality, and therefore we have:

$$\begin{aligned} \mathbf{E}_{\text{out}} = &\begin{bmatrix} \cos^2\beta & \cos\beta\sin\beta \\ \cos\beta\sin\beta & \sin^2\beta \end{bmatrix} e^{-\frac{i\pi}{4}} \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix} e^{-\frac{i\eta}{2}} \cdot \\ &\begin{bmatrix} \cos^2\theta + e^{i\eta}\sin^2\theta & (1-e^{i\eta})\cos\theta\sin\theta \\ (1-e^{i\eta})\cos\theta\sin\theta & \sin^2\theta + e^{i\eta}\cos^2\theta \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \end{aligned} \tag{S14}$$

Then, using the terms before the linear polarizer, we have:

$$\mathbf{E}_{\text{before\_linear}} = \frac{1}{\sqrt{2}} e^{-\frac{i\pi}{4}} e^{-\frac{i\eta}{2}} \begin{bmatrix} i\cos^2\theta + ie^{i\eta}\sin^2\theta + \left(1-e^{i\eta}\right)\cos\theta\sin\theta \\ \left(1-e^{i\eta}\right)\cos\theta\sin\theta - i\sin^2\theta - ie^{i\eta}\cos^2\theta \end{bmatrix} \tag{S15}$$

To further simplify the equation, we use Euler's formula:

$$\mathbf{E}_{\text{before\_linear}} = \frac{1}{\sqrt{2}} e^{-\frac{i\pi}{4}} e^{-\frac{i\eta}{2}} E_{\text{before\_linear\_x}} \cdot \mathbf{x} + \frac{1}{\sqrt{2}} e^{-\frac{i\pi}{4}} e^{-\frac{i\eta}{2}} E_{\text{before\_linear\_y}} \cdot \mathbf{y} \tag{S16}$$

where **x** and **y** are unit vectors pointing toward the positive *x* and *y* axis directions, respectively, and $E_{\text{before\_linear\_x}}$ and $E_{\text{before\_linear\_y}}$ are the *x* and *y* components of $\mathbf{E}_{\text{before\_linear}}$, defined as:

$$\begin{cases} E_{\text{before\_linear\_x}} = i\cos^2\theta + \left(i\cos\eta - \sin\eta\right)\sin^2\theta + \left(1-\cos\eta - i\sin\eta\right)\cos\theta\sin\theta \\ E_{\text{before\_linear\_y}} = \left(1-\cos\eta - i\sin\eta\right)\cos\theta\sin\theta - i\sin^2\theta - \left(i\cos\eta - \sin\eta\right)\cos^2\theta \end{cases} \tag{S17}$$

Since the final polarizer is a linear polarizer, the output field can be written as:

$$\mathbf{E}_{\text{out}} = \frac{1}{\sqrt{2}} e^{-\frac{i\pi}{4}} e^{-\frac{i\eta}{2}} \left(\cos\beta E_{\text{before\_linear\_x}} + \sin\beta E_{\text{before\_linear\_y}}\right) \cdot \mathbf{k} \tag{S18}$$

where **k** is a unit vector pointing toward the direction with β degrees with respect to the positive x-axis. Based on equations S17-S18, we get:

$$\mathbf{E}_{\text{out}} = \frac{1}{\sqrt{2}} e^{-\frac{i\pi}{4}} e^{-\frac{i\eta}{2}} \left(a + ib\right) \cdot \mathbf{k} \tag{S19}$$

where *a* and *b* are both real and functions of θ and η that are defined by:

$$\begin{cases} a = -\cos\beta\sin\eta\sin^2\theta + \sin\beta\sin\eta\cos^2\theta \\ \quad + \cos\beta\cos\theta\sin\theta + \sin\beta\cos\theta\sin\theta \\ \quad - \cos\beta\cos\eta\cos\theta\sin\theta - \sin\beta\cos\eta\cos\theta\sin\theta \\ b = \cos\beta\cos^2\theta - \sin\beta\sin^2\theta \\ \quad + \cos\beta\cos\eta\sin^2\theta - \sin\beta\cos\eta\cos^2\theta \\ \quad - \cos\beta\sin\eta\cos\theta\sin\theta - \sin\beta\sin\eta\cos\theta\sin\theta \end{cases} \tag{S20}$$

Then, the amplitude and phase of the output field can be written as:

$$\begin{cases} A_{\text{out}} = \dfrac{\sqrt{2}}{2}\sqrt{a^2 + b^2} \\ \varphi_{\text{out}} = \text{atan2}\left(b, a\right) - \dfrac{\eta}{2} - \dfrac{\pi}{4} \end{cases} \tag{S21}$$

where atan2(*y*, *x*) is the four-quadrant inverse tangent function for point (*x*, *y*).

To relate these results from equation S21 to the amplitude and phase channels obtained by the holographic reconstruction method, a background normalization procedure is required, where the background region is normalized to have a unit amplitude and zero phase. For the background region, the output field can be written as:

$$\begin{aligned}\mathbf{E}_{\mathrm{bg}} &= \begin{bmatrix} \cos^2\beta & \cos\beta\sin\beta \\ \cos\beta\sin\beta & \sin^2\beta \end{bmatrix} e^{-\frac{i\pi}{4}} \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \\ &= \frac{1}{\sqrt{2}}\left(\cos\beta - \sin\beta\right) e^{\frac{i\pi}{4}} \cdot \mathbf{k}\end{aligned} \tag{S22}$$

Thus, we have:

$$\begin{cases} A_{\mathrm{bg}} = \dfrac{\sqrt{2}}{2}\left|\cos\beta - \sin\beta\right| \\ \varphi_{\mathrm{bg}} = \operatorname{atan2}\left[0, \cos\beta - \sin\beta\right] + \dfrac{\pi}{4} \end{cases} \tag{S23}$$

Hence, the reconstructed amplitude and phase can be expressed as:

$$\begin{cases} A_{\mathrm{recon}} = \dfrac{A_{\mathrm{out}}}{A_{\mathrm{bg}}} = \dfrac{\sqrt{a^2 + b^2}}{\left|\cos\beta - \sin\beta\right|} \\ \varphi_{\mathrm{recon}} = \varphi_{\mathrm{out}} - \varphi_{\mathrm{bg}} = \operatorname{atan2}\left(b, a\right) - \dfrac{\eta}{2} - \dfrac{\pi}{2} - \operatorname{atan2}\left[0, \cos\beta - \sin\beta\right] \end{cases} \tag{S24}$$

Sources of error in DL-HPM

An important source of error for DL-HPM comes from inaccuracies in the image co-registration step, which negatively affects both the quantitative analysis as well as the training of the network. As was demonstrated in Figure 1 of main text, two different image modalities (running on different set-ups) were used to capture the images, which constituted the training and testing datasets. Therefore, an accurate co-registration procedure is required to align the images resulting from the two-imaging systems. However, the registration algorithms are not always completely effective, leaving a relatively large registration error for e.g., MSU crystals. In Figure 4 of main text, an additional local registration step aligning the individual crystals was applied to help account for such co-registration errors. Furthermore, co-registration errors between the input and target images during the training phase also cause the loss function to be applied incorrectly, which relatively reduces the performance of the trained network.

Another cause of error is related to larger crystals, as they are more likely to be formed by multiple MSU crystals which are stacked or clustered next to one another. This can lead to a more complex relationship between the amplitude/phase channel and the retardance/orientation channel, as the input phase channel may start to present phase wrapping effects. Since such larger crystals are rarely observed in the training dataset, the ability of the network to accurately infer the retardance and orientation channels corresponding to these clustered objects is generally reduced.

In addition to these, various other aspects of the experimental set-up can also cause errors. Throughout the image acquisition procedures for training and testing phases, each sample slide was aligned after being moved between the two different imaging systems. This physical alignment procedure is used to ensure that the sample is aligned at the same angle relative to the polarizer-analyzer pair in each imaging system, and this alignment process can inevitably

cause experimental errors.

MSU dataset

The MSU dataset used for training and testing of DL-HPM contains 8 MSU slides obtained from different de-identified patients with different sample conditions in terms of the crystal density and dominant cell types. Figure S1 shows an example field of view for each of these slides obtained using the SCPLM method.

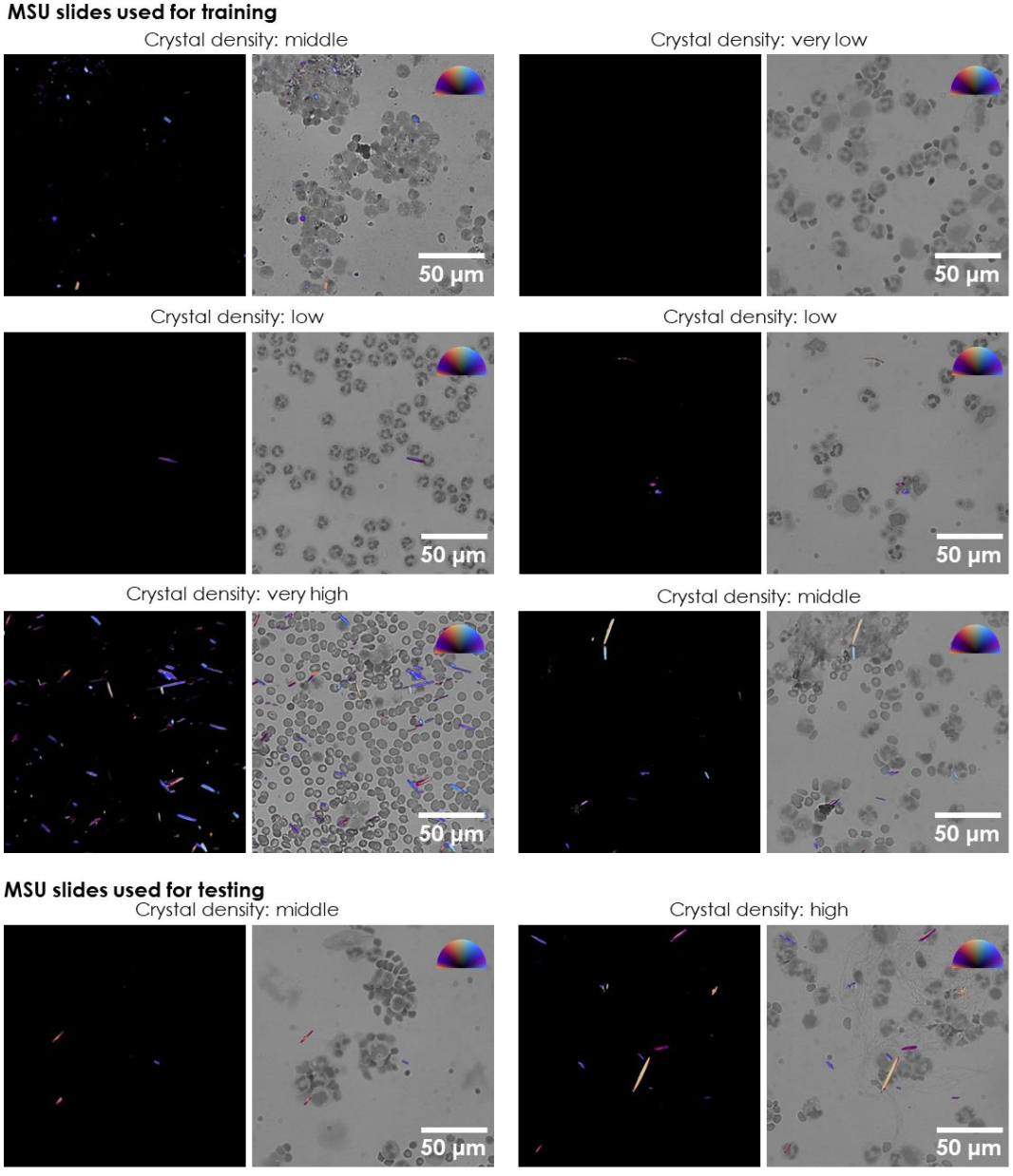

Figure S1. **Example fields of view of each MSU slide used for training and testing of DL-HPM.** Color bar: from left to top to right, respectively, represents π, π/2, 0 Rad in the orientation channel. Retardance is represented by the distance from the center of the color bar, ranging from 0 to 0.4.

Impact of using different loss functions on network training

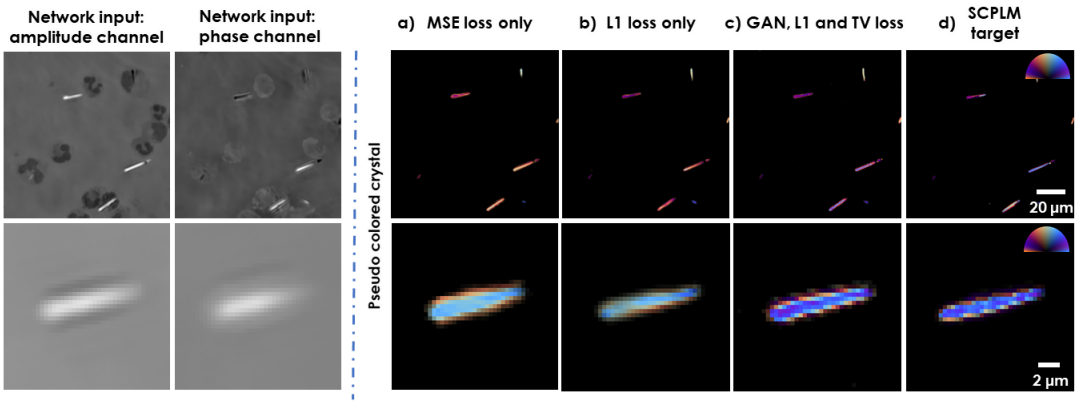

Figure S2. **Network output using different loss functions during the training.** a) Loss function used mean squared error (MSE) loss only. b) Loss function used L1 loss only. c) Loss function used in the main text – i.e., a combination of GAN, L1 and TV losses (see the Methods section of the main text). d) Ground truth imaged by SCPLM. Color bar: from left to top to right, respectively, represents π, $\pi/2$, 0 Rad in the orientation channel. Retardance is represented by the distance from the center of the color bar, ranging from 0 to 0.4.

References

(1) Goodman, J. W. *Introduction to Fourier Optics*; Roberts and Company Publishers, 2005.

(2) Zhang, Y.; Wang, H.; Wu, Y.; Tamamitsu, M.; Ozcan, A. Edge Sparsity Criterion for Robust Holographic Autofocusing. *Opt Lett* **2017**, *42* (19), 3824–3827.

(3) Rivenson, Y.; Wu, Y.; Wang, H.; Zhang, Y.; Feizi, A.; Ozcan, A. Sparsity-Based Multi-Height Phase Recovery in Holographic Microscopy. *Scientific Reports* **2016**, *6*, 37862. https://doi.org/10.1038/srep37862.

(4) Greenbaum, A.; Zhang, Y.; Feizi, A.; Chung, P.-L.; Luo, W.; Kandukuri, S. R.; Ozcan, A. Wide-Field Computational Imaging of Pathology Slides Using Lens-Free on-Chip Microscopy. *Science translational medicine* **2014**, *6* (267), 267ra175–267ra175.

(5) Greenbaum, A.; Ozcan, A. Maskless Imaging of Dense Samples Using Pixel Super-Resolution Based Multi-Height Lensfree on-Chip Microscopy. *Opt. Express, OE* **2012**, *20* (3), 3129–3143. https://doi.org/10.1364/OE.20.003129.

(6) Luo, W.; Zhang, Y.; Feizi, A.; Göröcs, Z.; Ozcan, A. Pixel Super-Resolution Using Wavelength Scanning. *Light Sci Appl* **2016**, *5* (4), e16060. https://doi.org/10.1038/lsa.2016.60.

(7) Greenbaum, A.; Sikora, U.; Ozcan, A. Field-Portable Wide-Field Microscopy of Dense Samples Using Multi-Height Pixel Super-Resolution Based Lensfree Imaging. *Lab Chip* **2012**, *12* (7), 1242–1245. https://doi.org/10.1039/C2LC21072J.

(8) Bai, B.; Wang, H.; Liu, T.; Rivenson, Y.; FitzGerald, J.; Ozcan, A. Pathological Crystal Imaging with Single-Shot Computational Polarized Light Microscopy. *Journal of biophotonics* **2020**, *13* (1), e201960036.